Validation of the Quality of the *A Posteriori* Error Estimator Based on Polynomial Preserving Recovery for Linear Elements

Zhimin Zhang*and Ahmed Naga
Department of Mathematics, Wayne State University

Abstract In this paper the quality of the error estimator based on the Polynomial Preserving Recovery (PPR) is investigated using the computer-based approach described in [1, 3]. Also, a comparison is made between the performance of the error estimator based on the PPR and the one based on the Superconvergence Patch Recovery (SPR).

Key Words. finite element method, *a posteriori* error estimator, least-squares fitting, SPR, PPR, superconvergence, effectivity index, robustness index

AMS Subject Classification. 65N30, 65N15, 65N12, 65D10, 74S05, 41A10, 41A25

1 Introduction

Judging the error in a finite element approximation of a partial differential equation by using *a priori* error estimates is not reliable in most cases as these estimates have unknown constants. One way to overcome this problem is to use "very small" elements which, in practical problems, lead to large linear systems. The time cost for solving such systems is very high and the round off errors may decrease the accuracy of the approximation. A different approach is to use higher order elements, but this has many drawbacks. First, the implementation of higher order is very expensive. Secondly, the band width of the resulting linear systems increases, and hence, computational time cost gets higher. Finally, higher order elements require the problem solution to have a higher degree of regularity in order to achieve the expected *a priori* estimated accuracy. In many practical situations the regularity of the problem solution is unknown or is not enough to use higher order elements.

Another strategy is adaptation where an initial mesh is used to get an initial finite element approximation which is postprocessed to "measure" the error by using an *a posteriori* error estimator. An *a posteriori* error estimator, if "accurate", identifies parts of the mesh where

*This research was partially supported by the National Science Foundation grants DMS-0074301, DMS-0079743, and INT-0196139.

the error in the solution does not meet a prescribed tolerance. Such parts are then refined to obtain a new mesh that is used to get a new solution. This process is repeated till the error is within the specified tolerance. In many aspects, this strategy has been proved to be better than previous approaches on condition that the *a posteriori* error estimator is accurate. Because of that *a posteriori* error estimators have attracted many researchers and became the focus of intensive investigations; see, for example, [1, 3, 6, 7, 8, 9, 10, 11, 14, 15, 17]. Generally speaking, error estimators can be classified under two categories. The first category contains the residual type estimators, as in [6], and the second one contains recovery type estimators, as in [15]. In a recovery type estimator, a recovery operation uses the finite element solution (or its gradient) to build another gradient as in [16, 13] (or another solution as in [10, 11]).

Among gradient recovery techniques, the Superconvergence Patch Recovery (SPR) has been considered to be the best for many years. The Polynomial Preserving Recovery (PPR) is a new gradient recovery technique, see [13], that can be used to recover a superconvergent gradient under some mild conditions imposed on the mesh as was shown in [14]. This motivated the use of the PPR-recovered gradient in building an asymptotically exact *a posteriori* error estimator (the PPR estimator). By testing it on a set of benchmarks, the PPR estimator was found to be as good as or better than the estimator based on the SPR-recovered gradient (the SPR estimator; also known as the ZZ-SPR estimator). However, benchmark computations may lead to inaccurate conclusions as was shown through examples in [5]. A more accurate methodology to study the performance of an *a posteriori* error estimator was proposed in [4]. Using this methodology, many known estimators were studied in [3, 5] where it was found that the SPR estimator is the most robust.

The goal of this paper is to use this methodology to study the PPR estimator and to compare it with the SPR estimator. Since the PPR is still in its development phase, this paper considers only linear elements. Quadratic elements are currently under investigations and their results will be available soon.

2 Preliminaries

2.1 Model problem

The model problem considered in this study is the heat conduction in an orthotropic medium governed by the elliptic boundary value problem

$$\begin{cases} -Lu = -\nabla(\mathcal{D}\nabla u) = f & \text{in } \Omega \\ \boldsymbol{n}\cdot(\mathcal{D}\nabla u) = g & \text{on } \Gamma_N \\ u = 0 & \text{on } \Gamma_D \end{cases} \tag{2.1}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded domain with Lipschitz boundary $\partial\Omega = \overline{\Gamma_N} \cup \overline{\Gamma_D}$, the boundary segments Γ_N and Γ_D are assumed to be disjoint, $\boldsymbol{n}$ is the unit outward normal vector to $\partial\Omega$, and $\mathcal{D}$ is the thermal conductivity matrix that is constant all over Ω, symmetric, and positive definite. If the orientation of the material orthotropy principal axes with respect to the problem coordinate system is θ, then

$$\mathcal{D} = \begin{bmatrix} (\frac{d+1}{2}) + (\frac{d-1}{2})\cos(2\theta) & (\frac{d-1}{2})\sin(2\theta) \\ (\frac{d-1}{2})\sin(2\theta) & (\frac{d+1}{2}) - (\frac{d-1}{2})\cos(2\theta) \end{bmatrix}. \tag{2.2}$$

Without loss of generality, we may assume that the principal thermal conductivities are 1 and $d \geq 1$. If $\Gamma_N = \partial\Omega$, the compatibility condition $\int_\Omega f + \int_{\partial\Omega} g = 0$ must be satisfied and the condition $\int_\Omega u = 0$ is used to get a unique solution. As usual, $W_p^m(\Omega)$ and $H^m(\Omega)$ are the classical Sobolev spaces equipped with the norms $\| \quad \|_{m,p,\Omega}$, and $\| \quad \|_{m,\Omega}$, respectively, and the seminorms $| \quad |_{m,p,\Omega}$, and $| \quad |_{m,\Omega}$, respectively. The set of all polynomials defined on $\Omega' \subseteq \mathbb{R}^2$ of total degree $\leq r$ is denoted by $P_r(\Omega')$.

The variational form of this problem is to find $u \in V$ such that

$$B(u, v) = l(v) \text{ for all } v \in V, \tag{2.3}$$

where

$$V = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D\},$$
$$B(u, v) = \int_\Omega \mathcal{D}\nabla u \nabla v,$$

and

$$l(v) = \int_\Omega fv + \int_{\Gamma_N} gv.$$

Let $\mathcal{T}_h$ be a triangulation of Ω. For linear elements, the finite element space $S_h \subset V$ associated with $\mathcal{T}_h$ is defined by

$$S_h = \{v \in V : v \in P_1(\tau) \text{ for every triangle } \tau \in \mathcal{T}_h\}.$$

The subspace $S_h^0 \subseteq S_h$ is defined by

$$S_h^0 = \{v \in S_h : v|_{\Gamma_D} = 0\}.$$

If $\Gamma_N = \partial\Omega$, then S_h^0 is taken to be S_h. The finite element solution of the variational problem (2.3) is u_h, where

$$B(u_h, v) = l(v) \text{ for all } v \in S_h^0. \tag{2.4}$$

For $\Omega' \subseteq \Omega$, we define the space $S_h(\Omega')$ and the bilinear operator $B_{\Omega'}$, where

$$S_h(\Omega') = \{v|_{\Omega'} : v \in S_h\}$$

and

$$B_{\Omega'}(u,v) = \int_{\Omega'} \mathcal{D}\nabla u \nabla v.$$

2.2 Definitions of the SPR and the PPR

A C^0 finite element solution u_h has discontinuous gradient ∇u_h. In an attempt to better approximate ∇u, the SPR, as well as the PPR, constructs a continuous gradient $\mathrm{G_h}u_h \in S_h \times S_h$; called the recovered gradient. As it is known, any function in S_h is completely defined by its nodal values. So, it suffices to define $\mathrm{G_h}u_h$ at mesh nodes. This definition depends on the node location in $\overline{\Omega}$.

The definition of SPR-recovered gradient at a mesh node z is as follows (see [16] for more details).

- If $z \in \Omega$, we use a patch ω_z consisting of the triangles attached to z as shown in Fig. 1(a). To recover the $x-$derivative at z, we find a polynomial $p_x \in P_1(\omega_z)$ that best fits, in least squares sense, $\partial_x u_h$ at the triangles centroids in ω_z. The recovered $x-$derivative at z is defined to be $p_x(z)$. Similarly, we can define the recovered $y-$derivative at z.

- If $z \in \partial\Omega$, let $z_1, z_2, \ldots, z_{N_{z,s}}$ denote the mesh nodes in Ω that are directly connected to z. Let ω_i be the patch associated with z_i for $i = 1, 2, \ldots, N_{z,s}$ and let $p_{x,i} \in P_1(\omega_i)$ be the polynomial that best fits $\partial_x u_h$ sampled at triangles centroids in ω_i. Again, the patch ω_i consists of the triangles attached to z_i. The recovered $x-$derivative at z is defined to be $\dfrac{1}{N_{z,s}} \displaystyle\sum_{i=1}^{N_{z,s}} p_{x,i}(z)$. Similarly, we can define the recovered $y-$derivative at z.

- If $z \in \partial\Omega$ with no attached internal nodes, the recovered gradient at z is defined to be $\nabla u_h(z)$.

The construction of the PPR-recovered gradient at mesh nodes proceeds in two stages. In the first stage mesh nodes in Ω are considered while mesh nodes on $\partial\Omega$ are considered in the second stage.

- **Stage 1**

  As in SPR, we use a patch ω_z consisting of the triangles attached to z. To recover the gradient at z, we find a polynomial $p \in P_2(\omega_z)$ that best fits u_h sampled at the mesh nodes in ω_z, in least-squares sense, and $\mathrm{G_h}u_h(z)$ is defined to be $\nabla p(z)$. To get p, ω_z must contain at least 6 mesh nodes. If this is not the case and if ω_z does not share any edges with $\partial\Omega$, ω_z is extended by adding every triangle sharing an edge with ω_z as shown in

Fig. 1(b). If ω_z has less than 6 mesh nodes with some of its edges on $\partial\Omega$, recovering the gradient at z is delayed to the Stage 2. For an example of such case, see Fig. 2(a), the Criss Cross pattern.

- **Stage 2**

  Basically this stage uses the gradient recovered in the first stage and linear extrapolation to complete the gradient recovery at the rest of the mesh nodes. Let $\mathcal{N}_{h,0}$ be the set of the mesh nodes left in Stage 1 without recovery, which includes those on $\partial\Omega$. The gradient recovery is completed in a finite number of iterations where every iteration proceeds as follows. The iteration starts by defining $\mathcal{T}_{h,0} \subset \mathcal{T}_h$ where a mesh triangle $\tau \in \mathcal{T}_{h,0}$ if and only if $\mathrm{G_h}u_h$ is defined at each of its vertices. For $z \in \mathcal{N}_{h,0}$, let ω_z be the patch consisting of the triangles attached to z. We have two cases.

  1. The patch ω_z has common edges with triangles $\tau_1, \tau_2, \ldots, \tau_{N_{z,p}}$ in $\mathcal{T}_{h,0}$. Let ω_{τ_i} denote the union of the triangles in $\mathcal{T}_{h,0}$ that have common edges with τ_i along with τ_i. Note that ω_{τ_i} has at least 4 nodes for which $\mathrm{G_h}u_h$ is well-defined. Using least squares, we can find the linear polynomial $q_{x,i} \in P_1(\omega_{\tau_i})$ that best fits the x-components of $\mathrm{G_h}u_h$ at the mesh nodes in ω_{τ_i}. The recovered x-derivative at z is defined to be $\dfrac{1}{N_{z,p}} \displaystyle\sum_{i=1}^{N_{z,p}} q_{x,i}(z)$. The recovered y-derivative at z is defined in a similar way.
  2. The patch ω_z has no common edges with triangles in $\mathcal{T}_{h,0}$. In this case z is left for another iteration where it is add to $\mathcal{N}_{h,1}$, a set taken to be empty at the beginning of the iteration.

  After going over all the nodes in $\mathcal{N}_{h,0}$, if $\mathcal{N}_{h,1}$ is empty, we are done; otherwise we set $\mathcal{N}_{h,0} = \mathcal{N}_{h,1}$ and start another iteration.

Fig. 2(a) shows some examples that explain how iterations proceed to construct the PPR-recovered gradient. Nodes labelled with 0 are those processed in stage 1. Nodes Processed in stage 2 are labelled 1,2,3, or 4, depending on the iteration in which they are processed.

To better understand the extrapolation procedure used in stage 2, let us have an example. As depicted in Fig. 2(b), we want to recover the gradient at $z \in \partial\Omega$. The triangles attached to z are τ_1, τ_2, and τ_3. The triangles τ_1 and τ_3 do not have common edges with triangles in $\mathcal{T}_{h,0}$, but τ_2 does as it shares a common edge with $\tau_4 \in \mathcal{T}_{h,0}$. To use τ_4 solely in extrapolating the gradient at z, τ_2 and τ_4 should form a convex quadrilateral; otherwise extrapolation may be unstable. To avoid that, the strategy used in step 2 proposes the use of the triangles τ_5 and τ_6 along with τ_4 to carry out the extrapolation. The triangles τ_5 and τ_6 are chosen because

they are in $\mathcal{T}_{h,0}$ and every one of them share an edge with τ_4. Next, we compute two linear polynomials that best fit, in least squares sense, the components of the recovered gradient at the mesh nodes in τ_4, τ_5 and τ_6. Finally, the obtained linear polynomials are evaluated at z to get the recovered gradient at z.

Remark 2.1. The definition of the PPR-recovered gradient, especially for the nodes on $\partial\Omega$, is different from the one adopted in [13, 14]. After testing many definitions using the computer-based theory, it was found that the proposed definition leads to more robust error indicators as it will be explained later.

2.3 Definition of the *A Posteriori* Error Estimator

The recovered gradient $\mathrm{G_h}u_h$ can be used in constructing an *a posteriori* error estimator defined by

$$\varepsilon(u_h, \Omega, f) = \sqrt{\sum_{\tau \in \mathcal{T}_h} \eta_\tau^2}$$

where τ is a mesh triangle and η_τ is the element error indicator defined by

$$\eta_\tau = \| \mathrm{G_h}u_h - \nabla u_h \|_{L^2(\tau)}.$$

The reliability of *a posteriori* error estimator is measured using the effectivity index $\kappa(u_h, \Omega, f)$ with

$$\kappa(u_h, \Omega, f) = \frac{\varepsilon(u_h, \Omega, f)}{\|\nabla(u - u_h)\|_{L^2(\Omega_h)}}.$$

If $\mathrm{G_h}u_h$ is superconvergent to the ∇u, then $\kappa(u_h, \Omega, f) \to 1$ as $h \to 0$. As it is well known, superconvergence is a delicate property that requires the mesh and the solution to satisfy certain conditions which may be hard to achieve in practice. Therefore, varying u over a set $\mathcal{U}$ of solutions of interest and varying $\mathcal{T}_h$ over a set $\mathcal{M}$ of meshes that are used in computations, it is realistic to have

$$0 < \underline{\kappa} \leq \lim_{h \to 0} \kappa(u_h, \Omega, f) \leq \overline{\kappa} < \infty.$$

To measure the deviation of the effectivity index from 1, we use the robustness index $\mathcal{R}$ defined by

$$\mathcal{R} = \max\{\overline{\kappa} - 1, \frac{1}{\underline{\kappa}} - 1\}.$$

The smaller the value of $\mathcal{R}$, the more accurate the estimator is. In general, $\mathcal{R}$, as well as the bounds $\underline{\kappa}$ and $\overline{\kappa}$, depends on $\mathcal{M}$, $\mathcal{U}$, and the definition of the error estimator. Trying to find the bounds $\underline{\kappa}$ and $\overline{\kappa}$ analytically is tedious and inaccurate. Also, trying to estimate them using benchmarks is not reliable as benchmarks represent special cases. Moreover, $\underline{\kappa}$ and $\bar{\kappa}$ provide

only a global information about the error estimator and can not help too much in judging the accuracy of the local error indicators. Indeed, it is the local error indicator that is more important in adaptive design.

A computer-based theory was presented in [4] to find the asymptotic bounds for the effectivity index associated with an error indicator over an interior element patch. This methodology was extended in [5] to handle patches adjacent to $\partial\Omega$. The next section is devoted for a brief review of this theory.

3 Review of the Computer-Based Theory for Internal Patches

Let $\Omega_2 \subset \Omega_1 \subset \Omega_0 \subset \Omega$. The following three assumptions are essential for the computer-based theory.

