ICES REPORT 17-28

October 2017

Direct Serendipity Finite Elements on Convex Quadrilaterals

by

Todd Arbogast and Zhen Tao

The Institute for Computational Engineering and Sciences
The University of Texas at Austin
Austin, Texas 78712

Reference: Todd Arbogast and Zhen Tao, "Direct Serendipity Finite Elements on Convex Quadrilaterals," ICES REPORT 17-28, The Institute for Computational Engineering and Sciences, The University of Texas at Austin, October 2017.

Direct Serendipity Finite Elements on Convex Quadrilaterals

Todd Arbogast[a,1], Zhen Tao[b,1]

[a]*University of Texas at Austin; Institute for Computational Engineering and Sciences; 201 EAST 24th St., C0200; Austin, TX 78712–1229 and Department of Mathematics;2515 Speedway, C1200; Austin, TX 78712–1202; U.S.A.*
[b]*Institute for Computational Engineering and Sciences; 201 EAST 24th St., C0200; Austin, TX 78712–1229*

Abstract

The classical serendipity finite element spaces suffer from poor approximation on nondegenerate, convex quadrilaterals. In this paper, we develop the *direct* serendipity spaces $\mathcal{DS}_r$, a new family of finite elements for $r \geq 2$ that has the same number of degrees of freedom as the classical space but maintains optimal approximation properties. The set of local shape functions for $\mathcal{DS}_r$ contains the full set of polynomials of degree r defined directly on each element. Because there are not enough degrees of freedom, exactly two supplemental rational functions are added to each element.

Keywords: serendipity, finite element, quadrilateral, optimal convergence
2010 MSC: 65N30, 65M60, 65N12, 65M12, 65D05

1. Introduction

The serendipity finite elements on rectangles, especially the 8-node biquadratic and the 12-node bicubic ones, have been well studied for many years. They appear in almost any introductory reference on finite elements, e.g., [1, 2, 3], and they are provided by software packages both in academia [4] and industry [5]. The two dimensional serendipity elements appear in the periodic table of the finite elements of Arnold and Logg [6], where they are denoted as $\mathcal{S}_r\Lambda^0$ (they form the precursor to the Brezzi-Douglas-Marini spaces ($\mathcal{S}_r\Lambda^1$) [7]). Compared with the full tensor product Lagrange finite elements, serendipity finite elements use fewer degrees of freedom, and they are usually more efficient. It was not until recently, however, that a general definition of the serendipity finite element spaces of arbitrary order on rectangles in any space dimension was given by Arnold and Awanou [8, 9].

The serendipity finite element spaces, denoted $\mathcal{S}_r$ of index $r \geq 1$, work very well on computational meshes of rectangular element domains, but it is well known that their performance is degraded by bilinear distortion when the space is third order accurate and above, i.e., $r \geq 2$. This is not the case for tensor product Lagrange finite elements [10, 11, 12]. To be more precise, let the element domain E be a nondegenerate, convex quadrilateral. We can view it as the image

Email addresses: arbogast@ices.utexas.edu (Todd Arbogast), taozhen@ices.utexas.edu (Zhen Tao)
[1]This work was supported by the U.S. National Science Foundation under grant DMS-1418752.

of the unit square $\hat{E}$ under the bilinear map $\mathbf{F}_E : \hat{E} \to E$. Mapped serendipity elements do not approximate to optimal order $r+1$ on E, but the image of the full space of tensor product polynomials maintains accuracy on E. Rand, Gillette, and Bajaj [13] recently introduced a new family of Serendipity finite elements based on generalized barycentric coordinates of index $r = 2$ that is accurate to order 3 on any convex, planar polygon.

In this paper, we introduce a new family of finite element spaces, each of which has the same number of degrees of freedom as the corresponding classical serendipity finite element space but also works well on general non-degenerate convex quadrilaterals. We call these new elements *direct* serendipity finite elements, and denote them by $\mathcal{DS}_r$, $r \geq 2$. They are *direct* in the sense that the shape functions contain a full set of polynomials (not tensor product polynomials) defined directly on the element. Because there are not enough degrees of freedom, two supplemental rational functions need to be added to each element, much like as is done for the mixed finite element spaces af Arbogast and Correa [14].

We construct the new family of direct serendipity finite elements of any order $r \geq 2$ in the next section. In Section 3, we discuss the H^1-conforming implementation of the new elements. We make some remarks on approximation and stability of the new spaces in Section 4. We close with some numerical results comparing the performance of the mapped serendipity, the new direct serendipity, and the mapped tensor product Lagrange spaces in the last section.

2. Construction of the Direct Serendipity Elements

Let $\mathbb{P}_r(\Omega)$ denote the space of polynomials of degree up to r on $\Omega \subset \mathbb{R}$ or $\mathbb{R}^2$, and let $\mathbb{P}_{r,r}$ denote the tensor product space of polynomials on $\mathbb{R}^2$ of degree up to r in each coordinate variable.

Let the reference element $\hat{E}$ be $[0,1]^2$ and E be a closed, convex quadrilateral which does not degenerate to a triangle or a line. Define the bilinear and bijective map $\mathbf{F}_E : \hat{E} \to E$ as

$$\mathbf{F}_E(\hat{\mathbf{x}}) = \mathbf{F}_E(\hat{x},\hat{y}) = \mathbf{x}_c^1(1-\hat{x})(1-\hat{y}) + \mathbf{x}_c^2\hat{x}(1-\hat{y}) + \mathbf{x}_c^3\hat{x}\hat{y} + \mathbf{x}_c^4(1-\hat{x})\hat{y} \in \mathbb{P}_{1,1}, \tag{1}$$

where $\mathbf{x}_c^i$, $i = 1,2,3,4$, are the four corner vertices of the quadrilateral (see Figure 1). The reference element $\hat{E}$ and the physical element E are labeled counterclockwise, and for $i = 1,2,3,4$, they have edges $\hat{e}_i$ and e_i with outer unit normals $\hat{\nu}_i$ and ν_i, respectively, where e_i connects $\mathbf{x}_c^{i-1}$ to $\mathbf{x}_c^i$ (interpret $\mathbf{x}_c^0 = \mathbf{x}_c^4$).

2.1. Auxiliary functions

We define the linear polynomials giving the distance of any point $\mathbf{x}$ to each edge e_i, $i = 1,2,3,4$, in the normal direction as

$$\lambda_i(\mathbf{x}) = -(\mathbf{x} - \mathbf{x}_c^i)\cdot\nu_i, \quad i = 1,2,3,4. \tag{2}$$

If $\mathbf{x}$ is in the interior $\mathring{E}$ of E, the barycentric coordinates are strictly positive, and each vanishes on the edge which defines it.

Next we define the directional polynomials

$$\lambda_V(\mathbf{x}) = \lambda_1(\mathbf{x}) - \lambda_3(\mathbf{x}), \tag{3}$$

$$\lambda_H(\mathbf{x}) = \lambda_2(\mathbf{x}) - \lambda_4(\mathbf{x}), \tag{4}$$

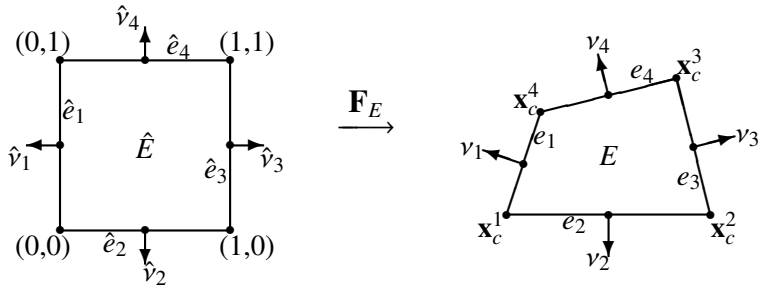


Figure 1: A reference element $\hat{E} = [0,1]^2$ and counterclockwise oriented convex quadrilateral E, with edges $\hat{e}_i$ and e_i and outer unit normals $\hat{\nu}_i$ and ν_i, respectively. Corners for the physical element E are labeled as $\mathbf{x}_c^i$, for $i = 1,2,3,4$.

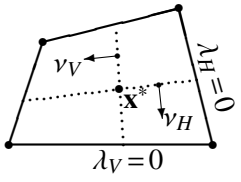


Figure 2: Illustration of lines $\lambda_H = 0$, $\lambda_V = 0$, point $\mathbf{x}^*$, and unit normal vectors ν_V and ν_H.

which are linear functions vanishing on a "vertical" or "horizontal" line, respectively, as shown in Figure 2. It is worth remarking that they do not necessary vanish at the mid-lines connecting mid-points of the opposite edges. Obviously, $\lambda_V(\mathbf{x})$ is negative on e_1 and positive on e_3, while $\lambda_H(\mathbf{x})$ is negative on e_2 and positive on e_4. Therefore, there exists a point $\mathbf{x}^* \in E$ such that $\lambda_V(\mathbf{x}^*) = \lambda_H(\mathbf{x}^*) = 0$. Combined with the convexity of the quadrilateral E, the polynomials can be rewritten as

$$\lambda_V(\mathbf{x}) = -\|\nu_1 - \nu_3\|(\mathbf{x} - \mathbf{x}^*) \cdot \nu_V, \tag{5}$$

$$\lambda_H(\mathbf{x}) = -\|\nu_2 - \nu_4\|(\mathbf{x} - \mathbf{x}^*) \cdot \nu_H, \tag{6}$$

where $\nu_V = \dfrac{\nu_1 - \nu_3}{\|\nu_1 - \nu_3\|}$ and $\nu_H = \dfrac{\nu_2 - \nu_4}{\|\nu_2 - \nu_4\|}$ are unit vectors.

The directional rational functions are defined as

$$R_V(\mathbf{x}) = \frac{\lambda_1(\mathbf{x}) - \lambda_3(\mathbf{x})}{\xi_V \lambda_1(\mathbf{x}) + \eta_V \lambda_3(\mathbf{x})} = \frac{\lambda_V(\mathbf{x})}{\xi_V \lambda_1(\mathbf{x}) + \eta_V \lambda_3(\mathbf{x})}, \tag{7}$$

$$R_H(\mathbf{x}) = \frac{\lambda_2(\mathbf{x}) - \lambda_4(\mathbf{x})}{\xi_H \lambda_2(\mathbf{x}) + \eta_H \lambda_4(\mathbf{x})} = \frac{\lambda_H(\mathbf{x})}{\xi_H \lambda_2(\mathbf{x}) + \eta_H \lambda_4(\mathbf{x})}, \tag{8}$$

where

$$\xi_V^{-1} = \sqrt{1 - (\nu_H \cdot \nu_1)^2} \quad \text{and} \quad \eta_V^{-1} = \sqrt{1 - (\nu_H \cdot \nu_3)^2}, \tag{9}$$

$$\xi_H^{-1} = \sqrt{1 - (\nu_V \cdot \nu_2)^2} \quad \text{and} \quad \eta_H^{-1} = \sqrt{1 - (\nu_V \cdot \nu_4)^2}. \tag{10}$$

Since E is convex and does not degenerate, the coefficients ξ_V, η_V, ξ_H, and η_H are constants greater than 1. Since $\lambda_1(\mathbf{x})\lambda_3(\mathbf{x}) \neq 0$ and $\lambda_2(\mathbf{x})\lambda_4(\mathbf{x}) \neq 0$ for any point $\mathbf{x} \in \mathring{E}$, the functions $R_V(\mathbf{x})$ and $R_H(\mathbf{x})$ are well defined on E and have linear denominators. Notice that

$$R_V(\mathbf{x})|_{e_1} = -\eta_V^{-1} \quad \text{and} \quad R_V(\mathbf{x})|_{e_3} = \xi_V^{-1}, \tag{11}$$

$$R_H(\mathbf{x})|_{e_2} = -\eta_H^{-1} \quad \text{and} \quad R_H(\mathbf{x})|_{e_4} = \xi_H^{-1}. \tag{12}$$

All these auxiliary functions can be calculated easily once the bilinear mapping $\mathbf{F}_E$ is provided.

2.2. Construction of shape functions

It is shown in [12] that the convergence of the linear serendipity finite element space ($r = 1$) does not degenerate on non-rectangular geometries. In fact, the parametric serendipity element $\mathcal{S}_1(E)$ and the tensor product space $\mathbb{P}_{1,1}$ on $\hat{E}$ mapped to E are identical. Therefore, we only develop our new direct serendipity finite elements for indices $r \geq 2$.

Recall Ciarlet's definition [2] of a finite element.

Definition 2.1 (Ciarlet 1978)**.** *Let*

1. $K \subset \mathbb{R}^d$ *be a bounded closed set with nonempty interior and a Lipschitz continuous boundary,*
2. $\mathcal{P}$ *be a finite-dimensional space of functions on* K*, and*
3. $\mathcal{N} = \{N_1, N_2, \ldots, N_k\}$ *be a basis for* $\mathcal{P}'$*.*

Then $(K, \mathcal{P}, \mathcal{N})$ *is called a* finite element.

A convex quadrilateral E satisfies the requirements of Property 1 for $d = 2$. Following [8], we define the shape functions (a basis for $\mathcal{P}$ in Definition 2.1) of the direct serendipity element based on the geometric decomposition of quadrilaterals (see Table 1). For a quadrilateral in two dimensional space $\mathbb{R}^2$, the geometric objects are vertices, edges, and the interior domain, with dimension $d = 0, 1, 2$, respectively. There are $2^{2-d}\binom{2}{d}$ objects for each dimension d. Each object has $\binom{r-d}{d}$ degrees of freedom (DOFs) associated to it when $r - 2d \geq 0$, and otherwise there are no DOFs associated with the object. We remark that there are no interior DOFs for $r = 2, 3$.

Table 1: A summary of the geometric decomposition and the degrees of freedom (DOFs) associated to each geometric object of a quadrilateral for the serendipity element of index $r \geq 2$.

dimension	object name	object count	total DOFs	DOFs per object
0	vertex	4	4	1
1	edge	4	$4(r-1)$	$r-1$
2	interior	1	$\max(0, \frac{1}{2}(r-2)(r-3))$	$\max(0, \frac{1}{2}(r-2)(r-3))$

Shape functions associated with the vertices are

$$\begin{aligned} \phi_1(\mathbf{x}) &= \lambda_3(\mathbf{x})\lambda_4(\mathbf{x}), \quad \phi_2(\mathbf{x}) = \lambda_1(\mathbf{x})\lambda_4(\mathbf{x}), \\ \phi_3(\mathbf{x}) &= \lambda_1(\mathbf{x})\lambda_2(\mathbf{x}), \quad \phi_4(\mathbf{x}) = \lambda_2(\mathbf{x})\lambda_3(\mathbf{x}). \end{aligned} \tag{13}$$

The vertex shape functions are second order polynomials.

Next we define shape functions associated with edges e_1 and e_3. There are $2(r-1)$ of them, and they are

$$\phi_{4+k}(\mathbf{x}) = \lambda_2(\mathbf{x})\lambda_4(\mathbf{x})\lambda_H^{k-1}(\mathbf{x}), \qquad k = 1, 2, \ldots, r-1, \tag{14}$$

$$\phi_{4+(r-1)+k}(\mathbf{x}) = \lambda_2(\mathbf{x})\lambda_4(\mathbf{x})\lambda_V(\mathbf{x})\lambda_H^{k-1}(\mathbf{x}), \qquad k = 1, 2, \ldots, r-2, \tag{15}$$

$$\phi_{4+2(r-1)}(\mathbf{x}) = \lambda_2(\mathbf{x})\lambda_4(\mathbf{x})\lambda_H^{r-2}(\mathbf{x})R_V(\mathbf{x}). \tag{16}$$

In a similar way we define shape functions associated with edge e_2 and e_4.

$$\phi_{4+2(r-1)+k}(\mathbf{x}) = \lambda_1(\mathbf{x})\lambda_3(\mathbf{x})\lambda_V^{k-1}(\mathbf{x}), \qquad k = 1, 2, \ldots, r-1, \tag{17}$$

$$\phi_{4+3(r-1)+k}(\mathbf{x}) = \lambda_1(\mathbf{x})\lambda_3(\mathbf{x})\lambda_H(\mathbf{x})\lambda_V^{k-1}(\mathbf{x}), \qquad k = 1, 2, \ldots, r-2, \tag{18}$$

$$\phi_{4+4(r-1)}(\mathbf{x}) = \lambda_1(\mathbf{x})\lambda_3(\mathbf{x})\lambda_V^{r-2}(\mathbf{x})R_H(\mathbf{x}). \tag{19}$$

The edge shape functions are regular polynomials of degree r except the last two functions in each direction, which are rational functions (which we call supplemental functions).

Finally, there are $\dim \mathbb{P}_{r-4} = \binom{r-2}{2}$ DOFs associated with the interior of the quadrilateral, which only appear when $r \geq 4$. Let

$$\phi_{4+4(r-1)+k}(\mathbf{x}) = \lambda_1(\mathbf{x})\lambda_3(\mathbf{x})\lambda_2(\mathbf{x})\lambda_4(\mathbf{x})\lambda_V^m(\mathbf{x})\lambda_H^n(\mathbf{x}), \tag{20}$$
$$m \geq 0,\ n \geq 0,\ m+n \leq r-4, \text{ and } k = \frac{1}{2}(m+n)(m+n+1) + n + 1.$$

The interior shape functions are polynomials of degree r.

The space of shape functions is

$$\mathcal{DS}_r(E) = \operatorname{span}\{\phi_1, \phi_2, \ldots, \phi_{D_r}\}, \quad D_r = \frac{1}{2}(r+2)(r+1) + 2. \tag{21}$$

We claim that the dimension is

$$\begin{aligned} \dim \mathcal{DS}_r(E) &= \sum_{d=0}^{\min(2,\lfloor r/2 \rfloor)} 2^{2-d}\binom{2}{d}\binom{r-d}{d} = 4r + \frac{1}{2}(r-2)(r-3) \\ &= \frac{1}{2}(r+2)(r+1) + 2 = \dim \mathbb{P}_r(E) + 2 = D_r, \end{aligned} \tag{22}$$

that is, for any order $r \geq 2$, the dimension of the direct serendipity finite element shape function space is D_r, two plus the dimension of the polynomials of degree r in $\mathbb{R}^2$. This claim will be proven by the independence of the degrees of freedom, i.e., Theorem 2.5. The two supplemental functions that do not belong to $\mathbb{P}_r(E)$ are $\phi_{4+2(r-1)}$ and $\phi_{4+4(r-1)}$.

2.3. Degrees of freedom

We define the DOFs ($\mathcal{N}$ in Definition 2.1) as a set of nodal functionals N_i defined at a nodal point $\mathbf{x}_i$, i.e.,

$$\mathcal{N} = \{N_i : N_i(\phi) = \phi(\mathbf{x}_i) \text{ for all } \phi(\mathbf{x}),\ i = 1, 2, \ldots, D_r\}. \tag{23}$$

For vertex DOFs, the nodal points $\mathbf{x}_i = \mathbf{x}_c^i$, $i = 1, 2, 3, 4$, are exactly the corner points of quadrilateral E. For edge DOFs, the nodal points are equally distributed on each edge. We choose to number them consecutively from "bottom" to "top" or "left" to "right", and numbered for e_1, e_3, e_2, and then e_4. Examples for $r = 2, 3, 4, 5$ can be found in Figure 3. To be specific, points $\mathbf{x}_{4+i}, i = 1, \ldots, r-1$, are distributed on edge e_1 equidistantly in the direction from $\mathbf{x}_c^1$ to $\mathbf{x}_c^4$, and $\mathbf{x}_{4+(r-1)+i}, i = 1, \ldots, r-1$, are distributed on edge e_3 in a similar way from $\mathbf{x}_c^2$ to $\mathbf{x}_c^3$. Nodal points

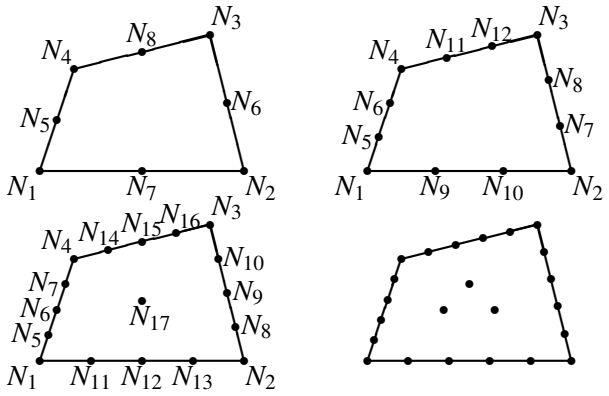


Figure 3: The degrees of freedom of the direct serendipity finite element space for $r = 2$ (top left), $r = 3$ (top right), $r = 4$ (botton left) and $r = 5$ (bottom right).