Assumption T1. (*Local uniformity of the mesh*). The mesh on Ω_0 is *uniform*, i.e., it is obtained by tessellating Ω_0 using a basic mesh cell of size h as shown in Fig. 3. The centroids of the cells are denoted by c_α, $\alpha \in \mathcal{I}_0$, where $\mathcal{I}_0$ is a suitable indexing set and the cell centered at c_α is denoted by $\mathcal{C}(c_\alpha, h)$. These cells are disjoint and constitute a partition of Ω_0. Furthermore, there are indexing sets $\mathcal{I}_2 \subset \mathcal{I}_1 \subset \mathcal{I}_0$ such that

$$\Omega_k = \bigcup_{\alpha \in \mathcal{I}_k} \mathcal{C}(c_\alpha, h)$$

for $k = 1, 2$. The subdomain Ω_k is convex and regular in the sense that the diameter of the largest ball inscribed inside Ω_k is at least CD_k where $D_k = \operatorname{diam} \Omega_k$ and C is a positive constant. For every $\alpha \in \mathcal{I}_0, \mathcal{C}(c_\alpha, h)$ is an image of a reference cell $\mathcal{C}$ under an invertible transformation $F_\alpha : \mathcal{C} \to \mathcal{C}(c_\alpha, h)$ where

$$F_\alpha(\hat{x}, \hat{y}) = (h\hat{x}, h\hat{y}) + c_\alpha \quad \forall (\hat{x}, \hat{y}) \in \mathcal{C}.$$

The Reference cell can be any convex domain that tessellates the plane, like hexagon, rectangle, or rhombus. For simplicity, $\mathcal{C}$ is a square of side length 2, but the theory can be generalized to cover the other shapes. The mesh on Ω_0 is *translation invariant*, i.e., a mesh $\mathcal{T}_\mathcal{C}$ is constructed on the reference cell $\mathcal{C}$ and the mesh on $\mathcal{C}(c_\alpha, h)$, $\alpha \in \mathcal{I}_0$, is the image of $\mathcal{T}_\mathcal{C}$ under F_α. In order to have a conforming finite element partition on Ω_0, $\mathcal{T}_\mathcal{C}$ must satisfy the following condition: the nodes on opposite horizontal (vertical) edges of $\mathcal{C}$ are symmetrically distributed with respect to $\hat{x}$-axis ($\hat{y}$-axis). A partition on $\mathcal{C}$ that satisfies this condition will be called *admissible partition*.

Assumption T2. (*Regularity of the exact solution*). The exact solution u must satisfy the conditions

- $u \in W^3_\infty(\Omega_0)$, and
- there exists a positive constant μ_0 such that

$$\inf_{(x,y)\in\Omega_0} \sum_{|\beta|=2} |D^\beta u(x,y)|^2 \geq \mu_0^2 > 0.$$

This rules out trivial cases when the second derivatives of u vanish identically.

Assumption T3. (L^2*-convergence of the error in* Ω_1). There exists $\epsilon \in (0, \frac{1}{2})$ such that

$$\|u - u_h\|_{L^2(\Omega_1)} \leq Ch^{2-\epsilon} D_1$$

where C is a positive constant that depends on u, but independent of h and D_1. This is to avoid the influence of outside effects such as singularities around re-entrant corners. For that, the mesh must be sufficiently refined in the neighborhood of such corners and this is only condition imposed on the mesh outside Ω_0.

Remark 3.1. The error inside any mesh element has two components. The first one is the local error attributed to the residuals in the element and its neighbors while the second one is the pollution error resulting from residuals in the rest of the mesh, especially those in the neighborhood of singularities. An element error indicator estimates only the local error and can not capture the pollution error. Indeed, the existence of the pollution error overshadows the local error and deteriorates the accuracy of local error indicators. To avoid that, Assumption T3 must be satisfied.

Assumption T2 enables us to approximate u by a quadratic polynomial $q \in P_2(\Omega_1)$. Specifically, q is taken to be the quadratic part of the Taylor series for u about the centroid of Ω_1. Let q_h be the orthogonal projection over $S_h(\Omega_1)$ obtained by solving the problem

$$\begin{cases} B_{\Omega_1}(q - q_h, v) & = 0 \qquad \forall v \in S_h(\Omega_1) \\ \int_{\Omega_1}(q - q_h) & = 0 \end{cases}. \tag{3.1}$$

The following theorem says that q_h is a "good" approximation of u_h. For a full proof of next theorem and other related details, see [1].

Theorem 3.2. *Suppose Assumptions T2 and T3 hold, let* $\sigma = \frac{2}{5}\epsilon$*, and let* q_h *be defined as in* (3.1)*. If*

$$ch^{\frac{1+\sigma}{2}} \leq D_2 \leq D_1 \leq Ch^{\frac{1+\sigma}{2}}$$

and

$$\operatorname{dist}(\partial\Omega_1, \partial\Omega_2) \geq Ch^{\sigma D_1}$$

for positive constants c and C independent of h, then

$$|\|\nabla(u-u_h)\|_{L^2(\Omega_2)} - \|\nabla(q-q_h)\|_{L^2(\Omega_2)}| \leq ch^{\sigma}\|\nabla(u-u_h)\|_{L^2(\Omega_2)}. \tag{3.2}$$

Let $\varepsilon(u_h, f, \Omega_2)$ denotes an error indicator constructed on Ω_2 for the true error $\|\nabla(u-u_h)\|_{L^2(\Omega_2)}$. Similarly, we construct the error indicator $\varepsilon(q_h, -Lq, \Omega_2)$ for the true error $\|\nabla(q-q_h)\|_{L^2(\Omega_2)}$. Intuitively, one would conjecture that this theorem implies

$$\lim_{h\to 0} \frac{\varepsilon(u_h, f, \Omega_2)}{\|\nabla(u-u_h)\|_{L^2(\Omega_2)}} = \lim_{h\to 0} \frac{\varepsilon(q_h, -Lq, \Omega_2)}{\|\nabla(q-q_h)\|_{L^2(\Omega_2)}}. \tag{3.3}$$

Actually, this conjecture is true provided that the error indicator is *stable*. Basically, an error indicator is stable if small changes in the finite element solution u_h and the data f result in small changes in the estimated error. Luckily, this requirement is satisfied by both recovery type and residual type error estimators as shown in [1].

Remark 3.3. It is important to note that the result in (3.3) is still true when Assumption T1 is not satisfied.

Let $\mathcal{Q}$ be the set of quadratic polynomials obtained by Taylor expansion of functions in $\mathcal{U}$ at the centroid of Ω_1. Equation (3.3) implies that the asymptotic range of $\kappa(u_h, f, \Omega_2)$, defined by the left hand side of (3.3), over $\mathcal{U}$ and $\mathcal{M}$ is the same as the asymptotic range of $\kappa(q_h, -Lq, \Omega_2)$, defined by the right hand side of (3.3), over $\mathcal{Q}$ and $\mathcal{M}$. However, the result in this form is not practical as we do not have access to q_h as $h \to 0$. The next step is to find a "good" approximation for q_h that is accessible as $h \to 0$.

To this end, we assume that the mesh on Ω_0 satisfies Assumption T1. We need to define the following set of subspaces and operators. The subspace of the periodic functions on $\mathcal{C}$ is defined by

$$H^{1,per}(\mathcal{C}) = \{v \in H^1(\mathcal{C}) : v(\hat{x}, -1) = v(\hat{x}, 1), v(1, \hat{y}) = v(-1, \hat{y}) \forall \hat{x}, \hat{y} \in (-1, 1)\}$$

which is equipped with the norm for the space $H^1(\mathcal{C})$. The corresponding finite element subspace constructed using the mesh $\mathcal{T}_{\mathcal{C}}$ is defined by

$$S^{per}(\mathcal{C}) = \{\hat{v} \in H^{1,per}(\mathcal{C}) : \hat{v}|_{\tau} \in P_1(\tau) \forall \tau \in \mathcal{T}_{\mathcal{C}}\}.$$

The projection operator $\Pi_{\mathcal{C}}^{per} : H^{1,per}(\mathcal{C}) \to S^{per}(\mathcal{C})$ is defined for each $\hat{u}$ in $H^{1,per}(\mathcal{C})$ such that $\Pi_{\mathcal{C}}^{per}\hat{u} = \hat{w}$ where $\hat{w} \in S^{per}(\mathcal{C})$ is satisfying the conditions

$$\begin{cases} B_{\mathcal{C}}(\hat{u} - \hat{w}, \hat{v}) = 0 & \forall \hat{v} \in S^{per}(\mathcal{C}) \\ \int_{\mathcal{C}}(\hat{u} - \hat{w}) = 0 & \end{cases}.$$

For $\alpha \in \mathcal{I}_0$, the space of the periodic functions on the cell $\mathcal{C}(c_\alpha, h)$ is defined by

$$H^{1,per}(\mathcal{C}(c_\alpha, h)) = \{\hat{v} \circ F_\alpha^{-1} : \hat{v} \in H^{1,per}(\mathcal{C})\}$$

and the finite element subspace constructed on $\mathcal{C}(c_\alpha, h)$ is defined by

$$S^{per}(\mathcal{C}(c_\alpha, h)) = \{\hat{v} \circ F_\alpha^{-1} : \hat{v} \in S^{per}(\mathcal{C})\}$$

The projection operator $\Pi^{per}_{\mathcal{C}(c_\alpha,h)} : H^{1,per}(\mathcal{C}(c_\alpha, h)) \to S^{per}(\mathcal{C}(c_\alpha, h))$ is defined for each u in $H^{1,per}(\mathcal{C}(c_\alpha, h))$ such that $\Pi^{per}_{\mathcal{C}(c_\alpha,h)} u = w$ where $w \in S^{per}(\mathcal{C}(c_\alpha, h))$ is satisfying the conditions

$$\left\{ \begin{array}{ll} B_{\mathcal{C}(c_\alpha,h)}(u - w, v) & = 0 \qquad \forall v \in S^{per}(\mathcal{C}(c_\alpha, h)) \\ \int_{\mathcal{C}(c_\alpha,h)}(u - w) & = 0 \end{array} \right. .$$

Note that F_α can be viewed as an invertible mapping from $H^1(\mathcal{C}(c_\alpha, h))$ to $H^1(\mathcal{C})$ in the following sense: for each $u \in H^1(\mathcal{C}(c_\alpha, h))$,

$$F_\alpha(u) = u \circ F_\alpha.$$

Using the above definitions, it is easy to prove that

$$\Pi^{per}_{\mathcal{C}(c_\alpha,h)} = F_\alpha \circ \Pi^{per}_{\mathcal{C}} \circ F_\alpha^{-1}.$$

Let $\hat{u} \in H^{1,per}(\mathcal{C})$, then we can define $u_\alpha \in H^{1,per}(\mathcal{C}(c_\alpha, h))$ by $u_\alpha = \hat{u} \circ F_\alpha^{-1}$. The conditions imposed on functions in $H^{1,per}(\mathcal{C})$ implies that the functions $u_\alpha, \alpha \in \mathcal{I}_0$, match on cells interfaces. Hence, piecing them together produces a function $u \in H^1(\Omega_0)$ defined by $u|_{\mathcal{C}(c_\alpha,h)} = u_\alpha$. The function u is called the periodic extension associated with $\hat{u}$ on Ω_0. The space of all periodic extensions over Ω_0 is denoted by $H^{1,per}(\Omega_0)$ and it is defined by

$$H^{1,per}(\Omega_0) = \{v \in H^1(\Omega_0) : v|_{\mathcal{C}(c_\alpha,h)} = \hat{v} \circ F_\alpha^{-1} \text{ with } \hat{v} \in H^{1,per}(\mathcal{C})\}.$$

The corresponding finite element subspace is defined by

$$S_h^{per}(\Omega_0) = \{v \in S_h(\Omega_0) : v|_{\mathcal{C}(c_\alpha,h)} = \hat{v} \circ F_\alpha^{-1} \text{ with } \hat{v} \in S^{per}(\mathcal{C})\}.$$

The periodic extension procedure described above can be represented by the invertible mapping $E^{per}_{\Omega_0} : H^{1,per}(\mathcal{C}) \to H^{1,per}(\Omega_0)$ such that for each $\hat{u} \in H^{1,per}(\mathcal{C})$, $(E^{per}_{\Omega_0}\hat{u})|_{\mathcal{C}(c_\alpha,h)} = \hat{u} \circ F_\alpha^{-1}$. The projection operator $\Pi^{per}_{\Omega_0} : H^{1,per}(\Omega_0) \to S_h^{per}(\Omega_0)$ is defined for each $u \in H^{1,per}(\Omega_0)$ such that $\Pi^{per}_{\Omega_0} u = w$ where $w \in S_h^{per}(\Omega_0)$ is satisfying the conditions

$$\left\{ \begin{array}{ll} B_{\Omega_0}(u - w, v) & = 0 \qquad \forall v \in S_h^{per}(\Omega_0) \\ \int_{\Omega_0}(u - w)dxdy & = 0 \end{array} \right. .$$

It is easy to verify that

$$\Pi^{per}_{\Omega_0} = E^{per}_{\Omega_0} \circ \Pi^{per}_{\mathcal{C}} \circ (E^{per}_{\Omega_0})^{-1}.$$

If $q_I \in S_h(\Omega_0)$ is the usual Lagrange interpolation of q, then it can be shown that $q - q_I$ is actually in $H^{1,per}(\Omega_0)$. Using this property, we can define a new approximation q_h^{asy} for q over Ω_0 where

$$q_h^{asy} = q_I + \Pi^{per}_{\Omega_0}(q - q_I).$$

The function q_h^{asy} is called the asymptotic finite element approximation of q over Ω_0. The following theorem says that q_h^{asy} is a "good" approximation of q_h. Before we state the theorem, note that q_h^{asy} is accessible as $h \to 0$. For the complete proof of this theorem, see [1].

Theorem 3.4. *If the assumptions in Theorem 3.2 and the Assumption T1 are satisfied, then*

$$\|\nabla(q_h^{asy} - q_h)\|_{L^2(\Omega_2)} \leq ch^{\sigma}\|\nabla(u - u_h)\|_{L^2(\Omega_2)},$$

and

$$|\|\nabla(u - u_h)\|_{L^2(\Omega_2)} - \|\nabla(q - q_h^{asy})\|_{L^2(\Omega_2)}| \leq ch^{\sigma}\|\nabla(u - u_h)\|_{L^2(\Omega_2)}. \tag{3.4}$$

Again, having stable error indicator enables us to conclude that

$$\lim_{h\to 0} \frac{\varepsilon(u_h, f, \Omega_2)}{\|\nabla(u - u_h)\|_{L^2(\Omega_2)}} = \lim_{h\to 0} \frac{\varepsilon(q_h^{asy}, -Lq, \Omega_2)}{\|\nabla(q - q_h^{asy})\|_{L^2(\Omega_2)}}. \tag{3.5}$$

Since the error $q - q_h^{asy}$ is periodic over Ω_2, the local error indicators approximate it in the same way for every cell. Hence, we have

$$\lim_{h\to 0} \frac{\varepsilon(u_h, f, \Omega_2)}{\|\nabla(u - u_h)\|_{L^2(\Omega_2)}} = \lim_{h\to 0} \frac{\varepsilon(q_h^{asy}, -Lq, \mathcal{C}(c_\alpha, h))}{\|\nabla(q - q_h^{asy})\|_{L^2(\mathcal{C}(c_\alpha,h))}} \tag{3.6}$$

for any $\alpha \in \mathcal{I}_2$. Hence, we need only to consider the error indicator on one cell in Ω_2. Furthermore, recovery type error estimators, as well as residual type error estimators, are scale invariant, i.e., transforming the problem through the mapping $(x, y) \mapsto (cx, cy)$ leaves the estimator invariant. This is basically because these error estimators use the energy norm, or the L^2-norm of the gradient, and both of them are scale invariant. Using this fact along with the the result in (3.6), it is enough to study the effectivity index of the error indicator over $\mathcal{C}$ considering it a cell in a reference domain $\hat{\Omega}$ which is the union of non overlapping translations of $\mathcal{C}$. For our purposes, it is enough to take $\hat{\Omega}$ as a 3 by 3 cell matrix with the $\mathcal{C}$ itself being in the center of $\hat{\Omega}$ as shown in Fig. 4(a).