$\mathbf{x}_{4+2(r-1)+i}, i = 1, \ldots, r-1$, are evenly set on edge e_2 from $\mathbf{x}_c^1$ to $\mathbf{x}_c^2$ and $\mathbf{x}_{4+3(r-1)+i}, i = 1, \ldots, r-1$, on edge e_4 from $\mathbf{x}_c^4$ to $\mathbf{x}_c^3$. For interior DOFs, find a triangle T strictly inside E and set the nodal points the same as the nodes of the Lagrange element of order $r-4$ for the triangle T. We can label the interior nodal points as $\mathbf{x}_{4+4(r-1)+i}, i = 1, \ldots, D_r = \frac{1}{2}(r-3)(r-2)$.

In order to prove the unisolvence of the degrees of freedom (i.e., that $\mathcal{N}$ is a basis for $\mathcal{P}'$), we require several lemmas. The proof of the following lemma can be found in [3, (3.1.10)].

Lemma 2.2. *Let p be a polynomial of degree $r \geq 1$ that vanishes on a hyperplane defined as the zero set of a linear polynomial L. Then $p = Lq$, where q is a polynomial of degree $r-1$.*

Lemma 2.3. *There exist $\alpha_V, \alpha_H > 0$ and $\beta_V, \beta_H \in \mathbb{R}$ such that, for all $x \in E$,*

$$\xi_V \lambda_1(\mathbf{x}) + \eta_V \lambda_3(\mathbf{x}) = \alpha_V + \beta_V \lambda_H(\mathbf{x}), \tag{24}$$

$$\xi_H \lambda_2(\mathbf{x}) + \eta_H \lambda_4(\mathbf{x}) = \alpha_H + \beta_H \lambda_V(\mathbf{x}). \tag{25}$$

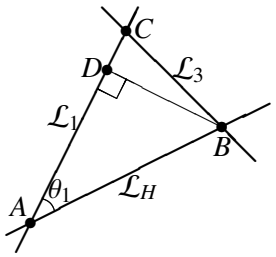


Figure 4: The triangle formed by vertices A, B, and C. The quadrilateral edge e_1 is on the line $\mathcal{L}_1$ and e_3 is on $\mathcal{L}_3$, and these are *not* parallel. The line $\mathcal{L}_H$ is the zero set of λ_H. The acute angle θ_1 is between $\mathcal{L}_1$ and $\mathcal{L}_H$ at A. A right triangle is formed with vertices A, B, and D, where D is on $\mathcal{L}_1$.

Proof. By symmetry, we need only show (24). We first suppose that e_1 is *not* parallel to e_3. As shown in Figure 4, we can then consider the triangle bounded by the lines $\mathcal{L}_1 = \{\mathbf{x} : \lambda_1(\mathbf{x}) = 0\}$, $\mathcal{L}_3 = \{\mathbf{x} : \lambda_3(\mathbf{x}) = 0\}$, and $\mathcal{L}_H = \{\mathbf{x} : \lambda_H(\mathbf{x}) = 0\}$. It has the vertices $A = \mathcal{L}_1 \cap \mathcal{L}_H$, $B = \mathcal{L}_3 \cap \mathcal{L}_H$,

and $C = \mathcal{L}_1 \cap \mathcal{L}_3$. Since $\lambda_1(\mathbf{x})$, $\lambda_3(\mathbf{x})$, and $\lambda_H(\mathbf{x})/\|\nu_2 - \nu_4\|$ are the distances from $\mathbf{x}$ to the lines $\mathcal{L}_1$, $\mathcal{L}_3$, and $\mathcal{L}_H$, respectively, the barycentric coordinates on this triangle are $\lambda_1(\mathbf{x})/\lambda_1(B)$, $\lambda_3(\mathbf{x})/\lambda_3(A)$, and $\lambda_H(\mathbf{x})/(\lambda_H(C)\,\|\nu_2 - \nu_4\|)$. These sum to one, so after rearranging,

$$\frac{|AB|}{\lambda_1(B)}\lambda_1(\mathbf{x}) + \frac{|AB|}{\lambda_3(A)}\lambda_3(\mathbf{x}) = |AB| - \frac{|AB|}{\lambda_H(C)\,\|\nu_2 - \nu_4\|}\lambda_H(\mathbf{x}), \tag{26}$$

where $|AB|$ is the length of the line segment AB. If e_1 and e_3 are parallel, then we have more simply that $\lambda_1(\mathbf{x})/\lambda_1(B)$ and $\lambda_3(\mathbf{x})/\lambda_3(A)$ sum to one, so

$$\frac{|AB|}{\lambda_1(B)}\lambda_1(\mathbf{x}) + \frac{|AB|}{\lambda_3(A)}\lambda_3(\mathbf{x}) = |AB|. \tag{27}$$

Let $0 < \theta_1, \theta_3 \leq \pi/2$ be the acute angles defined by the lines $\mathcal{L}_1$ and $\mathcal{L}_3$ with the line $\mathcal{L}_H$, respectively (see Figure 4). It can be easily verified that

$$\sin\theta_1 = \frac{\lambda_1(B)}{|AB|} \quad \text{and} \quad \sin\theta_3 = \frac{\lambda_3(A)}{|AB|}. \tag{28}$$

Moreover, if τ_j are unit vectors along the lines $\mathcal{L}_j$, for $j = 1, 3, H$, then

$$\cos\theta_1 = |\tau_H \cdot \tau_1| = |\nu_H \cdot \nu_1| \quad \text{and} \quad \cos\theta_3 = |\tau_H \cdot \tau_3| = |\nu_H \cdot \nu_3|,$$

and so

$$\sin\theta_1 = \sqrt{1 - \cos^2\theta_1} = \sqrt{1 - (\nu_H \cdot \nu_1)^2} = \xi_V^{-1} \quad \text{and} \quad \sin\theta_3 = \eta_V^{-1}. \tag{29}$$

The result follows from this and (28) combined with either (26) or (27). □

Lemma 2.4. *For any $k \geq 1$, the following function spaces are identical:*

$$\mathcal{P}_1 = \text{span}\,\{1, \lambda_H, \ldots, \lambda_H^k, \lambda_V, \lambda_H\lambda_V, \ldots, \lambda_H^{k-1}\lambda_V, \lambda_H^k R_V\},$$
$$\mathcal{P}_2 = \text{span}\,\left\{\{1, \lambda_H, \ldots, \lambda_H^k\} \otimes \{1, R_V\}\right\}.$$

Moreover, the following two function spaces are also identical:

$$\mathcal{P}_3 = \text{span}\,\{1, \lambda_V, \ldots, \lambda_V^k, \lambda_H, \lambda_V\lambda_H, \ldots, \lambda_V^{k-1}\lambda_H, \lambda_V^k R_H\},$$
$$\mathcal{P}_4 = \text{span}\,\left\{\{1, \lambda_V, \ldots, \lambda_V^k\} \otimes \{1, R_H\}\right\}.$$

Proof. The dimension of $\mathcal{P}_1$ is $2(k+1)$ and it is equal to the dimension of $\mathcal{P}_2$. From Lemma 2.3 and the definition of R_V (7), we know that

$$\lambda_V = (\alpha_V + \beta_V\lambda_H)R_V = \alpha_V R_V + \beta_V\lambda_H R_V, \tag{30}$$

where α_V is a strictly positive constant depending on the geometry of the element E.

We order the base of the function space $\mathcal{P}_2$ as

$$\{1, \lambda_H, \ldots, \lambda_H^k, R_V, \lambda_H R_V, \ldots, \lambda_H^k R_V\}. \tag{31}$$

In order to prove the equivalence of the two spaces, we only need to show that the transformation matrix $\mathbf{M}$ of the bases is invertible. The transformation matrix $\mathbf{M}$ from $\mathcal{P}_2$ to $\mathcal{P}_1$ is a $2(k+1)\times 2(k+1)$ matrix, and we can write it as a 2×2 block matrix, with each block being a $(k+1)\times(k+1)$ matrix. It can be easily verified that $\mathbf{M}_{11}$ is an identity matrix and $\mathbf{M}_{12}$ and $\mathbf{M}_{21}$ are zero. By (30),

$$\mathbf{M} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{22} \end{pmatrix}, \quad \mathbf{M_{22}} = \begin{pmatrix} \alpha_V & \beta_V & & & \\ & \alpha_V & \beta_V & & \\ & & \ddots & \ddots & \\ & & & \alpha_V & \beta_V \\ & & & & 1 \end{pmatrix}.$$

The matrix $\mathbf{M}_{22}$ is upper triangular with strictly positive diagonal entries, so the transformation matrix $\mathbf{M}$ is invertible and the two function spaces are the same.

In a similar way, $\mathcal{P}_3$ and $\mathcal{P}_4$ are identical as well. □

We now show the unisolvence of the degrees of freedom, which completes the requirements of Ciarlet's Definition 2.1 for $\mathcal{DS}_r(E)$ to be a well defined finite element, and shows $\dim \mathcal{DS}_r(E) = D_r = \dim \mathbb{P}_r + 2$.

Theorem 2.5. *Let $(E, \mathcal{DS}_r(E), \mathcal{N})$ be the r-th order direct serendipity finite element defined by* (21) (*i.e.,* (13)–(20)) *and* (23). *If $\psi \in \mathcal{DS}_r(E)$ and $N_k(\psi) = 0$, for all $k = 1, 2, \ldots, D_r$, then $\psi = 0$.*

Proof. Since $\psi \in \mathcal{DS}_r(E)$,

$$\psi(\mathbf{x}) = \sum_{k=0}^{D_r} a_k \phi_k(\mathbf{x}) \tag{32}$$

for some coefficients a_k. The corner DOFs vanish for all ϕ_k when $k > 4$, and when $j, k \le 4$, $j \ne k$, $N_j(\phi_k) = 0$ and $N_k(\phi_k) > 0$. We conclude that $a_k = 0$ for all $k \le 4$.