Let $\hat{q}_1 = \hat{x}^2, \hat{q}_2 = \hat{x}\hat{y}$, and $\hat{q}_3 = \hat{y}^2$, where $(\hat{x}, \hat{y}) \in \hat{\Omega}$, and the coordinate system of $\hat{\Omega}$ has its origin at the center of $\mathcal{C}$. Let $\hat{q}_{j,I}$ be the lagrange interpolation of $\hat{q}_j$ on $\hat{\Omega}$ for $j = 1, 2, 3$, and set

$\hat{q}_j^* = \hat{q}_{j,I} + \Pi_{\hat{\Omega}}^{per}(\hat{q}_j - \hat{q}_{j,I})$. Without loss of generality, we may consider quadratic polynomials of the form $\hat{q} = \sum_{j=1}^{3} c_j \hat{q}_j$ for some c_1, c_2, and c_3 in $\mathbb{R}$. Then, by linearity of $\Pi_{\hat{\Omega}}^{per}$, $\hat{q}^* = \sum_{j=1}^{3} c_j \hat{q}_j^*$. Equation (3.6), along with the fact that recovery error indicators are scale invariant, leads to

$$\lim_{h\to 0} \frac{\varepsilon(u_h, f, \Omega_2)}{\|\nabla(u - u_h)\|_{L^2(\Omega_2)}} = \frac{\varepsilon(\hat{q}^*, -L\hat{q}, \mathcal{C})}{\|\nabla(\hat{q} - \hat{q}^*)\|_{L^2(\mathcal{C})}}. \tag{3.7}$$

Let $\boldsymbol{c} = \begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix}^T$ and let $\mathrm{G}\hat{q}_j^*$ denotes the recovered gradient for $\hat{q}_j^*$ on $\mathcal{C}$ for $j = 1, 2, 3$. It is easy to verify that $\varepsilon(\hat{q}^*, -L\hat{q}, \mathcal{C})^2 = \boldsymbol{c}^T \boldsymbol{M}_e \boldsymbol{c}$ where $\boldsymbol{M}_e \in \mathbb{R}^{3\times 3}$ with

$$\boldsymbol{M}_e(i,j) = \int_{\mathcal{C}} (\mathrm{G}\hat{q}_i^* - \nabla\hat{q}_i^*)(\mathrm{G}\hat{q}_j^* - \nabla\hat{q}_j^*) \quad \text{for } i, j = 1, 2, 3.$$

Also, it can be shown that $\|\nabla(\hat{q} - \hat{q}^*)\|_{L^2(\mathcal{C})}^2 = \boldsymbol{c}^T \boldsymbol{M}_a \boldsymbol{c}$ where $\boldsymbol{M}_a \in \mathbb{R}^{3\times 3}$ with

$$\boldsymbol{M}_a(i,j) = \int_{\mathcal{C}} \nabla(\hat{q}_i - \hat{q}_i^*)\nabla(\hat{q}_j - \hat{q}_j^*) \qquad \text{for } i, j = 1, 2, 3.$$

Using these results in (3.7), we get

$$\lim_{h\to 0} (\kappa(u_h, f, \Omega_2))^2 = \frac{\boldsymbol{c}^T M_e \boldsymbol{c}}{\boldsymbol{c}^T M_a \boldsymbol{c}}. \tag{3.8}$$

Note that $\boldsymbol{M}_e$ and $\boldsymbol{M}_a$ are symmetric positive definite matrices. To see the beauty of this result, let $\mathcal{Q} = \{\sum_{i=1}^{3} c_i \hat{q}_i : c_1, c_2, c_3 \in \mathbb{R}\}$ and suppose that $\mathcal{M}$ contains just one mesh. Then,

$$\sqrt{\lambda_{min}} \le \lim_{h\to 0} \kappa(u_h, f, \Omega_2) \le \sqrt{\lambda_{max}}$$

where λ_{min} and λ_{max} are, respectively, the minimum and the maximum eigenvalues of the generalized eigenvalue problem

$$\boldsymbol{M}_e \boldsymbol{c} = \lambda \boldsymbol{M}_a \boldsymbol{c}.$$

Let us now have an example that explains this theory.

Example 1. Consider the reference cell $\mathcal{C} = [-1, 1] \times [-a, a]$, with the aspect ratio a taken to be the height to width ratio. The mesh on $\mathcal{C}$ is obtained by partitioning $\mathcal{C}$ into triangles arranged in Chevron pattern, as shown in Fig. 4(b). We want to study the effect of a on the effectivity index of the PPR error indicator when $\mathcal{D}$ in (2.1) is the identity matrix. This requires computing $\hat{q_j}^*$ for $j = 1, 2, 3$ as was explained at the end of last section. It is enough to illustrate the computations procedure for $\hat{q_1}^*$.

To simplify notations, let $\hat{u} = \hat{x}^2$ and let $\hat{u}_I$ be its Lagrange interpolation over $\hat{\Omega}$. Then, by definition,

$$\begin{aligned} \hat{u}^* &= \hat{u}_I + \Pi_{\hat{\Omega}}^{per}(\hat{u} - \hat{u}_I) \\ &= \hat{u}_I + E_{\hat{\Omega}}^{per} \circ \Pi_{\mathcal{C}}^{per} \circ (E_{\hat{\Omega}}^{per})^{-1}(\hat{u} - \hat{u}_I). \end{aligned} \tag{3.9}$$

Note that $(E^{per}_{\hat{\Omega}})^{-1}(\hat{u}-\hat{u}_I) = (\hat{u}-\hat{u}_I)|_{\mathcal{C}}$. Let $\hat{e} = (\hat{u}-\hat{u}_I)|_{\mathcal{C}}$ and $\hat{w} = \Pi^{per}_{\hat{\Omega}}(\hat{e})$, then

$$\hat{u}^* = \hat{u}_I + E^{per}_{\hat{\Omega}}\hat{w}, \tag{3.10}$$

and our task of finding $\hat{u}^*$ is reduced to find $\hat{w}$ which satisfies the equations

$$\begin{cases} B_{\mathcal{C}}(\hat{w},\hat{v}) &= B_{\mathcal{C}}(\hat{e},\hat{v}) \qquad \forall \hat{v} \in S^{per}(\mathcal{C}) \\ \int_{\mathcal{C}} \hat{w} &= \int_{\mathcal{C}} \hat{e} \end{cases}. \tag{3.11}$$

Let $\hat{u}_i = \hat{u}(\hat{z}_i)$ and $\hat{w}_i = \hat{w}(\hat{z}_i)$ for $i = 1,\ldots,9$. Also, let $\hat{\varphi}_i$ be the standard Lagrange basis function associated with $\hat{z}_i$, i.e., $\hat{\varphi}_i$ is piecewise continuous linear function satisfying $\hat{\varphi}_i(\hat{z}_j) = \delta_{ij}$ for $i,j = 1,\ldots,9$. Then,

$$\hat{u}_I = \sum_{i=1}^{9} \hat{u}_i\hat{\varphi}_i = \hat{\boldsymbol{\varphi}}^T\hat{\boldsymbol{u}}_L, \text{ and } \hat{w} = \sum_{i=1}^{9} \hat{w}_i\hat{\varphi}_i = \hat{\boldsymbol{\varphi}}^T\hat{\boldsymbol{w}}_L \tag{3.12}$$

where

$$\hat{\boldsymbol{u}}_L = \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_9 \end{bmatrix}, \hat{\boldsymbol{w}}_L = \begin{bmatrix} \hat{w}_1 \\ \hat{w}_2 \\ \vdots \\ \hat{w}_9 \end{bmatrix}, \text{ and } \hat{\boldsymbol{\varphi}} = \begin{bmatrix} \hat{\varphi}_1 \\ \hat{\varphi}_2 \\ \vdots \\ \hat{\varphi}_9 \end{bmatrix}.$$

Preforming the standard finite element sub-assembling procedure on the first equation in (3.11), we get

$$K_L\hat{\boldsymbol{w}}_L = \hat{\boldsymbol{e}}_L \tag{3.13}$$

where

$$K_L = \frac{1}{2a}\begin{bmatrix} a^2{+}1 & -a^2 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -a^2 & 2(a^2{+}1) & -a^2 & 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & -a^2 & a^2{+}1 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 2(a^2{+}1) & -2a^2 & 0 & -1 & 0 & 0 \\ 0 & -2 & 0 & -2a^2 & 4(a^2{+}1) & -2a^2 & 0 & -2 & 0 \\ 0 & 0 & -1 & 0 & -2a^2 & 2(a^2{+}1) & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & a^2{+}1 & -a^2 & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 & -a^2 & 2(a^2{+}1) & -a^2 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -a^2 & a^2{+}1 \end{bmatrix} \tag{3.14}$$

and

$$\hat{\boldsymbol{e}}_L = \frac{a}{6}\begin{bmatrix} -1 & 2 & -1 & 0 & 0 & 0 & 1 & -2 & 1 \end{bmatrix}^T.$$

Since $\hat{w} \in S^{per}(\mathcal{C})$, then, by the periodicity conditions,

$$\begin{cases} \hat{w}_3 = \hat{w}_7 = \hat{w}_9 = \hat{w}_1 \\ \hat{w}_8 = \hat{w}_2 \\ \hat{w}_6 = \hat{w}_4 \end{cases}. \tag{3.15}$$

Let $\hat{\boldsymbol{w}}_P = \begin{bmatrix} \hat{w}_1 & \hat{w}_2 & \hat{w}_4 & \hat{w}_5 \end{bmatrix}^T$, then the relations in (3.15) implies that

$$\hat{\boldsymbol{w}}_L = R^T \hat{\boldsymbol{w}}_P \tag{3.16}$$

where

$$R = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Premultiplying the linear system in (3.13) by R and using (3.16), we get

$$K_P \hat{\boldsymbol{w}}_P = \hat{\boldsymbol{e}}_P \tag{3.17}$$

where

$$K_P = R K_L R^T = \frac{2}{a} \begin{bmatrix} a^2+1 & -a^2 & -1 & 0 \\ -a^2 & a^2+1 & 0 & -1 \\ -1 & 0 & a^2+1 & -a^2 \\ 0 & -1 & -a^2 & a^2+1 \end{bmatrix} \text{ and } \hat{\boldsymbol{e}}_P = R\hat{\boldsymbol{e}}_L = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

It is easy to see that K_P is singular, and the second condition in (3.11) comes into action. Using the representations in (3.12), $\int_{\mathcal{C}} \hat{w} = \int_{\mathcal{C}} \hat{e}$ leads to

$$\left(\int_{\mathcal{C}} \boldsymbol{\varphi}^T \right) \hat{\boldsymbol{w}}_L = \int_{\mathcal{C}} \hat{u} - \left(\int_{\mathcal{C}} \boldsymbol{\varphi}^T \right) \hat{\boldsymbol{u}}_L$$

or

$$\frac{a}{6} \begin{bmatrix} 2 & 2 & 2 & 3 & 6 & 3 & 1 & 4 & 1 \end{bmatrix} \hat{\boldsymbol{w}}_L = \frac{-2a}{3}. \tag{3.18}$$

Using (3.16) in (3.18), we get

$$a \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \hat{\boldsymbol{w}}_P = \frac{-2a}{3} \tag{3.19}$$

Equations (3.17) and (3.19) uniquely defines $\hat{\boldsymbol{w}}_P$, where

$$\hat{\boldsymbol{w}}_P = \frac{-1}{6} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T.$$

Hence,

$$\Pi_{\mathcal{C}}^{per}(\hat{q}_1 - \hat{q}_{1,I}) = \hat{w} = \frac{-1}{6} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \hat{\boldsymbol{\varphi}}.$$

Similarly,

$$\Pi_{\mathcal{C}}^{per}(\hat{q}_2 - \hat{q}_{2,I}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \hat{\boldsymbol{\varphi}}$$

and

$$\Pi_{\mathcal{C}}^{per}(\hat{q}_3 - \hat{q}_{3,I}) = \frac{-a^2}{6} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \hat{\boldsymbol{\varphi}}.$$

The next step is to recover the gradient at the mesh nodes of $\mathcal{C}$ using the PPR. Having the PPR-recovered gradient, M_a and M_e were found to be equal with

$$\boldsymbol{M}_a = \boldsymbol{M}_e = \frac{2a}{3}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1+a^2 & 0 \\ 0 & 0 & 2a^2 \end{bmatrix}.$$

Hence, the PPR error indicator is asymptotically exact in case of Chevron pattern regardless of the reference cell aspect ratio.

4 Asymptotic Behavior of the PPR Error Indicators over Internal Patches

In this section we evaluate the performance of the PPR error indicator over internal patches through some tests similar to Example 1. At the same time, the performance of the PPR error indicator is compared with that of the SPR error indicator. Beginning with the reference cell $\mathcal{C} = [-1,1] \times [-1,1]$ partitioned into triangles arranged into one of the well-known patterns shown in Fig. 5, we want to study how the effectivity index (or, equivalently, the robustness index) would change in the following cases:

1. The mesh on $\mathcal{C}$ is distorted from the given pattern by moving the central node along one of the four lines: $\hat{y} = 0$, $\hat{x} = 0$, $\hat{y} = \hat{x}$, or $\hat{y} = -\hat{x}$. An example for the Criss Cross pattern is depicted in Fig. 6. In this case $\mathcal{D}$ is taken to be the identity.

2. The height of the reference cell is changing while the width is fixed at 2. The height is taken to be $2a$. In this case $\mathcal{D}$ is taken to be the identity.

3. The material properties, represented in the model problem by $\mathcal{D}$, are changed by varying $\theta \in [-\pi/2, \pi/2]$ and $d \geq 1$.

In these tests $\mathcal{Q} = \{\sum_{i=1}^{3} c_i \hat{q}_i : c_1, c_2, c_3 \in \mathbb{R}\}$. When $\mathcal{D}$ is the identity and $f = 0$ in (2.1), the solution u is harmonic. In this case $\mathcal{Q}$ contains only harmonic quadratic polynomials. Note that if $\hat{q} = \sum_{i=1}^{3} c_i \hat{q}_i$ is harmonic, then $c_3 = -c_1$ and $\hat{q} = c_1(\hat{q}_1 - \hat{q}_3) + c_2\hat{q}_2$. For the class of harmonic polynomials, the matrices $\boldsymbol{M}_e$ and $\boldsymbol{M}_a$ in (3.8) are replaced by $\boldsymbol{H}^T\boldsymbol{M}_e\boldsymbol{H}$ and $\boldsymbol{H}^T\boldsymbol{M}_a\boldsymbol{H}$, respectively, where

$$\boldsymbol{H} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Since $\mathcal{Q}$ contains all general polynomials, or all harmonic polynomials, studying the effectivity index (or the robustness index) is reduced to the study of the minimum and maximum eigenvalues, λ_{min} and λ_{max}, of the generalized eigenvalue problem $\boldsymbol{M}_e\boldsymbol{c} = \lambda \boldsymbol{M}_a\boldsymbol{c}$.

The response of robustness index to the distortion of the mesh on $\mathcal{C}$ is reported in Fig. 8 through Fig. 15. To better understand how to carry out these distortion tests, let us consider one of them. We distort Chevron pattern by moving the middle node along $\hat{y} = \hat{x}$ to (δ, δ) where δ has admissible values in $(-1, 0.5)$. By admissible values, we mean those values that would not destroy the triangulation of the mesh on $\mathcal{C}$. Varying the polynomial $\hat{q}$ over $\mathcal{Q}$, one gets an expression of λ_{min} and λ_{max} as a function of δ. This can be done using a symbolic computing package. In a similar way, the rest of the distortion tests are carried out with different values for δ depending on the basic pattern. From these tests, we observe the following:

- The PPR error indicator is exact for zero distortion in all of the four patterns while the SPR error indicator is not in case of the Chevron pattern.
- The PPR error indicator is less sensitive to distortion than SPR error indicator. The robustness index for the PPR error indicator is almost zero when the distortion is small.
- The robustness index of both error indicators is bounded when some of the mesh triangles in $\mathcal{C}$ degenerate as δ reaches its limits.
- Within practical distortion limits, the PPR error indicator is more robust than the SPR error indicator.
- The robustness index of the SPR error indicator gets smaller when the tests target the class of harmonic polynomials. The same is true for the PPR error indicator although the change is not significant in case of the Regular pattern.

The response of the effectivity index to changes in aspect ratio and material properties is reported in Table 1.