Next we consider the DOFs for nodal values on edges e_1 and e_3. Let the span of the shape functions associated with these edges, i.e., those in (14)–(16), be

$$\operatorname{span}\{\phi_5, \ldots, \phi_{4+2(r-1)}\} = (\lambda_2\lambda_4)\mathcal{P}_1 = (\lambda_2\lambda_4)\mathcal{P}_2,$$

where $\mathcal{P}_1$ and $\mathcal{P}_2$ are defined in Lemma 2.4. These functions vanish on e_2 and e_4, and they are polynomials of degree r on e_1 and e_3 by (11). To proceed, we simplify the basis of $\mathcal{P}_2$ for the purposes of this proof. Construct the functions

$$\zeta_1(x) = \eta_V^{-1} + R_V(x) \quad \text{and} \quad \zeta_3(x) = -\xi_V^{-1} + R_V(x), \tag{33}$$

and note that

$$\zeta_1 = \begin{cases} 0 & \text{on } e_1, \\ \xi_V^{-1} + \eta_V^{-1} & \text{on } e_3, \end{cases} \quad \text{and} \quad \zeta_3 = \begin{cases} -\xi_V^{-1} - \eta_V^{-1} & \text{on } e_1, \\ 0 & \text{on } e_3. \end{cases}$$

Therefore the functions ζ_1 and ζ_3 are non-zero constants on the interior of e_3 and e_1, respectively. It is clear from the definition (33) that

$$\operatorname{span}\{1, R_V\} = \operatorname{span}\{\zeta_1, \zeta_3\},$$

so

$$\operatorname{span}\{\phi_5,\dots,\phi_{4+2(r-1)}\} = (\lambda_2\lambda_4)\mathcal{P}_2 = (\lambda_2\lambda_4)\operatorname{span}\{\{1,\lambda_H,\dots,\lambda_H^{r-2}\}\otimes\{\zeta_1,\zeta_3\}\}.$$

Write the new basis as

$$\{\phi'_5,\dots,\phi'_{4+(r-1)}\} = \lambda_2\lambda_4\zeta_3\cdot\{1,\lambda_H,\dots,\lambda_H^{r-2}\}, \tag{34}$$

$$\{\phi'_{4+(r-1)+1},\dots,\phi'_{4+2(r-1)}\} = \lambda_2\lambda_4\zeta_1\cdot\{1,\lambda_H,\dots,\lambda_H^{r-2}\}, \tag{35}$$

and expand ψ as

$$\psi(\mathbf{x}) = \sum_{k=5}^{4+(r-1)} b_k\phi'_k + \sum_{k=4+(r-1)+1}^{4+2(r-1)} b_k\phi'_k + \sum_{k=4+2(r-1)+1}^{D_r} a_k\phi_k. \tag{36}$$

Now for $j = 5,\dots,4+(r-1)$,

$$0 = N_j(\psi) = \sum_{k=5}^{4+(r-1)} b_k\, N_j(\phi'_k), \tag{37}$$

and so $\psi|_{e_1} \in \mathbb{P}_r(e_1)$ vanishes at $r-1$ interior and two vertex points (see Figure 3). Thus $b_k = 0$ when $k \le 4+(r-1)$. Similarly, $b_k = 0$ for $4+(r-1)+1 \le k \le 4+2(r-1)$, and so $a_k = 0$ when $k \le 4+2(r-1)$. By symmetry for edges e_2 and e_4, we conclude that $a_k = 0$ for all $k \le 4+4(r-1)$.

If $r < 4$, there are no interior degrees of freedom, so we conclude that $\psi \equiv 0$ and we are done. Otherwise, we know that $\psi \in \mathbb{P}_r(E)$ and $\psi|_{e_j} = 0$, $j = 1,2,3,4$. Lemma 2.2 shows that $\psi = \lambda_1\lambda_2\lambda_3\lambda_4 q$, where $q \in \mathbb{P}_{r-4}(\mathbb{R}^2)$. The choice of interior nodal points N_i, $i = 4+4(r-1)+1,\dots,D_r$ leads us to conclude that $q \equiv 0$ and then $\psi \equiv 0$. □

The following Corollary can be easily derived from Theorem 2.5 and the dimension count (22).

Corollary 2.6. *The polynomial space $\mathbb{P}_r(E) \subset \mathcal{DS}_r(E)$. Moreover, $\mathbb{P}_r(E)$ is the span of all ϕ_i defined in* (13)–(20) *for which $1 \le i \le \dim \mathcal{DS}_r(E)$ and $i \ne 4+2(r-1)$, $i \ne 4+4(r-1)$.*

3. Implementation as an H^1-Conforming Space

Following [15, pp. 104–105], we require that the quadrilateral E be shape-regular, which implies that E does not degenerate to a triangle or a line. Denote by T_i, $i = 1,2,3,4$, the subtriangle of E with vertices $\mathbf{x}_c^i$, $\mathbf{x}_c^{i+1}$ and $\mathbf{x}_c^{i+2}$ (where $\mathbf{x}_c^5 = \mathbf{x}_c^1$ and $\mathbf{x}_c^6 = \mathbf{x}_c^2$). We define parameters to describe the geometry of the element E, which are

$$h_E = \text{diameter of } E, \tag{38}$$

$$\rho_E = 2\min_{1\le i\le 4}\{\text{diameter of circle inscribed in } T_i\}, \tag{39}$$

$$d_E = \sqrt{|E|}. \tag{40}$$

Shape regularity requires the existence of σ_* such that the ratio $\dfrac{\rho_E}{h_E} \ge \sigma_* > 0$. (We remark that σ_* exists for any single non-degenerate convex quadrilateral, but we will need these shape-regular parameters here and in Section 4 for a mesh of elements.) It should be clear that ρ_E, d_E, and h_E remain comparable as $|E| \to 0$.

3.1. A local, nearly nodal basis for the finite element

Let $\sharp(\phi_j)$ be the number of "λ" factors in ϕ_j, i.e., $\sharp(\phi_j)$ is equal to the polynomial degree if ϕ_j does not include a rational factor (R_V or R_H), and $\sharp(\phi_{4+2(r-1)}) = \sharp(\phi_{4+4(r-1)}) = r$. To reduce rounding errors in the following computations, we scale the shape functions as

$$\tilde{\phi}_j = \frac{\phi_j}{d_E^{\sharp(\phi_j)}}, \quad j = 1, 2, \ldots, \dim \mathcal{DS}_r(E). \tag{41}$$

To implement H^1-conforming direct serendipity elements, we need to find a proper basis for the finite element space, i.e., one that is continuous. We observe that the interior shape functions can be extended by zero. There are $4+4(r-1) = 4r$ vertex and edge degrees of freedom. Therefore, we construct a $4r \times 4r$ matrix $\mathbf{A}$, where $\mathbf{A}_{ij} = N_i(\tilde{\phi}_j)\ \forall i, j \le 4r$. This matrix has a block structure based on the DOFs on the vertices, edges e_1 and e_3, and edges e_2 and e_4, i.e.,

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{0} \\ \mathbf{A}_{23} & \mathbf{0} & \mathbf{A}_{33} \end{pmatrix}, \tag{42}$$

where $\mathbf{A}_{11}$ is of size 4×4 and $\mathbf{A}_{22}$ and $\mathbf{A}_{33}$ are of size $2(r-1) \times 2(r-1)$. From the proof of Theorem 2.5, we know that $\mathbf{A}$ is invertible.

We want to construct $\{\varphi_1, \ldots, \varphi_{\dim \mathcal{DS}_r(E)}\}$ as a basis for $\mathcal{DS}_r(E)$ such that the vertex and edge part is nodal, i.e., $N_k(\varphi_j) = \delta_{kj}$ for $k, j \le 4r$. (If we can do this, the shape functions can be joined continuously.) Assume φ_j, $j \le 4r$, is a linear combination of only the vertex and edge shape functions, i.e., $\varphi_j = \sum_{i=1}^{4r} b_{ji}\tilde{\phi}_i$. Since, for $k, j \le 4r$,

$$N_k(\varphi_j) = \sum_{i=1}^{4r} b_{ji} N_k(\tilde{\phi}_i) = \sum_{i=1}^{4r} b_{ji}\mathbf{A}_{ki} = \delta_{kj}, \tag{43}$$

we obtain that $b_{ji} = \mathbf{B}_{ji}$, where $\mathbf{B} = \mathbf{A}^{-T}$. For $4r < j \le \dim \mathcal{DS}_r(E)$, let $\varphi_j = \tilde{\phi}_j$.

3.2. An example

In this section, we provide an example of how to compute the direct serendipity finite element basis functions for $r = 2$ on a specific element E with vertices $(0,0)$, $(1,0)$, $(0.25, 0.5)$ and $(0.75, 0.75)$ (see Figures 5). The degrees of freedom for $r = 2$ are shown in Figure 3, i.e., there are no interior degrees of freedom and the $\dim \mathcal{DS}_2(E) = 8$ degrees of freedom are distributed at the four vertices and the middle points of the four edges.

First, we calculate the auxiliary functions λ_1 to λ_4, λ_V, λ_H, R_V and R_H defined in Section 2.1. That is,

$$\begin{aligned} {}_1(x,y) &= \frac{1}{\sqrt{5}}(2x - y), & \lambda_2(x,y) &= x, \\ \lambda_3(x,y) &= \frac{1}{\sqrt{10}}(3 - 3x - y), & \lambda_4(x,y) &= \frac{1}{4\sqrt{5}}(3 + 4x - 8y). \end{aligned} \tag{44}$$

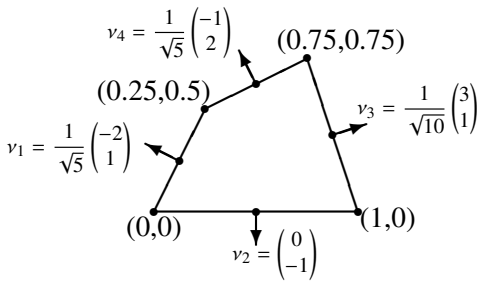


Figure 5: Example element E.

We can obtain λ_V and λ_H from (3)–(4), and also define $\nu_V = \begin{pmatrix}-0.9975\\0.0709\end{pmatrix}$ and $\nu_H = \begin{pmatrix}0.2298\\-0.9732\end{pmatrix}$ from (5)–(6). From (9)–(10), we know that $\xi_V = 1.3025$, $\eta_V = 1.0041$, $\xi_H = 1.0025$ and $\eta_H = 1.1621$. Therefore, the directional rational functions R_V and R_H can be calculated from (7)–(8). The functions $\lambda_1, \ldots, \lambda_4$, λ_V, λ_H, R_V and R_H are plotted in Figure 6.