As it is clear from the table, the PPR error indicator is asymptotically exact in all of the four patterns. This is true for any cell aspect ratio and for any material properties. The SPR error indicator is a little bit sensitive in case of the Chevron pattern, but it is asymptotically exact in the other three patterns.

Pattern	Changing Factor					
	Cell Aspect Ratio				Material Properties	
	General Polynomials		Harmonic Polynomials		PPR	SPR
	PPR	SPR	PPR	SPR		
Regular	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$
Chevron	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=\frac{37a^2+49}{49(a^2+1)}$ $\lambda_{max}=\frac{50}{49}$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=\frac{37a^2+49}{49(a^2+1)}$ $\lambda_{max}=\frac{50a^2+49}{49(a^2+1)}$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=\frac{43}{49}$ $\lambda_{max}=\frac{50}{49}$
Union Jack	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$
Criss Cross	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$	$\lambda_{min}=1$ $\lambda_{max}=1$

Table 1. Response of λ_{min} and λ_{max} to changes in cell aspect ratio and material properties when the patch is inside the domain

5 Extending the Computer-Based Approach to Patches Adjacent to the Boundary

So far we have seen how to study the asymptotic quality of error indicators over internal patches. Naturally, one may wonder if this methodology could be used for patches adjacent to the boundary. Recall that the core of this methodology is to use asymptotic finite element approximation. If this approximation satisfies the boundary conditions, the methodology is still applicable. Unfortunately, the procedure described in Section 3 does not take the boundary conditions into account. Therefore, we should not expect the asymptotic finite element approximation to satisfy the boundary conditions. The target of this section is to show how to build this approximation so that it satisfies the boundary conditions. If we can do that, the asymptotic quality of local error indicators over patches adjacent to the boundary can be studied in the same way as before.

Again, let q be the quadratic part of the Taylor expansion of u at some point on $\partial\Omega$ and let $q_h^{asy,1}$ be the asymptotic finite element approximation constructed as explained in Section 3. If $q_h^{asy,1}$ does not satisfy the boundary conditions for q, then we need to compute another component q_h^{bl} such that $q_h^{asy,2} = q_h^{asy,1} + q_h^{bl}$ satisfies the boundary conditions for q. The component q_h^{bl} is called *the boundary layer*, and $q_h^{asy,2}$ will be called *total asymptotic finite element approximation*. Without loss of generality, we may assume that the segment of $\partial\Omega$

under study is a horizontal edge along $x-$axis with Ω being in the upper half plane. We should observe the following.

1. Extending $q_h^{asy,1}$ up to the boundary requires the mesh adjacent to $\partial\Omega$ to satisfy Assumption T1.

2. Since $q_h^{asy,1}$ approximates the finite element solution of $-Lu = f$, q_h^{bl} should approximate the finite element solution of $-Lu = 0$. Moreover, q_h^{bl} should decay as we go inside the domain so that its effect on local error indicators for internal patches is negligible, as explained in Remark 3.1. Hence, q_h^{bl} must satisfy the *decay condition*

$$\lim_{y\to\infty} \nabla q_h^{bl} = 0. \tag{5.1}$$

3. Since $q - q_h^{asy,1}$ is periodic, q_h^{bl} is periodic in the horizontal direction, but not in the vertical direction because of the decay condition in (5.1).

As it was explained in Section 3, it suffices to do computations over a reference domain $\hat{\Omega}$. Since we are doing asymptotic computations, $\hat{\Omega}$ may be taken as the union of five infinite vertical strips, which is enough to do the gradient recovery computations as depicted in Fig. 7(a). Each of these strips is a horizontal translation of the strip $\mathcal{S} = [-1, 1] \times [0, \infty]$ with $\mathcal{S}$ itself being the middle strip in $\hat{\Omega}$. The strip $\mathcal{S}$ is partitioned using vertical translates of a reference cell $\mathcal{C}$. The cells are numbered as $1, 2, \ldots$, starting at the bottom and the horizontal edge of the first cell along $\hat{x}-$axis is denoted by $\hat{\Gamma}$.

The periodicity of the boundary layer component in the horizontal direction motivates the definition of the finite element subspace

$$S^{bl}(\mathcal{S}) = \{\hat{v} : \hat{v}|_\tau \in P_1(\tau)\, \forall \tau \in \mathcal{T}_\mathcal{S}\}$$

where $\mathcal{T}_\mathcal{S}$ is the triangular mesh used on $\mathcal{S}$. Note that every function $\hat{v} \in S^{bl}(\mathcal{S})$ can be extended to all of $\hat{\Omega}$ by translating it horizontally to every vertical strip in $\hat{\Omega}$, and, then, piecing these translates together. The extension of $\hat{v} \in S^{bl}(\mathcal{S})$ is well-defined as the translates match on vertical edges of the strips. The space of these extended functions is denoted by $S^{bl}(\hat{\Omega})$ and this extension procedure is represented by the operator $E^{bl}_{\hat{\Omega}} : S^{bl}(\mathcal{S}) \to S^{bl}(\hat{\Omega})$ and

$$S^{bl}(\hat{\Omega}) = \{E^{bl}_{\hat{\Omega}}\hat{v} : \hat{v} \in S^{bl}(\mathcal{S})\}.$$

As before, let $\hat{q}_1 = \hat{x}^2, \hat{q}_2 = \hat{x}\hat{y}$, and $\hat{q}_3 = \hat{y}^2$ where $(\hat{x}, \hat{y}) \in \hat{\Omega}$. Also, we may assume that any quadratic polynomial $\hat{q}$ on $\hat{\Omega}$ takes the form $\sum_{j=1}^3 c_j \hat{q}_j$ for some c_1, c_2, and c_3 in $\mathbb{R}$. To

compute $\boldsymbol{M}_e$ and $\boldsymbol{M}_a$, we need to compute $\hat{q}_j^*$ for $j = 1, 2, 3$, but the definition of $\hat{q}_j^*$ has to change to include the boundary layer component. The new definition is

$$\hat{q}^* = \hat{q}_I + E^{per}_{\hat{\Omega}} \hat{q}^{per} + E^{bl}_{\hat{\Omega}} \hat{q}^{bl} \tag{5.2}$$

where components of $\hat{q}^*$ are as follows. The first component, $\hat{q}_I$, is the Lagrange interpolation of $\hat{q}$ over $\hat{\Omega}$. The second component, $E^{per}_{\hat{\Omega}} \hat{q}^{per} \in S^{per}(\hat{\Omega})$, is the periodic extension of $\hat{q}^{per} \in S^{per}(\mathcal{C})$ where

$$\hat{q}^{per} = \Pi^{per}_{\mathcal{C}}(\hat{q} - \hat{q}_I).$$

The third component, $E^{bl}_{\hat{\Omega}} \hat{q}^{bl} \in S^{bl}(\hat{\Omega})$, is the extension of $\hat{q}^{bl} \in S^{bl}(\mathcal{S})$, and it accounts for the boundary layer. Therefore, and according to properties of the boundary layer component, $\hat{q}^{bl}$ is constructed as follows.

1. In case of Dirichlet boundary conditions, the nodal values of $\hat{q}^*$ on $\hat{\Gamma}$ must be equal to those of $\hat{q}$. Therefore, $\hat{q}^{bl}$ solves the problem

$$\begin{cases} B_{\mathcal{S}}(\hat{q}^{bl}, \hat{v}) = 0 & \forall \hat{v} \in S^{bl,0}(\mathcal{S}) \\ \hat{q}^{bl}|_{\hat{\Gamma}} = -\hat{q}^{per}|_{\hat{\Gamma}} \end{cases} \tag{5.3}$$

   where $S^{bl,0}(\mathcal{S}) = \{\hat{v} \in S^{bl}(\mathcal{S}) : \hat{v}|_{\hat{\Gamma}} = 0\}$

2. In case of Neumann boundary conditions $\hat{q}^{bl}$ solves the problem

$$B_{\mathcal{S}}(\hat{q}^{bl}, \hat{v}) = \int_{\hat{\Gamma}} (\boldsymbol{n} \cdot \mathcal{D} \nabla \hat{g}) \hat{v} \qquad \forall \hat{v} \in S^{bl}(\mathcal{S}) \tag{5.4}$$

   where

$$\hat{g} = \hat{q} - \hat{q}_I - E^{per}_{\mathcal{S}} \hat{q}^{per}. \tag{5.5}$$

   For compatibility, we must have

$$\int_{\hat{\Gamma}} (\boldsymbol{n} \cdot \mathcal{D} \nabla \hat{g}) = 0 \tag{5.6}$$

   and for uniqueness of $\hat{q}^{bl}$ we use $\int_{\mathcal{S}} \hat{q}^{bl} = 0$.

3. Regardless of the boundary conditions type, $\hat{q}^{bl}$ must satisfy the decay condition

$$\lim_{\hat{y} \to \infty} \nabla \hat{q}^{bl} = 0. \tag{5.7}$$

Remark 5.1. Computing $\hat{q}^{bl}$ is tedious. For general situations, the steps are explained in [5], Appendix A.2. In the current work, studying the asymptotic quality of the PPR error indicators is restricted to the cases in which $\mathcal{C} = [-1, 1] \times [0, 2a]$. The triangulation in $\mathcal{C}$ is a distorted version of one of the basic patterns shown in Fig. 5 where the basic pattern is distorted by moving its central node. This will make the computations easier as we will see soon.

Theorem 5.2. *Let $\mathcal{C}$ and the triangulation defined on it be as in Remark 5.1, and let the boundary conditions along $\hat{\Gamma}$ be of Neumann type. Then, the compatibility condition in* (5.6) *is equivalent to $\int_{\hat{\Gamma}} \partial_{\hat{y}} \hat{g} = 0$ and $\hat{q}^{bl}$ is identically 0 on $\mathcal{S}$.*

Proof. Let $\hat{v} \in S^{bl}(\mathcal{S})$, and let $\hat{g}$, defined in (5.5), satisfy the compatibility condition in (5.6). The key to this proof is to note that $\hat{v}|_{\hat{\Gamma}}$, which implies that $(\partial_{\hat{x}} \hat{v})|_{\hat{\Gamma}}$ is odd. The same thing is true for $\hat{q} - \hat{q}_I, E_{\mathcal{S}}^{per} \hat{q}^{per}$, and, hence, $\hat{g}$.

Since $(\partial_{\hat{x}} \hat{g})|_{\hat{\Gamma}}$ is odd and $\mathcal{D}$ is positive definite, the compatibility condition in (5.6) is equivalent to

$$\int_{\hat{\Gamma}} \partial_{\hat{y}} \hat{g} = 0 \tag{5.8}$$

and the first claim in this lemma is true. To prove the second claim, it is enough to show that $\int_{\hat{\Gamma}} (\boldsymbol{n} \cdot \mathcal{D} \nabla \hat{g}) \hat{v} = 0$. Again, since $\partial_{\hat{x}} \hat{g}|_{\hat{\Gamma}}$ is odd, $\hat{v}|_{\hat{\Gamma}}$ is even, and $\mathcal{D}$ is positive definite, we get

$$\int_{\hat{\Gamma}} (\partial_{\hat{x}} \hat{g}) \hat{v} = 0.$$

Therefore, the proof of the first claim is complete if

$$\int_{\hat{\Gamma}} (\partial_{\hat{y}} \hat{g}) \hat{v} = 0, \tag{5.9}$$

and this is trivially the case if $(\partial_{\hat{y}} \hat{g})|_{\hat{\Gamma}}$ is odd.

Set $\hat{g}_j = \hat{g}$ when $\hat{q} = \hat{q}_j$ for $j = 1, 2, 3$. Using the definition of $\hat{q}$, $\hat{g} = \sum_{j=1}^{3} c_j \hat{g}_j$. Since $(\partial_{\hat{y}} \hat{q}_1)|_{\hat{\Gamma}} = 0$, $(\partial_{\hat{y}} \hat{g}_1)|_{\hat{\Gamma}} = \hat{h}_1(\hat{x})$ for some piecewise constant function $\hat{h}_1$. Similarly, $(\partial_{\hat{y}} \hat{g}_3)|_{\hat{\Gamma}} = \hat{h}_3(\hat{x})$ for some piecewise constant function $\hat{h}_3$. For $j = 2, (\partial_{\hat{y}} \hat{q}_2)|_{\hat{\Gamma}} = \hat{x}$, and, hence, $(\partial_{\hat{y}} \hat{g}_2)|_{\hat{\Gamma}} = \hat{x} + \hat{h}_2(\hat{x})$ for some piecewise constant function $\hat{h}_2$. Therefore, $(\partial_{\hat{y}} \hat{g})|_{\hat{\Gamma}} = c_2 \hat{x} + \hat{h}(\hat{x})$ where $\hat{h} = \sum_{j=1}^{3} c_j \hat{h}_j$, and $(\partial_{\hat{y}} \hat{g})|_{\hat{\Gamma}}$ is odd if and only if $\hat{h}$ is odd. The compatibility condition in (5.8) leads to $\int_{\hat{\Gamma}} \hat{h} = 0$, and $\hat{h}$ must be odd as it is piecewise constant. □

Theorem 5.2 says that if boundary conditions on $\hat{\Gamma}$ are of Neumann type, $\hat{q}^{bl}$ is identically 0 on $\mathcal{S}$ and it suffices to find the combinations of c_1, c_2, and c_3, used in the definition of $\hat{q}$, that makes $\int_{\hat{\Gamma}} \partial_{\hat{y}} \hat{g} = 0$. The situation in case of Dirichlet boundary conditions is different and $\hat{q}^{bl}$. The following example illustrates the computations procedure.

Example 2. In this example we construct the boundary layer $\hat{q}_1^{bl}$ corresponding to $\hat{q}_1$ on $\mathcal{S}$ when boundary conditions on $\hat{\Gamma}$ are of Dirichlet type. The reference cell $\mathcal{C} = [-1, 1] \times [0, 2a]$, the triangulation pattern on $\mathcal{C}$ is Chevron, and mesh nodes in the mth cell are numbered as shown in Fig. 7(b). To simplify notations, we set $\hat{w} = \hat{q}_1^{bl}$. Hence, $\hat{w}$ is the solution of

$$\begin{cases} B_{\mathcal{S}}(\hat{w}, \hat{v}) = 0 \quad \forall \hat{v} \in S^{bl,0}(\mathcal{S}) \\ \hat{w}|_{\hat{\Gamma}} = -\hat{q}_1^{per}|_{\hat{\Gamma}} \\ \lim\limits_{\hat{y} \to \infty} \nabla \hat{w} = 0 \end{cases} . \tag{5.10}$$

Set $\hat{w}_{(j,i)} = \hat{w}(\hat{z}_{(j,i)})$ for $j \geq 0$, and $i = 1, 2, 3$. For $j \geq 0$, let

$$\hat{\boldsymbol{w}}_P^{(j)} = \begin{bmatrix} \hat{w}_{(j,1)} & \hat{w}_{(j,2)} \end{bmatrix}^T \text{ and } \hat{\boldsymbol{w}}_L^{(j)} = \begin{bmatrix} \hat{w}_{(j,1)} & \hat{w}_{(j,2)} & \hat{w}_{(j,3)} \end{bmatrix}^T.$$

Using the results of Example 1, the boundary conditions imposed on $\hat{w}$ in (5.10) implies that

$$\hat{\boldsymbol{w}}_L^{(0)} = \frac{1}{6} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T. \tag{5.11}$$

Performing standard finite element sub-assembling procedure on the mth cell using standard Lagrange basis functions, the stiffness matrix and the load vectors are K_L, and $\mathbf{0}$, respectively, where K_L has the form given in (3.14). Since $\hat{w} \in S^{bl}(\mathcal{S}), \hat{w}_{(j,3)} = \hat{w}_{(j,1)}$ for all $j \geq 0$. Therefore,

$$\begin{bmatrix} \hat{\boldsymbol{w}}_L^{(2m-2)} \\ \hat{\boldsymbol{w}}_L^{(2m-1)} \\ \hat{\boldsymbol{w}}_L^{(2m)} \end{bmatrix} = R^T \begin{bmatrix} \hat{\boldsymbol{w}}_P^{(2m-2)} \\ \hat{\boldsymbol{w}}_P^{(2m-1)} \\ \hat{\boldsymbol{w}}_P^{(2m)} \end{bmatrix} \tag{5.12}$$

where

$$R = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Using (5.12), the stiffness matrix K_L is reduced to K_P, where

$$K_P = RK_LR^T = \frac{1}{a} \begin{bmatrix} a^2+1 & -a^2 & -1 & 0 & 0 & 0 \\ -a^2 & a^2+1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 2(a^2+1) & -2a^2 & -1 & 0 \\ 0 & -1 & -2a^2 & 2(a^2+1) & 0 & -1 \\ 0 & 0 & -1 & 0 & a^2+1 & -a^2 \\ 0 & 0 & 0 & -1 & -a^2 & a^2+1 \end{bmatrix}.$$

Note that the degrees of freedom $\hat{\boldsymbol{w}}_P^{(2m-1)}$ appears only in the mth cell, and the load vector is zero. Hence, using the third and fourth rows of K_P, it can be shown that

$$\hat{\boldsymbol{w}}_P^{(2m-1)} = \frac{1}{2(2a^2+1)} \begin{bmatrix} a^2+1 & a^2 \\ a^2 & a^2+1 \end{bmatrix} \left(\hat{\boldsymbol{w}}_P^{(2m-2)} + \hat{\boldsymbol{w}}_P^{(2m)} \right). \tag{5.13}$$

Having (5.13), the degrees of freedom $\hat{\boldsymbol{w}}_P^{(2m-1)}$ are eliminated and the stiffness matrix K_P is reduced to $\tilde{K}_P$ where

$$\tilde{K}_P = \begin{bmatrix} K_1 & K_2 \\ K_2 & K_1 \end{bmatrix} = \frac{1}{2a(2a^2+1)} \begin{bmatrix} 4a^4+5a^2+1 & -a^2(4a^2+3) & -(a^2+1) & -a^2 \\ -a^2(4a^2+3) & 4a^4+5a^2+1 & -a^2 & (-a^2+1) \\ -(a^2+1) & -a^2 & 4a^4+5a^2+1 & -a^2(4a^2+3) \\ -a^2 & -(a^2+1) & -a^2(4a^2+3) & 4a^4+5a^2+1 \end{bmatrix}. \tag{5.14}$$

The block matrices K_1 and K_2 are 2×2.

Assembling the reduced stiffness matrices $\tilde{K}_P$ from mth and $(m+1)$th cells, we get

$$\begin{cases} K_2 \hat{\boldsymbol{w}}_P^{(2m-2)} + 2K_1 \hat{\boldsymbol{w}}_P^{(2m)} + K_2 \hat{\boldsymbol{w}}_P^{(2m+2)} = \mathbf{0} \text{ for } m \geq 1 \\ \hat{\boldsymbol{w}}_P^{(0)} = \frac{1}{6} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{cases} \tag{5.15}$$

The solution of the recurrence relation in (5.15) when $m \geq 2$ can be written in the form

$$\hat{\boldsymbol{w}}_P^{(2j)} = \mu^j \boldsymbol{b} \text{ for } j \geq 1 \tag{5.16}$$

where $\mu \in \mathbb{R}$ and $\boldsymbol{b} \in \mathbb{R}^2$. To determine μ and $\boldsymbol{b}$, use (5.16) in the recurrence equation in (5.15). The resulting equation can be simplified to the form

$$2K_1 \boldsymbol{b} = -K_2(\mu^{-1} + \mu)\boldsymbol{b}. \tag{5.17}$$

Setting $A = K_2^{-1} K_1$, (5.17) takes the form

$$A\boldsymbol{b} = -\frac{(\mu^{-1} + \mu)}{2}\boldsymbol{b}. \tag{5.18}$$

Note that $\boldsymbol{b}$ is an eigenvector of A and its corresponding eigenvalue is $-(\mu^{-1}+\mu)/2$. The eigenvalues of A are -2 and $2+16a^2+16a^4$, and their corresponding eigenvectors are $\boldsymbol{b}_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ and $\boldsymbol{b}_2 = \begin{bmatrix} -1 & 1 \end{bmatrix}^T$, respectively. The decay condition imposed on $\hat{w}$ limits accepted values of μ to $(-1, 1]$. With this in mind, μ corresponding to -2 is $\mu_1 = 1$, and μ corresponding to $2 + 16a^2 + 16a^4$ is $\mu_2 = 1 + 8a^4 + 8a^2 - 4a(1 + 2a^2)\sqrt{1 + a^2}$. Hence,

$$\hat{\boldsymbol{w}}_P^{(2j)} = \gamma_1 \mu_1^j \boldsymbol{b}_1 + \gamma_2 \mu_2^j \boldsymbol{b}_2 \text{ for } j \geq 1$$

for some constants γ_1 and γ_2 in $\mathbb{R}$. To determine γ_1 and γ_2, we use the recurrence relation in (5.15) for $m = 1$. This leads to

$$K_2 \hat{\boldsymbol{w}}_P^{(0)} = -[2K_1 \hat{\boldsymbol{w}}_P^{(2)} + K_2 \hat{\boldsymbol{w}}_P^{(4)}] = -\sum_{i=1}^{2} \gamma_i [2\mu_i K_1 + \mu_i^2 K_2] \boldsymbol{b}_i.$$

Using (5.17), and by invertibility of K_2, last equation can be simplified to

$$\hat{\boldsymbol{w}}_P^{(0)} = \sum_{i=1}^{2} \gamma_i \boldsymbol{b}_i = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}.$$

Solving this linear system, we get $\gamma_1 = \frac{1}{6}, \gamma_2 = 0$, and $\hat{w} = \frac{1}{6}$ everywhere on $\mathcal{S}$.

Remark 5.3. The result in the previous example is expected. In general, let $\mathcal{C} = [-1,1]\times[0,2a]$ and let the partition of $\mathcal{C}$ be admissible as defined at the end of Assumption T1. If $\hat{q}^{per}|_{\hat{\Gamma}} = \gamma$, where $\gamma \in \mathbb{R}$ is some constant, then, by uniqueness of $\hat{q}^{bl}$, $\hat{q}^{bl}$ is identically $-\gamma$ on $\mathcal{S}$. This is the typical situation when the number of the mesh nodes on $\hat{\Gamma}$ is 2, as in the Criss Cross pattern.

Although Example 2 has an expected result, many intermediate results can be generalized to cases in which $\mathcal{C}$ and its partition satisfy the assumptions in Remark 5.1. In these cases, using the information available about the triangulation and about $\hat{q}^{bl}$ can greatly simplify our computations as in Theorem 5.2. In what follows we will see how to do that. Using Remark 5.3, cases in which the partition on $\mathcal{C}$ is the Criss Cross pattern or a distorted version of this pattern are trivial, and we only need to consider the other cases. As we did in Example 2, we set $\hat{w} = \hat{q}^{bl}$ and let the nodes in the mth cell be numbered as in Fig. 7(b). Then, $\hat{w}$ is the solution of

$$\begin{cases} B_{\mathcal{S}}(\hat{w},\hat{v}) = 0 \quad \forall \hat{v} \in S^{bl,0}(\mathcal{S}) \\ \hat{w}|_{\hat{\Gamma}} = -\hat{q}^{per}|_{\hat{\Gamma}} \\ \lim\limits_{\hat{y}\to\infty} \nabla\hat{w} = 0 \end{cases} . \tag{5.19}$$

Using the notations of Example 2 and following the steps carried out on the mth cell, we can compute K_L, then reduce it to K_P which, in turn, is reduced to $\tilde{K}_P$. Note that all of these matrices are symmetric. Hence, $\tilde{K}_P$ can be written in the form

$$\tilde{K}_P = \begin{bmatrix} K_1 & K_2^T \\ K_2 & K_3 \end{bmatrix} \tag{5.20}$$

where K_1, K_2, and K_3 are 2×2 matrices. Assembling the reduced stiffness matrices $\tilde{K}_P$ from mth and $(m+1)$th cells, we get

$$\begin{cases} K_2\hat{\boldsymbol{w}}_P^{(2m-2)} + (K_1+K_3)\hat{\boldsymbol{w}}_P^{(2m)} + K_2^T\hat{\boldsymbol{w}}_P^{(2m+2)} = \mathbf{0} \text{ for } m \geq 1 \\ \hat{\boldsymbol{w}}_P^{(0)} = -\begin{bmatrix} \hat{q}^{per}(\hat{z}_{(0,1)}) \\ \hat{q}^{per}(\hat{z}_{(0,2)}) \end{bmatrix} \end{cases} \tag{5.21}$$

Note that the restriction of $\hat{w}$ to any of the horizontal edges of any of the cells in $\mathcal{S}$ is an even function of $\hat{x}$, and the space of piecewise linear even functions in $\hat{x}$ on $[-1,1]$ is the span of $\{1, 1-2|\hat{x}|\}$. Hence, $\hat{w}$ can be expressed in terms of two components. The restriction of the first component to any horizontal edge in any cell is a multiple of 1 while the restriction of the second component to any of such edges is a multiple of $1-2|\hat{x}|$. This implies that

$$\hat{\boldsymbol{w}}_P^{(2j)} = \mu_{(1,j)}\boldsymbol{b}_1 + \mu_{(2,j)}\boldsymbol{b}_2 \text{ for } j \geq 0 \tag{5.22}$$

where $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ are as in Example 2. The values $\mu_{(1,0)}$ and $\mu_{(2,0)}$ are determined from $\hat{\boldsymbol{w}}_P^{(0)}$, and without loss of generality, we assume $\mu_{(2,0)} \neq 0$; otherwise $\hat{w}|_{\hat{\Gamma}}$ is constant and this case is trivial by Remark 5.3.

By the independence of the two components of $\hat{w}$, the components of $\hat{\boldsymbol{w}}_P^{(2j)}$ satisfy the recurrence equation in (5.21) for $m \geq 2$. Consequently,

$$\mu_{(i,j-1)}K_2\boldsymbol{b}_i + \mu_{(i,j)}(K_1+K_3)\boldsymbol{b}_i + \mu_{(i,j+1)}K_2^T\boldsymbol{b}_i = \mathbf{0} \text{ for } j \geq 2, \text{ and } i = 1,2. \tag{5.23}$$

Premultiplying (5.23) by $\boldsymbol{b}_i^T$, we get

$$(\mu_{(i,j-1)} + \mu_{(i,j+1)})\boldsymbol{b}_i^T K_2\boldsymbol{b}_i + \mu_{(i,j)}\boldsymbol{b}_i^T(K_1+K_3)\boldsymbol{b}_i = \mathbf{0} \text{ for } j \geq 2, \text{ and } i = 1,2. \tag{5.24}$$

We should note that K_1+K_3 is positive definite by coercivity of the bilinear operator $B_{\mathcal{S}}$. This implies that $\boldsymbol{b}_i^T(K_1+K_3)\boldsymbol{b}_i > 0$ for $i = 1,2$.

Let us consider the case when $i = 1$ in (5.25). Since any constant function satisfies the first equation in (5.19), we must have

$$K_2\boldsymbol{b}_1 + (K_1+K_3)\boldsymbol{b}_1 + K_2^T\boldsymbol{b}_1 = \mathbf{0}, \tag{5.25}$$

which leads to $\boldsymbol{b}_1^T(K_1+K_3)\boldsymbol{b}_1 = -2\boldsymbol{b}_1^T K_2\boldsymbol{b}_1$. Hence, when $i = 1$ (5.24) takes the form

$$(\mu_{(1,j-1)} + \mu_{(1,j+1)}) - 2\mu_{(1,j)} = 0.$$

The general solution of this difference equation is $\mu_{(1,j)} = \gamma_1 + \gamma_2 j$ for some $\gamma_1, \gamma_2 \in \mathbb{R}$, and $j \geq 1$. Using the decay condition imposed on $\hat{w}$, γ_2 must be zero.

Next, we consider the case when $i = 2$ in (5.24). If $\boldsymbol{b}_2^T K_2\boldsymbol{b}_2 = 0$, then $K_2\boldsymbol{b}_2 = \gamma_3\boldsymbol{b}_1$ for some $\gamma_3 \in \mathbb{R}$ and

$$\mu_{(2,j)} = \begin{cases} 0 & \text{for } j = 1 \text{ and } \gamma_3 \neq 0 \\ \gamma_2 & \text{for } j = 1 \text{ and } \gamma_3 = 0 \\ 0 & \text{for } j \geq 1 \end{cases}$$

where γ_2 is some real constant. If $\boldsymbol{b}_2^T K_2\boldsymbol{b}_2 \neq 0$, then (5.24), for i=2, takes the form

$$(\mu_{(2,j-1)} + \mu_{(2,j+1)}) - 2\rho\mu_{(2,j)} = 0$$

where $-2\rho = \boldsymbol{b}_2^T(K_1+K_3)\boldsymbol{b}_2/\boldsymbol{b}_2^T K_2\boldsymbol{b}_2$. The characteristic equation for (5.24) is $t^2 - 2\rho t + 1 = 0$, and, therefore, the characteristics of the difference equation in (5.24) are reciprocal of each other and their magnitude is 1 if they are complex. Hence, and with the decay condition in mind, we have

$$\mu_{(2,j)} = \begin{cases} 0 & |\rho| \leq 1, j \geq 1 \\ \gamma_2[\text{sign}(\rho)(|\rho| - \sqrt{\rho^2-1})]^j & |\rho| > 1, j \geq 1 \end{cases}$$

for some real constant γ_2.

It remains to determine γ_1 and γ_2 to complete the solution. For that we use (5.21) when $m = 1$, i.e.,

$$\mu_{(1,0)}K_2\boldsymbol{b}_1 + \gamma_1((K_1 + K_3) + K_2^T)\boldsymbol{b}_1 + \mu_{(2,0)}K_2\boldsymbol{b}_2 + \mu_{(2,1)}(K_1 + K_3)\boldsymbol{b}_2 + \mu_{(2,2)}K_2^T\boldsymbol{b}_2 = 0.$$

Using (5.25), last equation is reduced to

$$(\mu_{(1,0)} - \gamma_1)K_2 b_1 + \mu_{(2,0)}K_2\boldsymbol{b}_2 + \mu_{(2,1)}(K_1 + K_3)\boldsymbol{b}_2 + \mu_{(2,2)}K_2^T\boldsymbol{b}_2 = 0. \tag{5.26}$$

At this point, we have one the following four cases.

Case 1. If $\boldsymbol{b}_2^T K_2\boldsymbol{b}_2 = 0$ and $K_2\boldsymbol{b}_2 = \gamma_3\boldsymbol{b}_1$ for some nonzero $\gamma_3 \in \mathbb{R}$, then (5.26) takes the form

$$(\mu_{(1,0)} - \gamma_1)K_2\boldsymbol{b}_1 + \mu_{(2,0)}\gamma_3\boldsymbol{b}_1 = 0$$

and we have only one unknown. Premultiplying by $\boldsymbol{b}_1^T$, we get

$$\gamma_1 = \mu_{(1,0)} + \frac{2\mu_{(2,0)}\gamma_3}{\boldsymbol{b}_1^T K_2\boldsymbol{b}_1}.$$

Since $\mu_{(2,0)} \neq 0$ by assumption, $\boldsymbol{b}_1$ must be an eigenvector of K_2; otherwise there will be no solution and this contradicts the existence of $\hat{w}$.

Case 2. If $\boldsymbol{b}_2^T K_2\boldsymbol{b}_2 = 0$ and $K_2\boldsymbol{b}_2 = \boldsymbol{0}$, then (5.26) takes the form

$$(\mu_{(1,0)} - \gamma_1)K_2\boldsymbol{b}_1 + \gamma_2(K_1 + K_3)\boldsymbol{b}_2 = 0$$

and we have two unknowns. In this case, $K_2\boldsymbol{b}_1$ and $(K_1+K_3)\boldsymbol{b}_2$ must be independent; otherwise we will have infinitely many solutions contradicting the fact that $\hat{w}$ is unique. This leads to $\gamma_2 = 0$ and $\gamma_1 = \mu_{(1,0)}$.