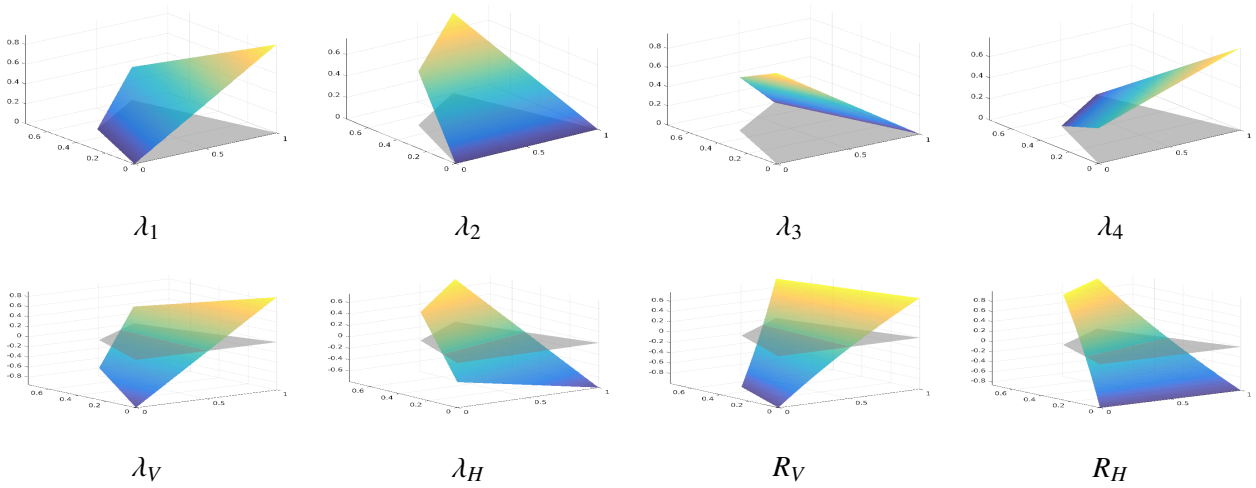
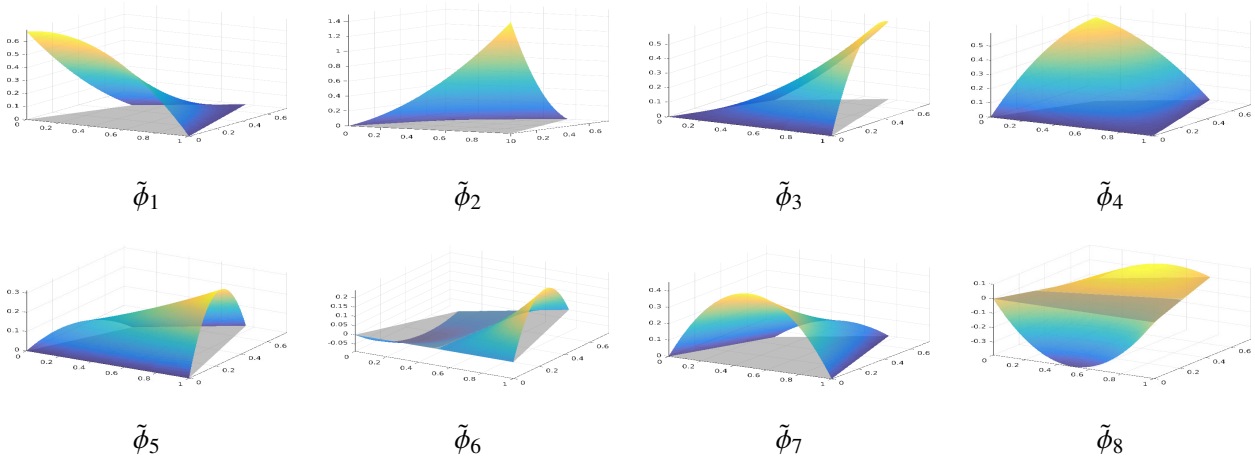


Figure 6: Auxiliary functions for element E in Figure 5.

The shape functions $\phi_1, \ldots, \phi_8$ are defined in (13)–(20). Easily, $|E| = 0.46875$ and therefore the scaling factor $d_E = 0.68465$. Combining the definitions in Section 2.2 and (41) (where $\sharp(\phi_j) = 2$ $\forall j$), we have the scaled shape functions $\tilde{\phi}_1, \ldots, \tilde{\phi}_8$, which are shown in Figure 7.

Finally, we set the entries of matrix $\mathbf{A}$, $\mathbf{A}_{ij} = N_i(\tilde{\phi}_j)$, which are

$$\mathbf{A} = \left(\begin{array}{cccc|cc|cc} 0.68 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.50 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.54 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.6 & 0 & 0 & 0 & 0 \\ \hline 0.27 & 0 & 0 & 0.4 & 0.09 & -0.09 & 0 & 0 \\ 0 & 0.51 & 0.49 & 0 & 0.31 & 0.24 & 0 & 0 \\ \hline 0.57 & 0.53 & 0 & 0 & 0 & 0 & 0.45 & -0.39 \\ 0 & 0 & 0.22 & 0.37 & 0 & 0 & 0.1 & 0.1 \end{array}\right). \tag{45}$$



$\tilde{\phi}_1$ $\tilde{\phi}_2$ $\tilde{\phi}_3$ $\tilde{\phi}_4$

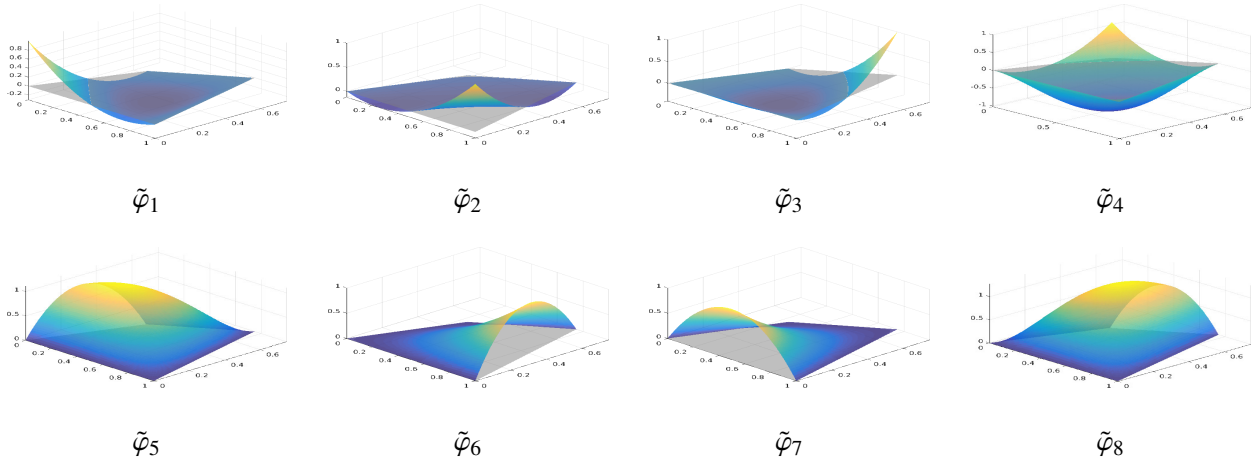
$\tilde{\phi}_5$ $\tilde{\phi}_6$ $\tilde{\phi}_7$ $\tilde{\phi}_8$

Figure 7: Scaled shape functions for element E in Figure 5 when $r = 2$.

Therefore, matrix $\mathbf{B} = \mathbf{A}^{-T}$ can be computed, and the basis functions are

$$\begin{aligned}
\varphi_1 &= 1.47\tilde{\phi}_1 - 1.92\tilde{\phi}_5 + 2.51\tilde{\phi}_6 - 0.99\tilde{\phi}_7 + 0.99\tilde{\phi}_8, & \varphi_5 &= 4.87\tilde{\phi}_5 - 6.34\tilde{\phi}_6,\\
\varphi_2 &= 0.67\tilde{\phi}_2 - 0.62\tilde{\phi}_5 - 0.62\tilde{\phi}_6 - 0.42\tilde{\phi}_7 + 0.42\tilde{\phi}_8, & \varphi_6 &= 1.80\tilde{\phi}_5 + 1.81\tilde{\phi}_6,\\
\varphi_3 &= 1.86\tilde{\phi}_3 - 1.65\tilde{\phi}_5 - 1.66\tilde{\phi}_6 - 1.95\tilde{\phi}_7 - 2.27\tilde{\phi}_8, & \varphi_7 &= 1.19\tilde{\phi}_7 - 1.19\tilde{\phi}_8,\\
\varphi_4 &= 1.69\tilde{\phi}_4 - 3.30\tilde{\phi}_5 + 4.30\tilde{\phi}_6 - 2.92\tilde{\phi}_7 - 3.40\tilde{\phi}_8, & \varphi_8 &= 4.68\tilde{\phi}_7 + 5.44\tilde{\phi}_8.
\end{aligned} \tag{46}$$

These are shown in Figure 8.



$\tilde{\varphi}_1$ $\tilde{\varphi}_2$ $\tilde{\varphi}_3$ $\tilde{\varphi}_4$

$\tilde{\varphi}_5$ $\tilde{\varphi}_6$ $\tilde{\varphi}_7$ $\tilde{\varphi}_8$

Figure 8: Basis shape functions for element E in Figure 5 when $r = 2$.

3.3. The global direct serendipity spaces

Let $\Omega \subset \mathbb{R}^2$ be a closed polygonal domain and impose a conforming finite element partition $\mathcal{T}_h$ of nondegenerate, convex quadrilaterals over Ω. The global direct serendipity finite element space of index $r \geq 2$ over $\mathcal{T}_h$ is

$$\mathcal{DS}_r = \{v_h \in C^0(\Omega) \ : \ v_h|_E \in \mathcal{DS}_r(E) \ \forall E \in \mathcal{T}_h\} \subset H^1(\Omega). \tag{47}$$

For the partition $\mathcal{T}_h$, let n_v, n_e, and n_i denote the total number of vertices, edges and quadrilaterals, respectively. The total number of nodal points is then

$$\dim \mathcal{DS}_r = n_v + (r-1)n_e + \frac{(r-2)(r-3)}{2} n_i.$$

These global nodal points can be ordered and expressed as $\mathbf{x}_j^h$, with N_j^h being the corresponding nodal linear functional for $\mathbf{x}_j^h$. A global basis for $\mathcal{DS}_r$ is given by

$$\mathcal{B}_h = \{\psi_j^h,\ 1 \le j \le \dim \mathcal{DS}_r\}, \tag{48}$$

where each ψ_j^h combines local basis functions associated with the same node.

4. Error Estimation and Stability

In Section 3, we defined a practical basis for $\mathcal{DS}_r$ which is nodal only at vertex and edge degrees of freedom, but simply extends the internal shape functions by zero (in order to reduce the computational cost). However, by Theorem 2.5, we know there exists a fully nodal basis of $\mathcal{DS}_r$ (i.e., one for which every basis function vanishes on all but one nodal function). Abusing notation, we denote it simply as $\{\psi_1, \ldots, \psi_{\dim \mathcal{DS}_r}\}$ in this section.

Definition 4.1. *Given the r-th order direct serendipity element $(E, \mathcal{DS}_r(E), \mathcal{N})$ and the nodal basis of $\mathcal{DS}_r(E)$, $\{\psi_1^E, \ldots, \psi_{\dim \mathcal{DS}_r(E)}^E\}$, let the operator $\mathcal{I}_E : L^2(E) \cap C^0(E) \longrightarrow \mathcal{DS}_r(E)$ be interpolation. That is, for a given function $v \in L^2(E) \cap C^0(E)$, $\mathcal{I}_E v \in \mathcal{DS}_r(E)$ and*

$$\mathcal{I}_E v = \sum_{j=1}^{\dim \mathcal{DS}_r(E)} N_j(v)\, \psi_j^E = \sum_{j=1}^{\dim \mathcal{DS}_r(E)} v(\mathbf{x}_j)\, \psi_j^E.$$

By Corollary 2.6, the interpolation operator preserves polynomials, so we have an important property [2, pp. 121–123] expressed in the following lemma.

Lemma 4.2. *The interpolation operator $\mathcal{I}_E$ is polynomial preserving, i.e., $\forall \psi \in \mathbb{P}_r(E)$, $\mathcal{I}_E \psi = \psi$. Moreover, $\|\mathcal{I}_E\|$ is bounded in the L^2-norm.*

With Lemma 4.2, the shape regularity defined in Section 3, and the inclusion Corollary 2.6, we have the analogue of the Bramble-Hilbert [16] or Dupont-Scott [17] lemma for local error estimation.

Lemma 4.3. *There exists a constant $C > 0$ such that for all functions $v \in H^{s+1}(E)$ ($H^1(E) \cap C^0(E)$ if $s = 0$),*

$$|v - \mathcal{I}_E v|_{m,E} \le C\, h_E^{s+1-m} |v|_{s+1,E}, \quad m = 0, 1 \text{ and } s = 0, 1, \ldots, r, \tag{49}$$

where $|\cdot|_{m,E}$ is the $H^m(E)$ seminorm.

Definition 4.4. *Given the finite element space $\mathcal{DS}_r$ with its basis nodal $\mathcal{B}_h$, let the operator $\mathcal{I}_h$ be global interpolation. That is, for a given function $v \in L^2(\Omega) \cap C^0(\Omega)$, $\mathcal{I}_h v \in \mathcal{DS}_r$ and*

$$\mathcal{I}_h v = \sum_{j=1}^{\dim \mathcal{DS}_r} N_j^h(v)\, \psi_j^h = \sum_{j=1}^{\dim \mathcal{DS}_r} v(\mathbf{x}_j^h)\, \psi_j^h.$$

From the definition and the way we construct ψ_j^h, we obtain that

$$(\mathcal{I}_h v)|_E = (\mathcal{I}_E v)|_E \quad \forall v \in L^2(\Omega) \cap C^0(\Omega),\ E \in \mathcal{T}_h. \tag{50}$$

Combining Lemma 4.3 and $|v - \mathcal{I}_h v|_{m,\Omega}^2 = \sum_{E \in \mathcal{T}_h} |v - \mathcal{I}_E v|_{m,E}^2$, we have the following global projection error estimate, provided we have shape regularity of the mesh. For the quadrilateral finite element partition $\mathcal{T}_h$ of Ω, let the maximal diameter $h = \max_{E \in \mathcal{T}_h} h_E$. Each element $E \in \mathcal{T}_h$ is *uniformly* shape regular, i.e., referring to (38)–(39), if there exists a constant σ_* such that $\dfrac{\rho_E}{h_E} \geq \sigma_* > 0\ \forall E \in \mathcal{T}_h$, where σ_* is independent of $\mathcal{T}_h$.

Lemma 4.5. *Let $\mathcal{T}_h$ be uniformly shape regular. There exists a constant C, independent of h, such that for all functions $v \in H^{s+1}(\Omega)$ ($v \in H^1(\Omega) \cap C^0(\overline{\Omega})$ if $s = 0$),*

$$|v - \mathcal{I}_h v|_{m,\Omega} \leq C\, h^{s+1-m} |v|_{s+1,\Omega}, \quad m = 0, 1 \text{ and } s = 0, 1 \ldots, r. \tag{51}$$

Consider a heterogeneous elliptic problem with a homogeneous Dirichlet boundary condition

$$\begin{aligned} -\nabla \cdot (\mathbf{a} \nabla u) &= f \quad \text{in} \quad \Omega, \\ p &= 0 \quad \text{on} \quad \partial\Omega, \end{aligned} \tag{52}$$

where the second order tensor $\mathbf{a}$ is uniformly positive definite and bounded, and $f \in L^2(\Omega)$. The boundary value problem can be written in the weak form: Find $u \in H_0^1(\Omega)$ such that

$$a(u, v) = f(v), \quad \forall v \in H_0^1(\Omega), \tag{53}$$

where $a(u, v) = (\mathbf{a}\nabla u, \nabla v)$ and $f(v) = (f, v)$, $(\cdot, \cdot)$ being the $L^2(\Omega)$-innerproduct.

Define the global finite element space over $\mathcal{T}_h$

$$X_{0,h} = \{v_h \in \mathcal{DS}_r \ :\ v_h = 0 \text{ on } \partial\Omega\} \subset H_0^1(\Omega).$$

We then obtain the Galerkin approximation: Find $u_h \in X_{0,h}$ such that

$$a(u_h, v_h) = f(v_h), \quad \forall v_h \in X_{0,h}. \tag{54}$$

Combining Céa's lemma [2, 3] and the global projection estimate Lemma 4.5, we obtain an H^1-error estimate for the problem. Since Ω is a polygonal domain, $\partial\Omega$ is a Lipschitz boundary. If we assume that Ω is also convex, we have elliptic regularity of the solution [18, Theorem 4.3.1.4], and the Aubin-Nitsche duality principle [2, 3] gives an L^2-error estimate.

Theorem 4.6. *Let Ω be a convex polygonal domain and let $\mathcal{T}_h$ be uniformly shape regular. There exists a constant C, independent of the subspace $X_{0,h}$, such that*

$$\|u - u_h\|_{1,\Omega} \leq C\, h^s |u|_{s+1,\Omega}, \tag{55}$$

$$\|u - u_h\|_{0,\Omega} \leq C\, h^{s+1} |u|_{s+1,\Omega}, \quad s = 0, 1, \ldots, r. \tag{56}$$

5. Numerical Results

In this section we present convergence studies for the direct serendipity spaces $\mathcal{DS}_r$, the regular serendipity spaces $\mathcal{S}_r$, and the space given by mapping the local tensor product spaces $\mathbb{P}_{r,r}$ defined on the reference $\hat{E}$ to the elements (hereafter simply called the $\mathbb{P}_{r,r}$ space). The test problem (52) is defined on the unit square $\Omega = [0,1]^2$ with the coefficient $\mathbf{a}$ being the 2×2 identity matrix. The exact solution is $u(x,y) = \sin(\pi x)\sin(\pi y)$ and the source term $f(x,y) = 2\pi^2 \sin(\pi x)\sin(\pi y)$. Solutions are computed on three different sequences of meshes. The first sequence, $\mathcal{T}_h^1$, is a uniform mesh of n^2 square elements (two sets of parallel edges per element). The second sequence, $\mathcal{T}_h^2$, is a mesh of n^2 trapezoids of base h and one pair of parallel edges of size $0.75h$ and $1.25h$, as proposed in [12]. The third sequence, $\mathcal{T}_h^3$, is chosen so as to have no pair of edges parallel. The first 4×4 meshes for each sequence are shown in Figure 9. Finer meshes are constructed by repeating the same pattern over the domain.

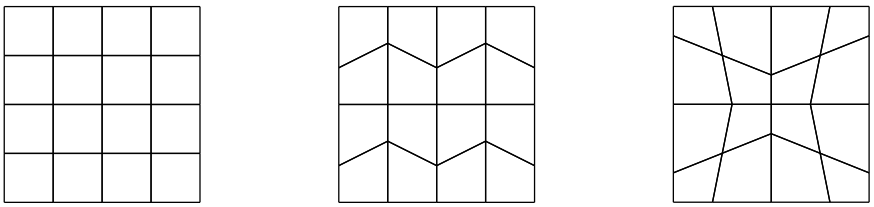


Figure 9: Mesh of 4×4 squares for the three base meshes. Finer meshes are constructed by repeating this base mesh pattern over the domain. Note that the meshes have 2, 1, and 0 parallel edges per element, respectively.

For an $n \times n$ mesh, we know that $n_v = (n+1)^2$, $n_e = 2n(n+1)$, and $n_i = n^2$. Therefore, the total number of degrees of freedom for $\mathbb{P}_{r,r}$ is $(nr+1)^2 = O(r^2n^2)$ and the total number of degrees of freedom for $\mathcal{S}_r$ and $\mathcal{DS}_r$, is

$$\begin{aligned}\dim(\mathcal{S}_r) = \dim(\mathcal{DS}_r) &= n_v + n_e(r-1) + n_i\frac{(r-2)(r-3)}{2} \\ &= (n+1)^2 + 2n(n+1)(r-1) + n^2\frac{(r-2)(r-3)}{2} \\ &= \frac{r^2-r+4}{2}n^2 + 2rn + 1 = O\left(\frac{r^2n^2}{2}\right).\end{aligned}$$

Therefore, the total number of degrees of freedom for a serendipity space is asymptotically about half the size of that for a tensor product space of the same order r, as shown in Table 2.