Case 3. If $\boldsymbol{b}_2^T K_2\boldsymbol{b}_2 \neq 0$ and $|\rho| \leq 1$, then (5.26) takes the form

$$(\mu_{(1,0)} - \gamma_1)K_2\boldsymbol{b}_1 + \mu_{(2,0)}K_2\boldsymbol{b}_2 = 0.$$

Since $\mu_{(2,0)} \neq 0$, we must have $K_2\boldsymbol{b}_2 = \gamma_3 K_2\boldsymbol{b}_1$ for some $\gamma_3 \in \mathbb{R}$. Hence,

$$\gamma_1 = \mu_{(1,0)} + \mu_{(2,0)}\gamma_3.$$

Case 4. If $\boldsymbol{b}_2^T K_2\boldsymbol{b}_2 \neq 0$ and $|\rho| > 1$, then (5.26) takes the form

$$(\mu_{(1,0)} - \gamma_1)K_2\boldsymbol{b}_1 + (\mu_{(2,0)} - \gamma_2)K_2\boldsymbol{b}_2 = 0.$$

For uniqueness of $\hat{w}$, we must have K_2 nonsingular Hence,

$$\gamma_i = \mu_{(i,0)} \text{ for } i = 1, 2.$$

The previous results are summarized in the following theorem.

Theorem 5.4. *Let $\mathcal{C} = [-1,1]\times[0,2a]$, where a is the cell aspect ratio, and let the triangulation on $\mathcal{C}$ be a distorted version of one of the basic patterns Regular, Chevron, or Union Jack, where the basic pattern is distorted by moving its central node. Let $\hat{q}^{bl}$ be the boundary layer corresponding to $\hat{q}$ on $\mathcal{S}$. To simplify notations, set $\hat{w} = \hat{q}^{bl}$, i.e., $\hat{w}$ is the solution of* (5.19). *Let $\boldsymbol{b}_1, \boldsymbol{b}_2$, and $\hat{\boldsymbol{w}}_P^{(2j)}$, for $j \geq 0$, be defined as in Example 2, and let K_1, K_2, K_3 be defined as in* (5.20). *Then,*

$$\hat{\boldsymbol{w}}_P^{(2j)} = \mu_{(1,j)}\boldsymbol{b}_1 + \mu_{(2,j)}\boldsymbol{b}_2 \text{ for } j \geq 0$$

where $\mu_{(1,0)}$ and $\mu_{(2,0)}$ are determined from $\hat{\boldsymbol{w}}_P^{(0)}$ given in the boundary conditions, and $\mu_{(1,j)} = \gamma_1$ for all $j \geq 1$ for some $\gamma_1 \in \mathbb{R}$. Without loss of generality, assume that $\mu_{(2,0)} \neq 0$; otherwise apply Remark 5.3. If $\boldsymbol{b}_2^T K_2 \boldsymbol{b}_2 \neq 0$, set $-2\rho = \boldsymbol{b}_2^T(K_1 + K_3)\boldsymbol{b}_2 / \boldsymbol{b}_2^T K_2 \boldsymbol{b}_2$. To determine γ_1 and $\mu_{(2,j)}$ for $j \geq 1$, there are four cases:

1. *If $K_2\boldsymbol{b}_2 = 0$, then $\mu_{(2,j)} = 0$ for all $j \geq 1$ and $\gamma_1 = \mu_{(1,0)}$.*

2. *If $K_2\boldsymbol{b}_2 = \gamma_3\boldsymbol{b}_1$ for some nonzero $\gamma_3 \in \mathbb{R}$, then $\mu_{(2,j)} = 0$ for all $j \geq 1$, $\boldsymbol{b}_1$ is an eigenvector of K_2, and*
$$\gamma_1 = \mu_{(1,0)} + \frac{2\mu_{(2,0)}\gamma_3}{\boldsymbol{b}_1^T K_2 \boldsymbol{b}_1}.$$

3. *If $\boldsymbol{b}_2^T K_2 \boldsymbol{b}_2 \neq 0$ and $|\rho| \leq 1$, then $\mu_{(2,j)} = 0$ for all $j \geq 1$, $K_2\boldsymbol{b}_2 = \gamma_3 K_2\boldsymbol{b}_1$ for some $\gamma_3 \in \mathbb{R}$, and $\gamma_1 = \mu_{(1,0)} + \mu_{(2,0)}\gamma_3$.*

4. *If $\boldsymbol{b}_2^T K_2 \boldsymbol{b}_2 \neq 0$ and $|\rho| > 1$, then $\gamma_1 = \mu_{(1,0)}$, $\mu_{(2,j)} = \mu_{(2,0)}[\text{sign}(\rho)(|\rho| - \sqrt{\rho^2 - 1})]^j$ for all $j \geq 1$, and K_2 is nonsingular.*

6 Asymptotic Quality of the PPR Error Indicators over Patches Adjacent to $\partial\Omega$

In Section 5 it was shown how to extend asymptotic finite element approximation up to the boundary in such a way that boundary conditions are satisfied. This implies that error indicators over patches adjacent to the boundary can be evaluated using the computer-based methodology explained in Section 3 after modifying the definition of $\hat{q}^*$ as was explained in Section 5. With this in mind, the performance of the PPR error indicators over patches adjacent to the boundary will be studied and compared to that of SPR error indicators using the tests in Section 4. In every one of these tests the asymptotic quality is evaluated through the effectivity index (or the robustness index) of the error indicator when applied to $\hat{q}^*$ defined by (5.2) over $\mathcal{C}$. The following sums up the steps to be done in every test:

1. Choose $\mathcal{D}$, aspect ratio a of $\mathcal{C} = [-1, 1] \times [0, 2a]$, and the triangulation pattern.

2. For $i = 1, 2, 3$, compute $\hat{q}_i^{per} = \Pi_{\mathcal{C}}^{per}(\hat{q}_i - \hat{q}_{i,I})$ as explained in Example 1.

3. If Neumann boundary conditions are assumed, $\hat{q}^{bl} = 0$, but we need to find the combinations c_1, c_2, and c_3 such that $\int_{\hat{\Gamma}} \partial_{\hat{y}} \hat{g} = 0$ where $\hat{g}$ is defined in (5.5).

4. If Dirichlet boundary conditions are assumed, compute $\hat{q}_i^{bl}$ for $i = 1, 2, 3$ as explained in Theorem 5.4.

5. At this stage, $\hat{q}^*$ is well defined and its recovered gradient can be constructed on $\mathcal{C}$. Recovering the gradient using the PPR is tricky and the recovered gradient has to be constructed for some nodes attached to $\mathcal{C}$ before processing its nodes. Fig. 2 is helpful in determining such required nodes.

6. Having $\hat{q}^*$ and its corresponding recovered gradients, M_e and M_a are computed using their definitions in Section 3.

Throughout the tests in this section, $\mathcal{Q} = \{\sum_{i=1}^{3} c_i \hat{q}_i : c_1, c_2, c_3 \in \mathbb{R}\}$. Of course, there are be some limitations on c_1, c_2, and c_3 in case of Neumann conditions. The results are very much the same if polynomials in $\mathcal{Q}$ are harmonic, and so the class of harmonic polynomials will not have special treatment in this section.

Among all the tests, excessive mesh distortion has the severest effect on the performance for both the PPR and the SPR error indicators. The results for mesh distortion tests are reported in figures Fig. 16 through Fig. 23. A glimpse over these figures reveals the following:

- For small distortion, the PPR error indicator is performing better than the SPR error indicator.

- In the Regular and the Chevron patterns, the PPR and the SPR error indicators are comparable while the PPR error indicator is doing better in both the Union Jack and the Criss Cross patterns.

- The PPR error indicator is asymptotically exact for zero distortion except for the case in which the pattern is the Union Jack and the boundary condition type is Dirichlet.

As shown in Table 2 and Fig. 24, changing the aspect ratio of $\mathcal{C}$ has very little effect on the performance of both the PPR and the SPR error indicators, but the PPR is relatively better. Note that the PPR error indicator is asymptotically exact regardless of a, except for the case in which the pattern is the Union Jack and the boundary condition type is Dirichlet.

Pattern	Dirichlet Boundary Conditions		Neumann Boundary Conditions	
	PPR	SPR	PPR	SPR
Regular	$\lambda_{\min}=1$ $\lambda_{\max}=1$	$\lambda_{\min}=1$ $\lambda_{\max}=1$	$\lambda_{\min}=1$ $\lambda_{\max}=1$	$\lambda_{\min}=1$ $\lambda_{\max}=1$
Chevron	$\lambda_{\min}=1$ $\lambda_{\max}=1$	$\lambda_{\min}=\frac{50a^2+63}{63(a^2+1)}$ $\lambda_{\max}=\frac{131}{126}$	$\lambda_{\min}=1$ $\lambda_{\max}=1$	$\lambda_{\min}=\frac{50a^2+63}{63(a^2+1)}$ $\lambda_{\max}=1$
Union Jack	See Fig. 24(b)	See Fig. 24(b)	$\lambda_{\min}=1$ $\lambda_{\max}=1$	See Fig. 24(a)
Criss Cross	$\lambda_{\min}=1$ $\lambda_{\max}=1$	See Fig. 24(d)	$\lambda_{\min}=1$ $\lambda_{\max}=1$	See Fig. 24(c)

Table 2. Response of λ_{min} and λ_{Max}, or $\mathcal{R}$, to change in cell aspect ratio when the patch is adjacent to the boundary.

The material properties have very little effect on the PPR error indicator as shown in Table 3 and Fig. 25. Again, the PPR error indicator is asymptotically exact regardless of θ and d except for the case in which the pattern is the Union Jack and the boundary condition type is Dirichlet.

Pattern	Dirichlet Boundary Conditions		Neumann Boundary Conditions	
	PPR	SPR	PPR	SPR
Regular	$\lambda_{\min}=1$ $\lambda_{\max}=1$	$\lambda_{\min}=1$ $\lambda_{\max}=1$	$\lambda_{\min}=1$ $\lambda_{\max}=1$	$\lambda_{\min}=1$ $\lambda_{\max}=1$
Chevron	$\lambda_{\min}=1$ $\lambda_{\max}=1$	$\lambda_{\min}=\frac{113}{126}$ $\lambda_{\max}=\frac{131}{126}$	$\lambda_{\min}=1$ $\lambda_{\max}=1$	$\lambda_{\min}=\frac{113}{126}$ $\lambda_{\max}=1$
Union Jack	See Fig. 25(b)	See Fig. 24(b)	$\lambda_{\min}=1$ $\lambda_{\max}=1$	See Fig. 25(a)
Criss Cross	$\lambda_{\min}=1$ $\lambda_{\max}=1$	See Fig. 24(d)	$\lambda_{\min}=1$ $\lambda_{\max}=1$	See Fig. 25(c)

Table 3. Response of λ_{min} and λ_{max}, or $\mathcal{R}$, to change in θ, the material orthotropy orientation, when $d = 100$, and when the patch is adjacent to the boundary

Remark 6.1. The definition of the recovered gradient at nodes on $\partial\Omega$ greatly affects the quality of error indicators over patches adjacent to the boundary. In some cases, like the PPR and the SPR, there may be more than one possible definition. Studying the asymptotic quality of the error indicators can help choosing the best one. This strategy was used in [5] to choose

the right definition for the SPR-recovered gradient at the boundary nodes, and the same can be done in the case of the PPR recovery. Following are some of the possible definitions and their disadvantages. One of the possible definitions treats boundary nodes in the same way they are treated in the SPR recovery. The disadvantages of this definition are: (1) incapability to handle nodes that are not attached to internal nodes, and (2) the resulting error indicators are sensitive to cell aspect ratio. Another definition treats boundary nodes in the same way internal nodes are treated in the PPR recovery. The patch corresponding to a boundary node z is constructed by either extending out (as in [13, 14]) or by including two or more layers of nodes around z. This definition leads to error indicators that are sensitive to cell aspect ratio. In both of these two definitions the resulting error indicators are very robust under changes in material properties and mesh distortion.

In conclusion, it was shown that the PPR error indicator performs as good as or better than the SPR error indicator. It seems that the material properties and the cell aspect ratio have very little effect on both of the two error indicators. Also, it seems that mesh geometry is the most important factor that affects the performance of both of the two error indicators.

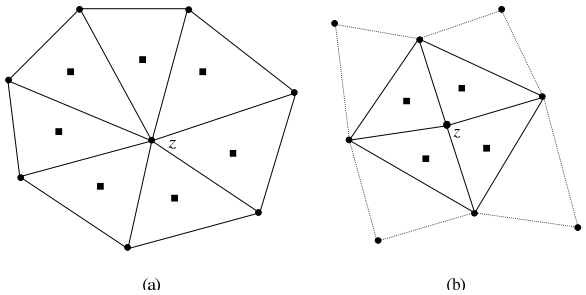


Fig. 1. Patch required for gradient recovery. Sampling points for SPR are marked with ■, while those needed for PPR are marked with •

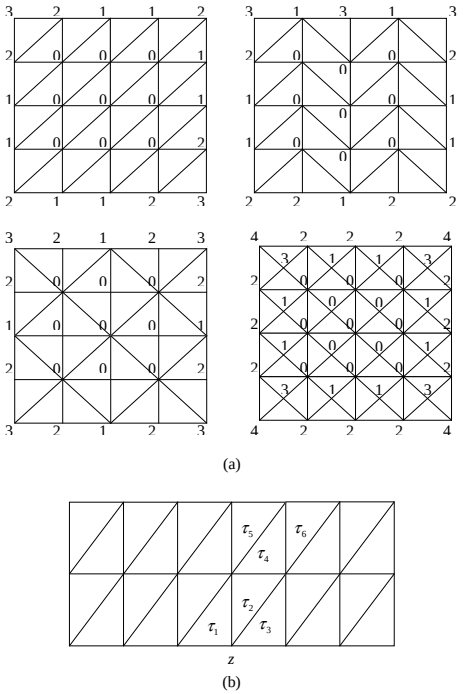


Fig. 2. Recovering the gradient at boundary nodes using PPR.

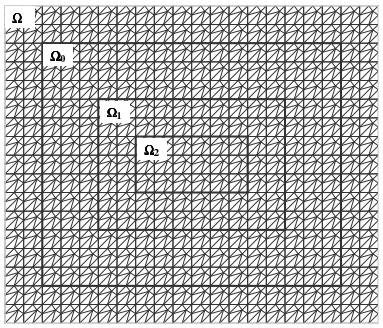
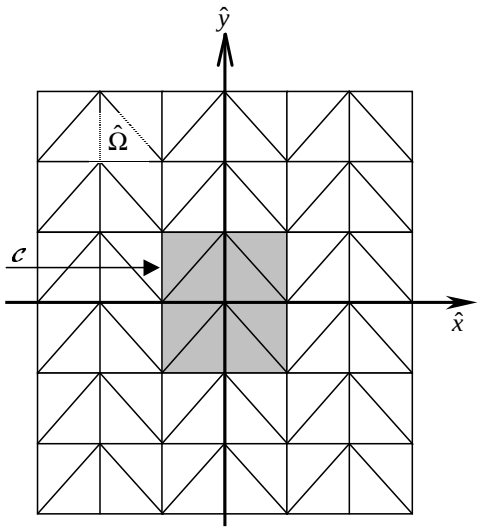
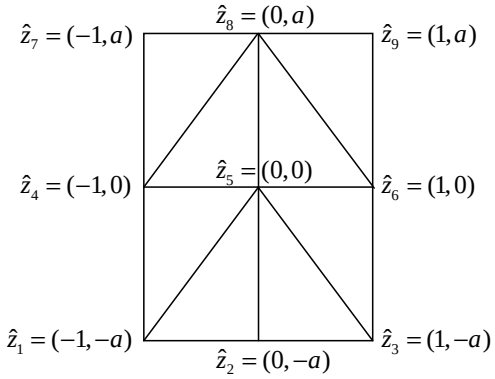


Fig. 3. An example of a uniform periodic translation invariant mesh on Ω_0.

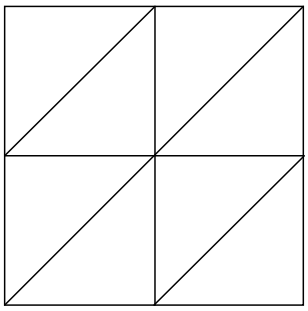


(a) The reference domain $\hat{\Omega}$

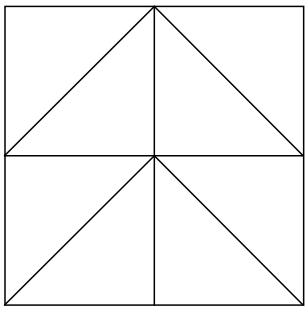


(b) Mesh nodes and triangles in $\mathcal{C}$

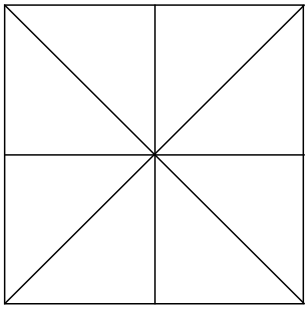
Fig. 4. Reference domain $\hat{\Omega}$, reference cell $\mathcal{C}$, and the mesh on $\mathcal{C}$ for example 1.