We report the L^2-errors and the orders of the convergence of the spaces $\mathbb{P}_{r,r}$, $\mathcal{S}_r$ and $\mathcal{DS}_r$ for $r = 2,3,4,5$ on mesh sequence $\mathcal{T}_h^1$ in Table 3. The errors and convergence rates of the same tests in the H^1-seminorm are presented in Table 4. Since $\mathcal{T}_h^1$ is a sequence of square meshes, the direct serendipity space $\mathcal{DS}_r$ and the regular serendipity space $\mathcal{S}_r$ coincide on $\mathcal{T}_h^1$, despite the fact that they are implemented in different ways. All three families of spaces show an $(r+1)$-st order convergence in the L^2-norm and an r-th order convergence in the H^1-seminorm, as we should expect. The errors for $\mathbb{P}_{r,r}$ are smaller than that for $\mathcal{DS}_r = \mathcal{S}_r$, but $\mathbb{P}_{r,r}$ uses many more degrees of freedom (see Table 2).

Table 2: A comparison of the global number of degrees of freedom of $\mathbb{P}_{r,r}$, $\mathcal{S}_r$, and $\mathcal{DS}_r$ on an $n \times n$ mesh.

	$\dim(\mathbb{P}_{r,r})$	$\dim(\mathcal{S}_r) = \dim(\mathcal{DS}_r)$
$r = 2$	$4n^2 + 4n + 1$	$3n^2 + 4n + 1$
$r = 3$	$9n^2 + 6n + 1$	$5n^2 + 6n + 1$
$r = 4$	$16n^2 + 8n + 1$	$8n^2 + 8n + 1$
$r = 5$	$25n^2 + 10n + 1$	$12n^2 + 10n + 1$

Table 3: L^2-errors and convergence rates for $\mathbb{P}_{r,r}$, $\mathcal{DS}_r$, and $\mathcal{S}_r$ spaces on square meshes.

	$r = 2$		$r = 3$		$r = 4$		$r = 5$	
n	error	rate	error	rate	error	rate	error	rate
$\mathbb{P}_{r,r}$ on $\mathcal{T}_h^1$ meshes								
8	2.451e-04	2.99	5.564e-06	3.99	1.054e-07	4.99	1.688e-09	6.00
12	7.282e-05	2.99	1.101e-06	4.00	1.389e-08	5.00	1.483e-10	6.00
16	3.075e-05	3.00	3.486e-07	4.00	3.298e-09	5.00	2.640e-11	6.00
24	9.116e-06	3.00	6.890e-08	4.00	4.344e-10	5.00	2.420e-12	5.89
$\mathcal{S}_r = \mathcal{DS}_r$ on $\mathcal{T}_h^1$ meshes								
8	2.457e-04	2.99	1.805e-05	4.09	1.422e-06	5.01	6.440e-08	5.93
12	7.289e-05	3.00	3.497e-06	4.05	1.870e-07	5.00	5.739e-09	5.96
16	3.076e-05	3.00	1.099e-06	4.02	4.437e-08	5.00	1.027e-09	5.98
24	9.118e-06	3.00	2.161e-07	4.01	5.841e-09	5.00	9.049e-11	5.99

Tables 5–6 show the errors (in the L^2 and H^1-seminorms, respectively) and the orders of convergence for the trapezoidal mesh sequence $\mathcal{T}_h^2$. The tensor product space $\mathbb{P}_{r,r}$ achieves the expected convergence rates on trapezoidal meshes. The direct serendipity space $\mathcal{DS}_r$ retains an $(r+1)$-st order of convergence in the L^2 norm and an r-th order convergence in the H^1-seminorm, as Theorem 4.6 predicts. Meanwhile, the regular serendipity spaces $\mathcal{S}_r$, $r = 2, 3, 4, 5$, have worse than optimal convergence rates in both norms as Arnold, Boffi and Falk observed in [12]. The errors and convergence rates for $\mathbb{P}_{r,r}$, $\mathcal{DS}_r$, and $\mathcal{S}_r$ on mesh sequence $\mathcal{T}_h^3$ are similar to those on $\mathcal{T}_h^2$, so we omit showing them here.

In Section 3, we showed that assembling a global linear system requires the inversion of a $4r \times 4r$ matrix on each element. Compared with $\mathbb{P}_{r,r}$, for which the degrees of freedom are precalculated and stored in a reference element and then mapped to each element, we may expect a longer assembly time for $\mathcal{DS}_r$. On the other hand, since $\mathcal{DS}_r$ results in smaller linear systems, time spent on solving linear systems could be saved.

In Figure 10, we measure wall clock times of assembling the linear systems and solving them with a sparse direct solver (UMFPACK [19]) or with a Jacobi preconditioned conjugate gradient solver. We consider $\mathbb{P}_{r,r}$ and $\mathcal{DS}_r$ for $r = 2, 5$ on $n \times n$ $\mathcal{T}_h^2$ meshes. The tests are conducted on a

Table 4: H^1-seminorm errors and convergence rates for $\mathbb{P}_{r,r}$, $\mathcal{DS}_r$, and $\mathcal{S}_r$ spaces on square meshes.

	$r=2$		$r=3$		$r=4$		$r=5$	
n	error	rate	error	rate	error	rate	error	rate
$\mathbb{P}_{r,r}$ on $\mathcal{T}_h^1$ meshes								
8	1.276e-02	2.00	4.233e-04	3.00	1.047e-05	4.00	2.066e-07	5.00
12	5.673e-03	2.00	1.255e-04	3.00	2.070e-06	4.00	2.723e-08	5.00
16	3.191e-03	2.00	5.295e-05	3.00	6.549e-07	4.00	6.462e-09	5.00
24	1.418e-03	2.00	1.569e-05	3.00	1.294e-07	4.00	8.511e-10	5.00
$\mathcal{S}_r = \mathcal{DS}_r$ on $\mathcal{T}_h^1$ meshes								
8	1.285e-02	2.02	1.537e-03	3.05	1.141e-04	3.99	5.201e-06	4.99
12	5.690e-03	2.01	4.507e-04	3.03	2.261e-05	3.99	6.856e-07	5.00
16	3.197e-03	2.00	1.894e-04	3.01	7.164e-06	4.00	1.628e-07	5.00
24	1.420e-03	2.00	5.597e-05	3.01	1.416e-06	4.00	2.144e-08	5.00

Table 5: L^2-errors and convergence rates for $\mathbb{P}_{r,r}$, $\mathcal{DS}_r$, and $\mathcal{S}_r$ spaces on trapezoidal meshes.

	$r=2$		$r=3$		$r=4$		$r=5$	
n	error	rate	error	rate	error	rate	error	rate
$\mathbb{P}_{r,r}$ on $\mathcal{T}_h^2$ meshes								
8	3.329e-04	2.99	9.740e-06	3.99	2.382e-07	4.99	5.076e-09	5.99
12	9.888e-05	2.99	1.928e-06	3.99	3.142e-08	5.00	4.462e-10	6.00
16	4.176e-05	3.00	6.107e-07	4.00	7.459e-09	5.00	7.946e-11	6.00
24	1.238e-05	3.00	1.207e-07	4.00	9.827e-10	5.00	6.979e-12	6.00
$\mathcal{S}_r$ on $\mathcal{T}_h^2$ meshes								
8	5.714e-04	2.92	4.844e-04	2.89	2.612e-05	3.72	2.005e-06	4.13
12	1.731e-04	2.94	1.482e-04	2.92	6.084e-06	3.59	3.884e-07	4.05
16	7.409e-05	2.95	6.383e-05	2.93	2.265e-06	3.43	1.234e-07	3.99
24	2.254e-05	2.94	1.963e-05	2.91	5.984e-07	3.28	2.516e-08	3.92
32	9.799e-06	2.90	8.635e-06	2.85	2.408e-07	3.16	8.342e-09	3.84
48	3.127e-06	2.82	2.825e-06	2.76	6.875e-08	3.09	1.850e-09	3.71
64	1.440e-06	2.70	1.332e-06	2.61	2.862e-08	3.05	6.644e-10	3.56
$\mathcal{DS}_r$ on $\mathcal{T}_h^2$ meshes								
8	3.492e-04	3.00	3.897e-05	4.07	2.187e-06	5.00	8.896e-08	5.96
12	1.036e-04	3.00	7.457e-06	4.08	2.889e-07	4.99	7.870e-09	5.98
16	4.373e-05	3.00	2.313e-06	4.07	6.868e-08	4.99	1.404e-09	5.99
24	1.296e-05	3.00	4.469e-07	4.05	9.058e-09	5.00	1.235e-10	6.00

Table 6: H^1-seminorm errors and convergence rates for $\mathbb{P}_{r,r}$, $\mathcal{DS}_r$, and $\mathcal{S}_r$ spaces on trapezoidal meshes.

	$r = 2$		$r = 3$		$r = 4$		$r = 5$	
n	error	rate	error	rate	error	rate	error	rate
$\mathbb{P}_{r,r}$ on $\mathcal{T}_h^2$ meshes								
8	1.734e-02	2.00	7.206e-04	2.99	2.310e-05	3.99	6.083e-07	4.99
12	7.710e-03	2.00	2.139e-04	3.00	4.570e-06	4.00	8.021e-08	5.00
16	4.337e-03	2.00	9.027e-05	3.00	1.447e-06	4.00	1.904e-08	5.00
24	1.928e-03	2.00	2.676e-05	3.00	2.859e-07	4.00	2.509e-09	5.00
$\mathcal{S}_r$ on $\mathcal{T}_h^2$ meshes								
8	2.413e-02	1.94	1.834e-02	1.90	1.818e-03	2.65	1.537e-04	3.18
12	1.105e-02	1.93	8.572e-03	1.88	6.582e-04	2.51	4.483e-05	3.04
16	6.432e-03	1.88	5.091e-03	1.81	3.345e-04	2.35	1.945e-05	2.90
24	3.104e-03	1.80	2.560e-03	1.70	1.360e-04	2.22	6.370e-06	2.75
32	1.920e-03	1.67	1.643e-03	1.54	7.378e-05	2.12	3.029e-06	2.58
48	1.043e-03	1.50	9.409e-04	1.37	3.190e-05	2.07	1.140e-06	2.41
64	7.097e-04	1.34	6.602e-04	1.23	1.776e-05	2.03	5.953e-07	2.26
$\mathcal{DS}_r$ on $\mathcal{T}_h^2$ meshes								
8	1.836e-02	2.01	2.517e-03	3.02	1.625e-04	3.99	7.384e-06	4.99
12	8.143e-03	2.00	7.400e-04	3.02	3.216e-05	4.00	9.757e-07	4.99
16	4.577e-03	2.00	3.109e-04	3.01	1.018e-05	4.00	2.318e-07	5.00
24	2.033e-03	2.00	9.170e-05	3.01	2.012e-06	4.00	3.056e-08	5.00

single computer with four Intel(R) Core(TM) i5-4570 CPUs at 3.20GHz and 8GB total memory (RAM). The code uses the deal.II library [20].

Solid lines in Figure 10 show the time cost of the assembly routines and dashed lines represent the linear solvers costs. Obviously, the slope versus $\log(n)$ for the linear solvers are steeper than the slope for the assembly routines. The assembly times depend linearly on the size of the problem in log scale. When $r = 2$, $\mathcal{DS}_2$ (purple line) takes a longer time to assemble the linear system compared with the pre-calculated $\mathbb{P}_{2,2}$ (blue line), as we expected. However, when $r = 5$, assembling $\mathcal{DS}_5$ (blue line) is actually quicker than assembling $\mathbb{P}_{5,5}$ (purple line). We attribute this to the fact that there are fewer degrees of freedom on each element.

The time costs of the conjugate gradient solver for the two finite elements (yellow and cyan lines) are very close when $r = 2$, and both of them are better than their sparse direct counterparts (orange and green lines). The sparse direct solver is quicker for the $\mathcal{DS}_2$ stiffness matrix than for the $\mathbb{P}_{2,2}$ stiffness matrix. But when $r = 5$, solving a $\mathcal{DS}_5$ system is always cheaper than solving a $\mathbb{P}_{5,5}$ system. The sparse direct solvers (orange and green lines) perform better than their conjugate gradient counterparts (yellow and cyan lines) when $r = 5$ and n is small, but the gap decreases as n increases.

The time costs of the other routines in the test program are either negligible (e.g., system setup)

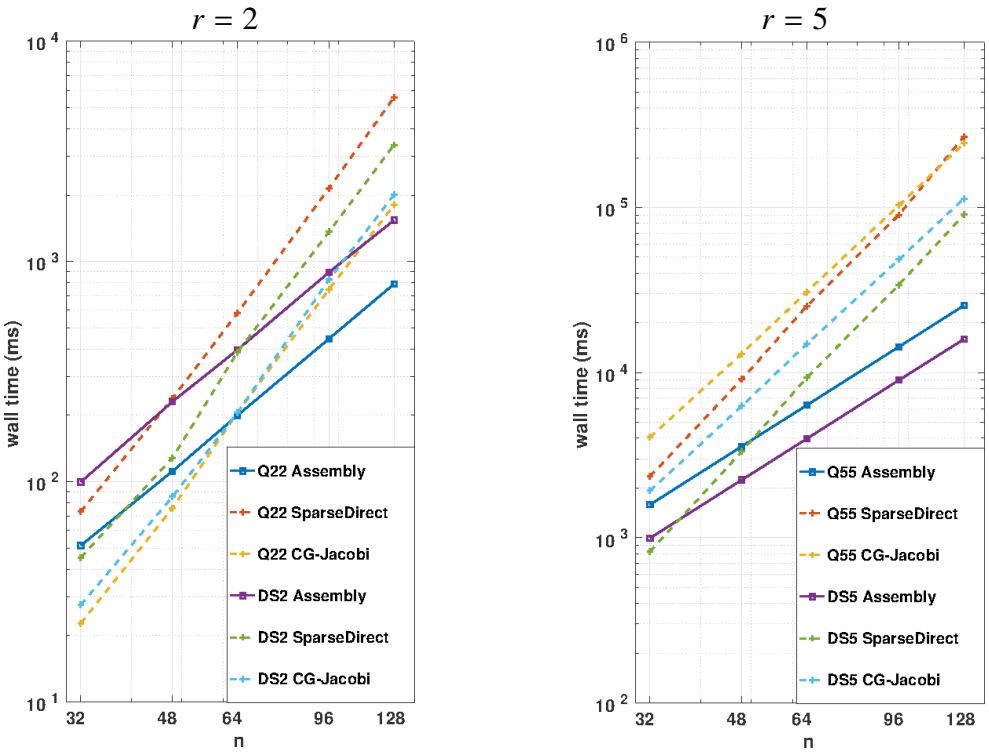


Figure 10: Wall clock times for assembling the linear systems and for solving them with a sparse direct solver or with a Jacobi preconditioned conjugate gradient solver. Results are for $\mathbb{P}_{r,r}$ and $\mathcal{DS}_r$, $r = 2$ (left) and $r = 5$ (right) on $n \times n$ $\mathcal{T}_h^2$ meshes. The time and n are shown in log scale.

Table 7: Total wall clock times (in milliseconds) for $\mathcal{DS}_r$ and $\mathbb{P}_{r,r}$ with different solvers when $r = 2, 5$.

	$\mathcal{DS}_2$		$\mathbb{P}_{2,2}$	
n	Sparse Direct	CG-Jacobi	Sparse Direct	CG-Jacobi
32	273.114	254.942	193.127	145.379
48	701.488	595.363	503.456	351.505
64	1297.63	1106.55	1056.56	707.311
96	3400.92	2918.73	3223.02	1847.66
128	6945.08	5779.14	7459.29	3748.74
	$\mathcal{DS}_5$		$\mathbb{P}_{5,5}$	
n	Sparse Direct	CG-Jacobi	Sparse Direct	CG-Jacobi
32	2468.98	3578.92	4213.28	5982.87
48	7428.24	10392.1	13308.0	17283.1
64	15908.5	21548.8	32774.2	38440.1
96	49768.3	63886.7	106822.	119959.
128	116427.	139270.	299793.	273998.

or follow a similar pattern as the assembly routine (e.g., post processing and output). Table 7 gives the total wall clock times for $\mathcal{DS}_r$ and $\mathbb{P}_{r,r}$ with different solvers when $r = 2, 5$. When $r = 2$, $\mathcal{DS}_2$ takes more time than $\mathbb{P}_{2,2}$ if we use the conjugate gradient solver. On the other hand, if we choose the sparse direct solver, $\mathbb{P}_{2,2}$ is quicker than $\mathcal{DS}_2$ when n is small, but $\mathcal{DS}_2$ will take less time if we keep refining the problem. As we might expect, $\mathcal{DS}_5$ always takes less time than $\mathbb{P}_{5,5}$ when $r = 5$. If we take a closer look, when $n = 128$, $\mathcal{DS}_5$ take about half of the time for $\mathbb{P}_{5,5}$. This observation is consistent with the number of degrees of freedom in Table 2.

We close by remarking that the time cost for the assembly routine can be scaled nearly perfectly in parallel since it basically involves only local computations. Therefore, reducing the global number of degrees of freedom, even perhaps at the expense of a slightly more expensive assembly, is worthwhile.

[1] G. Strang, G. J. Fix, An analysis of the finite element method, Series in Automatic Computation, Prentice-Hall, Englewood Cliffs, New Jersey, 1973.

[2] P. G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.

[3] S. C. Brenner, L. R. Scott, The Mathematical Theory of Finite Element Methods, Springer-Verlag, New York, 1994.

[4] T. Dupont, J. Hoffman, C. Johnson, R. C. Kirby, M. G. Larson, A. Logg, L. R. Scott, The fenics project, Tech. Rep. 2003–21, Chalmers Finite Element Center, Chalmers University of Technology, Goteborg, Sweden (2003).

[5] I. Hibbitt, Karlsson & Sorensen, ABAQUS/Standard User's Manual (2001).

[6] D. N. Arnold, A. Logg, Periodic table of the finite elements, SIAM News 47 (9).

[7] F. Brezzi, J. Douglas, Jr., L. D. Marini, Two families of mixed elements for second order elliptic problems, Numer. Math. 47 (1985) 217–235.

[8] D. N. Arnold, G. Awanou, The serendipity family of finite elements, Foundations of Computational Mathematics 11 (3) (2011) 337–344.

[9] D. N. Arnold, G. Awanou, Finite element differential forms on cubical meshes, Math. Comp. 83 (2014) 1551–1570.

[10] N.-S. Lee, K.-J. Bathe, Effects of element distortions on the performance of isoparametric elements, Intl. J. Numer. Meth. Engineering 36 (20) (1993) 3553–3576.

[11] V. N. Kaliakin, Introduction to approximate solution techniques, numerical modeling, and finite element methods, CRC Press, 2001.

[12] D. N. Arnold, D. Boffi, R. S. Falk, Approximation by quadrilateral finite elements, Math. Comp. 71 (239) (2002) 909–922.

[13] A. Rand, A. Gillette, C. Bajaj, Quadratic serendipity finite elements on polygons using generalized barycentric coordinates, Math. Comp. 83.

[14] T. Arbogast, M. R. Correa, Two families of H(div) mixed finite elements on quadrilaterals of minimal dimension, SIAM J. Numer. Anal. 54 (6) (2016) 3332–3356, DOI 10.1137/15M1013705.

[15] V. Girault, P. A. Raviart, Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms, Springer-Verlag, Berlin, 1986.

[16] J. H. Bramble, S. R. Hilbert, Estimation of linear functionals on Sobolev spaces with applications to Fourier transforms and spline interpolation, SIAM J. Numer. Anal. 7 (1970) 112–124.

[17] T. Dupont, L. R. Scott, Polynomial approximation of functions in Sobolev space, Math. Comp. 34 (1980) 441–463.

[18] P. Grisvard, Elliptic Problems in Nonsmooth Domains, Pitman, Boston, 1985.

[19] T. A. Davis, Algorithm 832: UMFPACK v4.3—an unsymmetric-pattern multifrontal method, ACM Transactions on Mathematical Software 30 (2) (2004) 196–199.

[20] W. Bangerth, D. Davydov, T. Heister, L. Heltai, G. Kanschat, M. Kronbichler, M. Maier, B. Turcksin, D. Wells, The `deal.II` library, version 8.4, Journal of Numerical Mathematics 24.