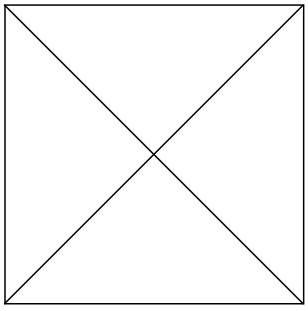
(a) Regular pattern



(b) Chevron pattern

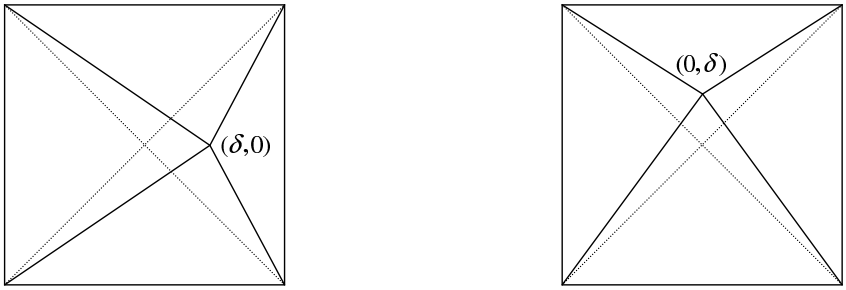


(c) Union Jack pattern



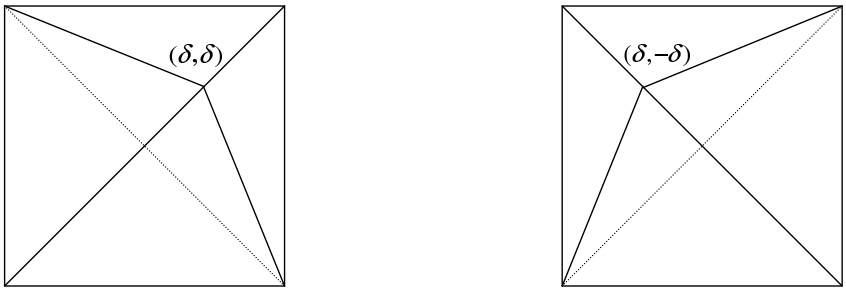
(c) Criss Cross pattern

Fig. 5. The four patterns used in partitioning the reference cell.



(a) Moving the central node along the line $\hat{y}=0$

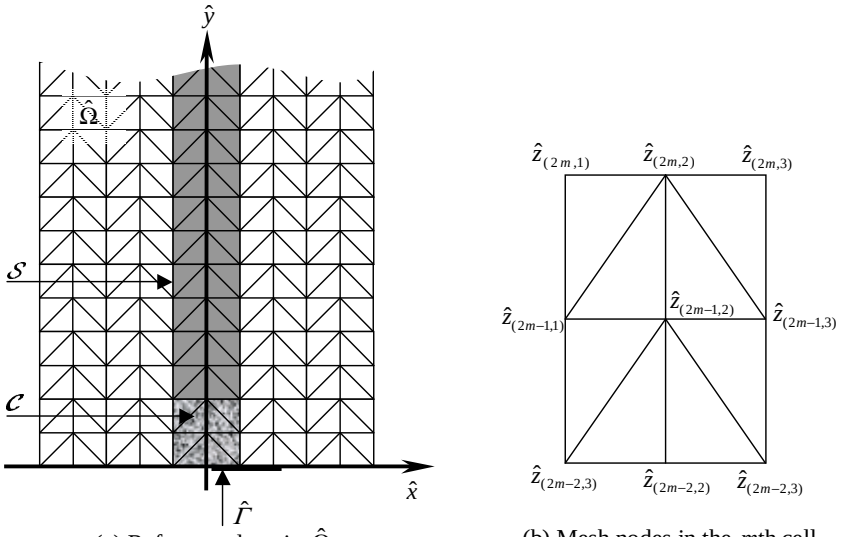
(b) Moving the central node along the line $\hat{x}=0$



(c) Moving the central node along the line $\hat{y}=\hat{x}$

(d) Moving the central node along the line $\hat{y}=-\hat{x}$

Fig. 6. Distorting the mesh on $\mathcal{C}$ from Criss Cross pattern by moving the central node.



(a) Reference domain $\hat{\Omega}$

(b) Mesh nodes in the mth cell

Fig. 7. Reference domain $\hat{\Omega}$, reference strip $\mathcal{S}$, and the mesh on mth cell for example 2.

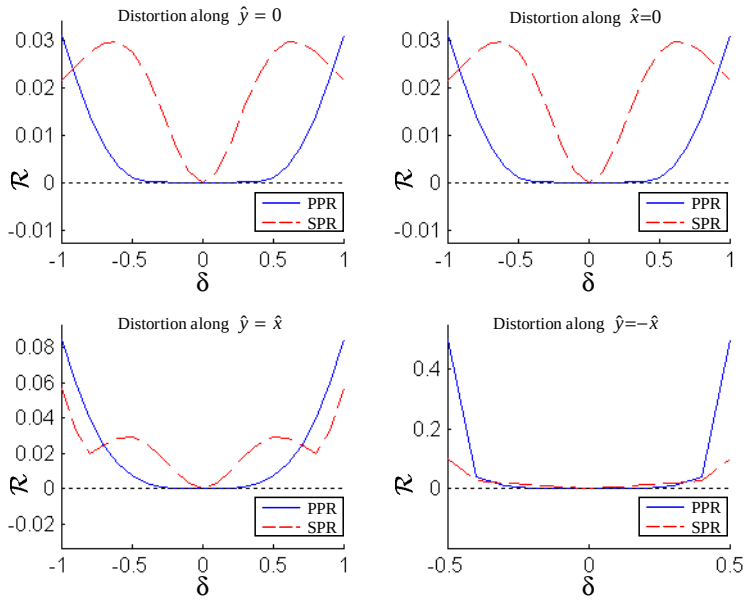


Fig. 8. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Regular pattern for the class of general polynomials.

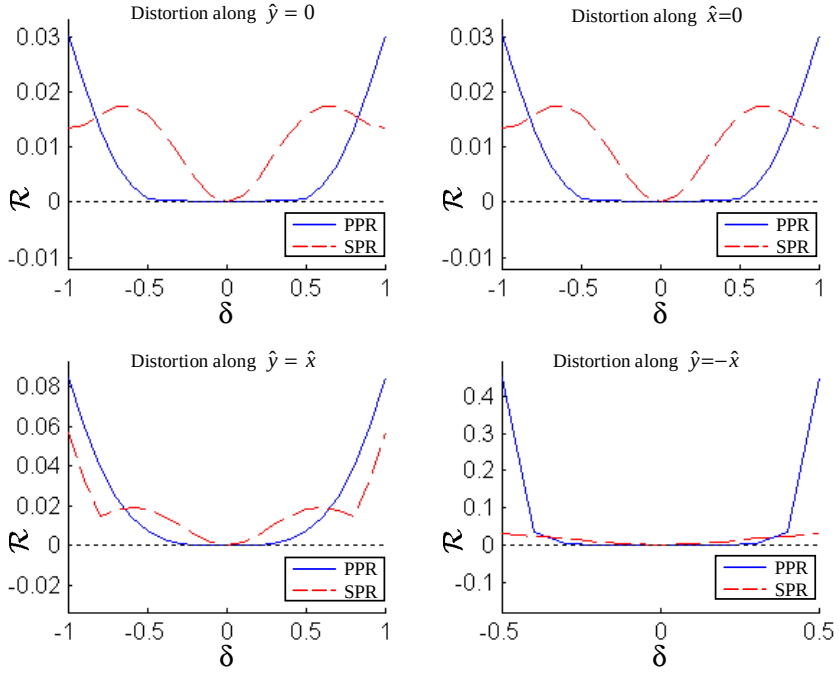


Fig. 9. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Regular pattern for the class of harmonic polynomials.

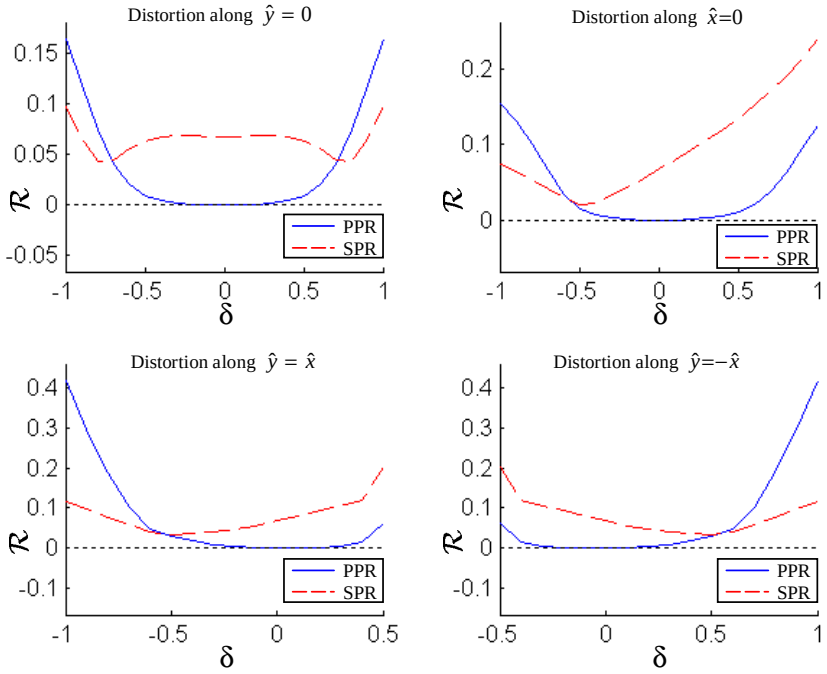


Fig. 10. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Chevron pattern for the class of general polynomials.

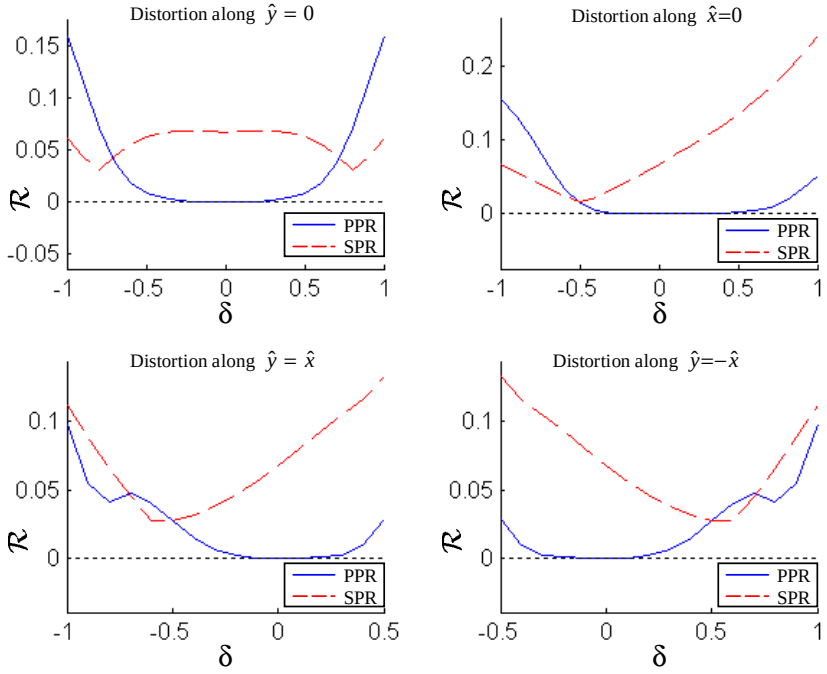


Fig. 11. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Chevron pattern for the class of harmonic polynomials.

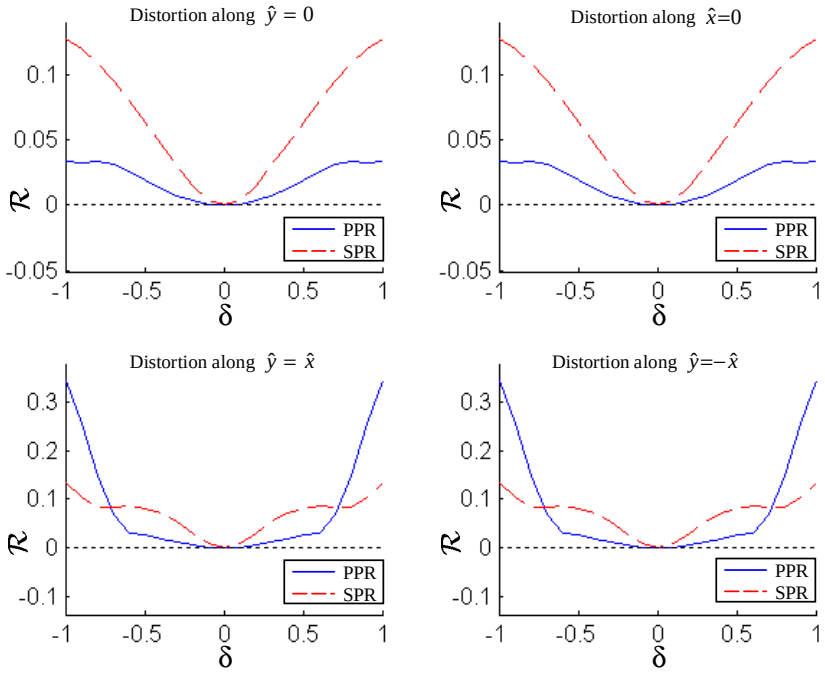


Fig. 12. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Union Jack pattern for the class of general polynomials.

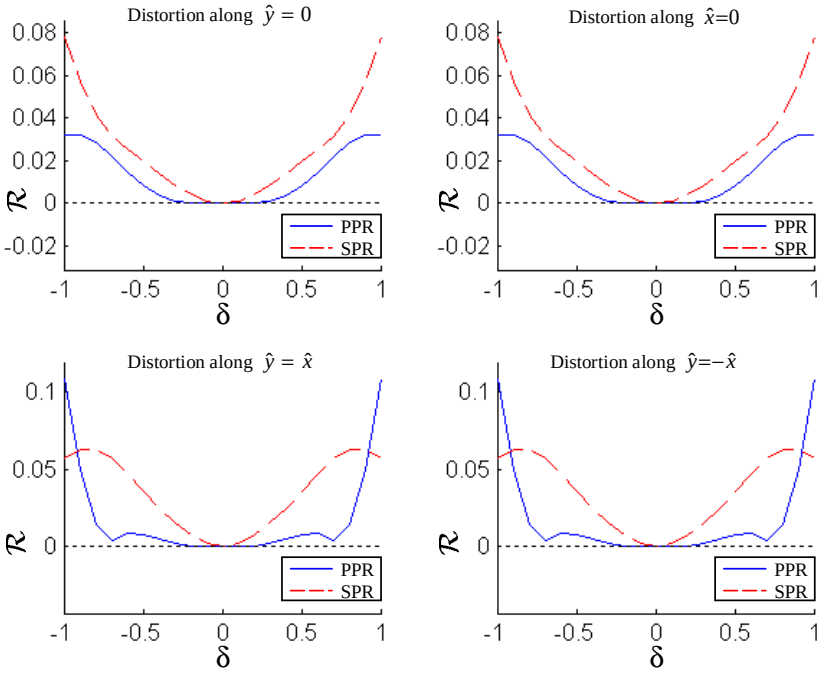


Fig. 13. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Union Jack pattern for the class of harmonic polynomials.

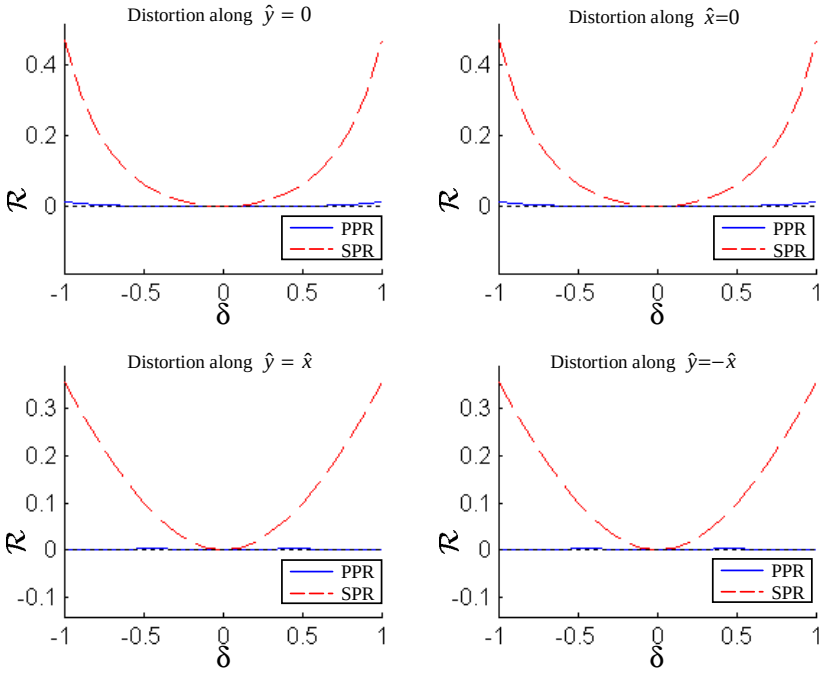


Fig. 14. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Criss Cross pattern for the class of general polynomials.

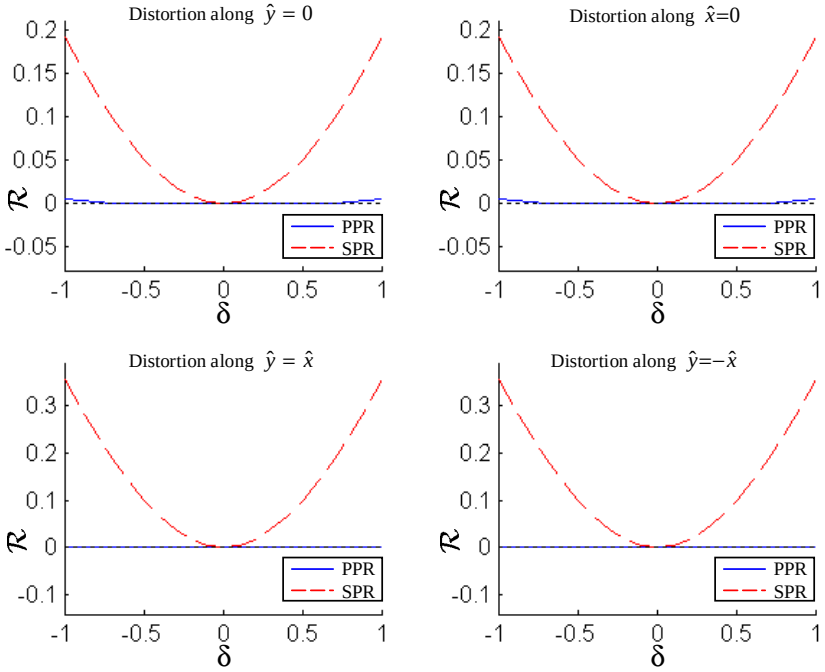


Fig. 15. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Criss Cross pattern for the class of harmonic polynomials.

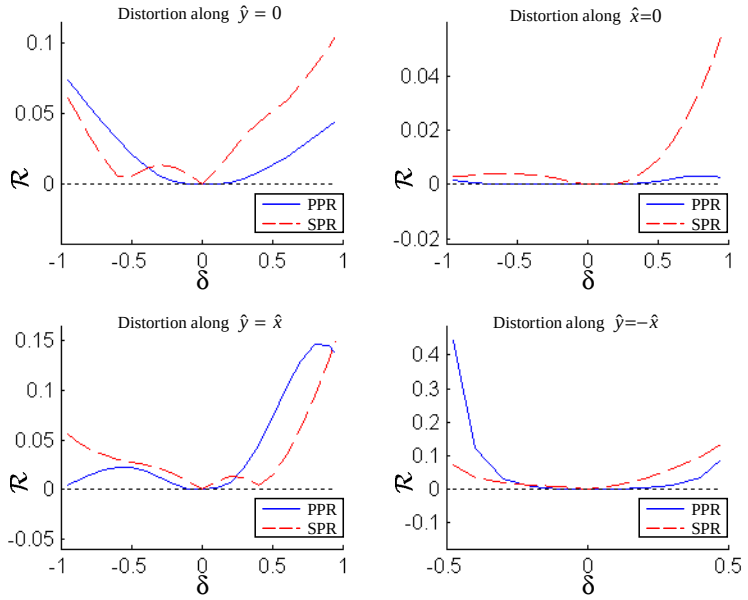


Fig. 16. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Regular pattern for the class of general polynomials. Boundary condition on $\hat{\Gamma}$ is of Neumann type.

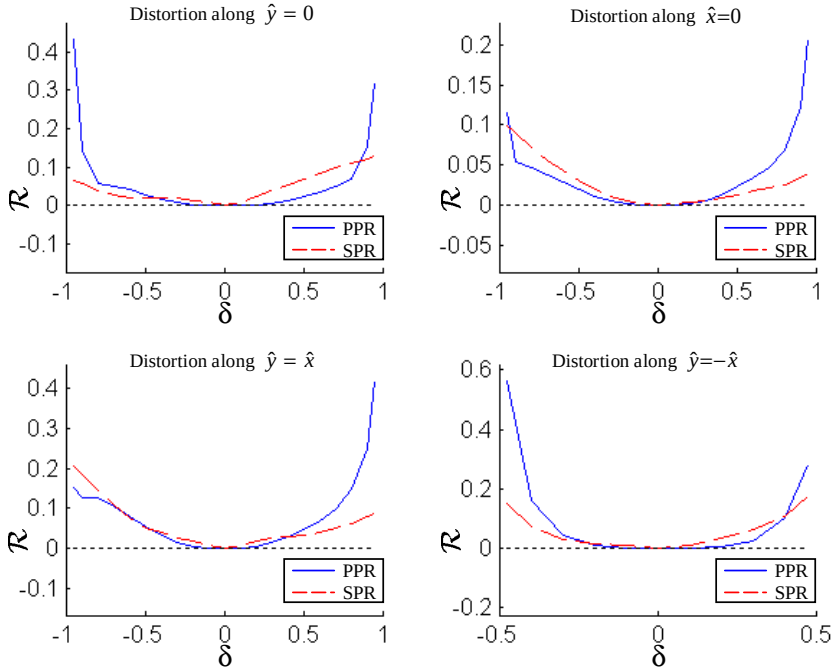


Fig. 17. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Regular pattern for the class of general polynomials. Boundary condition on $\hat{\Gamma}$ is of Dirichlet type.

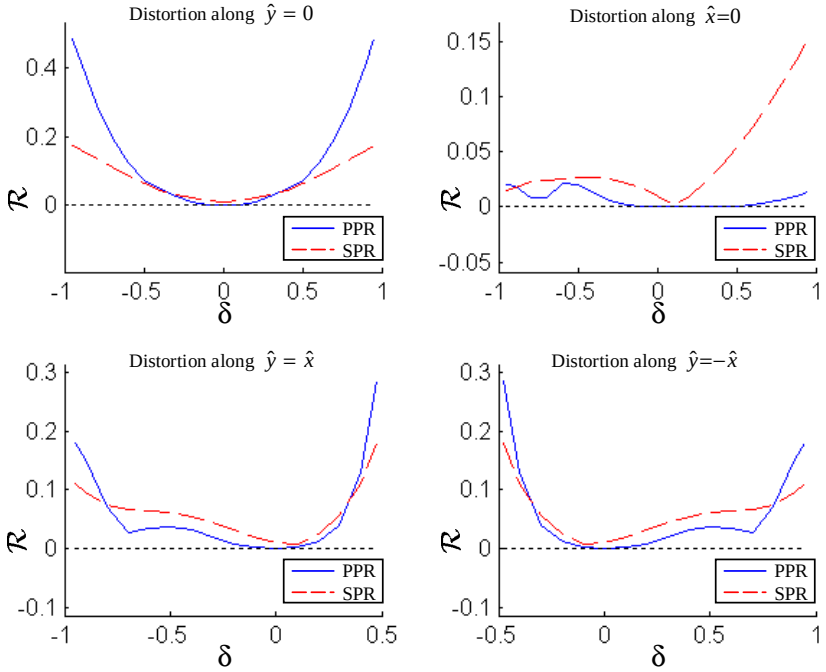


Fig. 18. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Chevron pattern for the class of general polynomials. Boundary condition on $\hat{\Gamma}$ is of Neumann type.

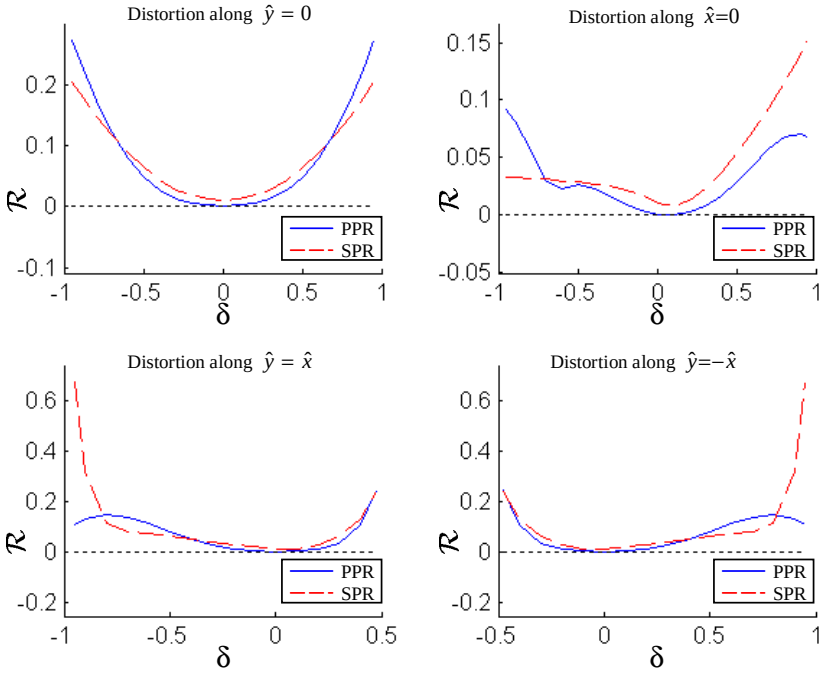


Fig. 19. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Chevron pattern for the class of general polynomials. Boundary condition on $\hat{\Gamma}$ is of Dirichlet type.

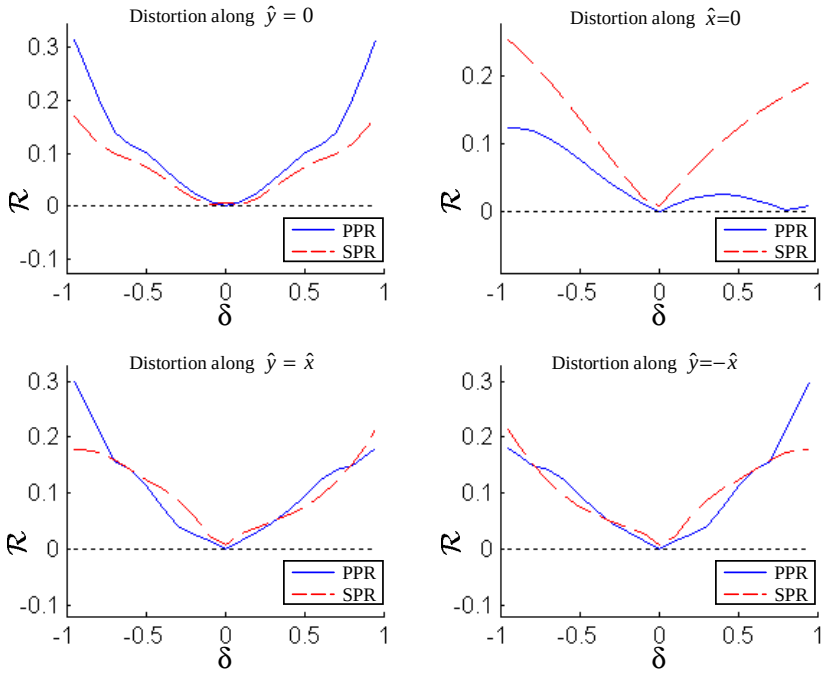


Fig. 20. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Union Jack pattern for the class of general polynomials. Boundary condition on $\hat{\Gamma}$ is of Neumann type.

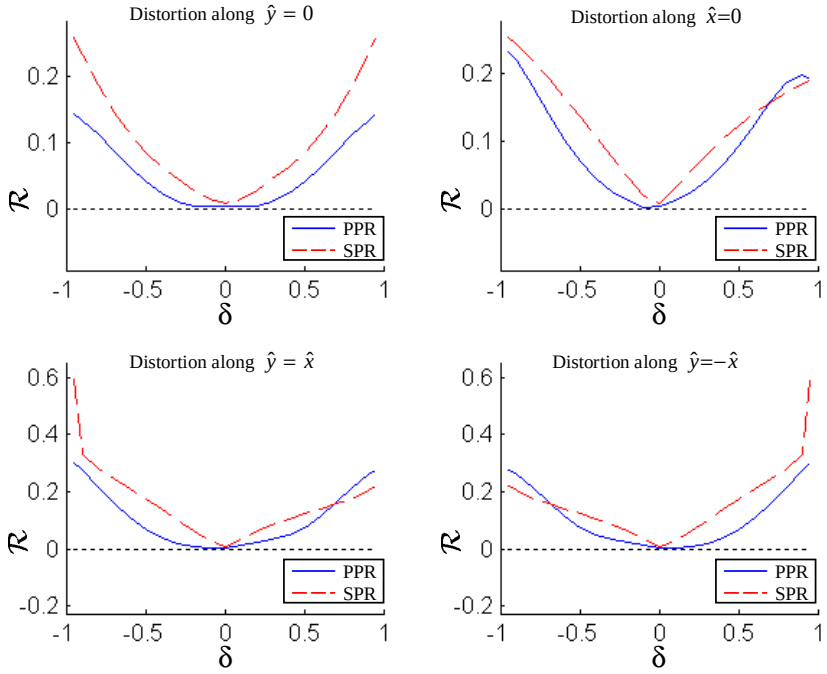


Fig. 21. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Union Jack pattern for the class of general polynomials. Boundary condition on $\hat{\Gamma}$ is of Dirichlet type.

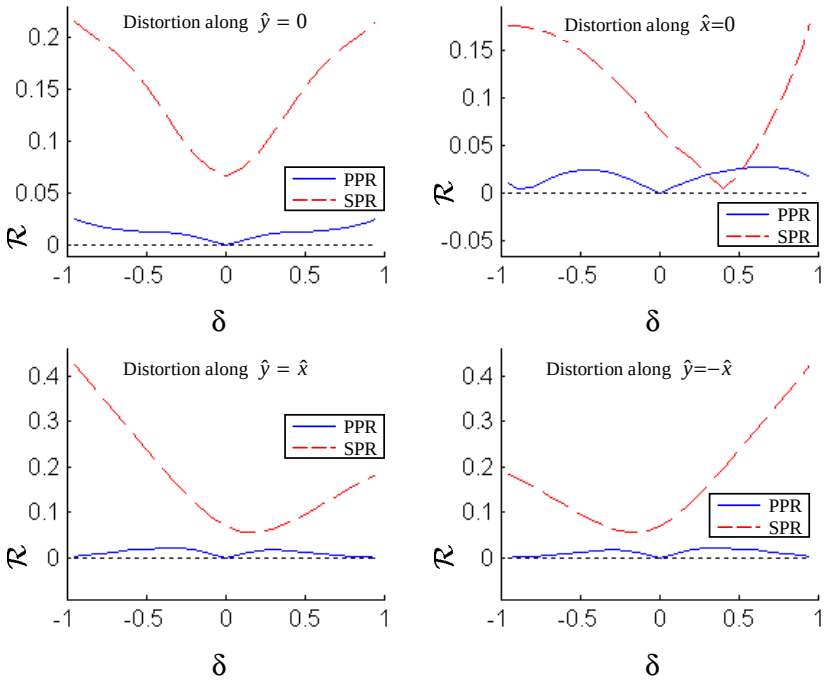


Fig. 22. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Criss Cross pattern for the class of general polynomials. Boundary condition on $\hat{\Gamma}$ is of Neumann type.

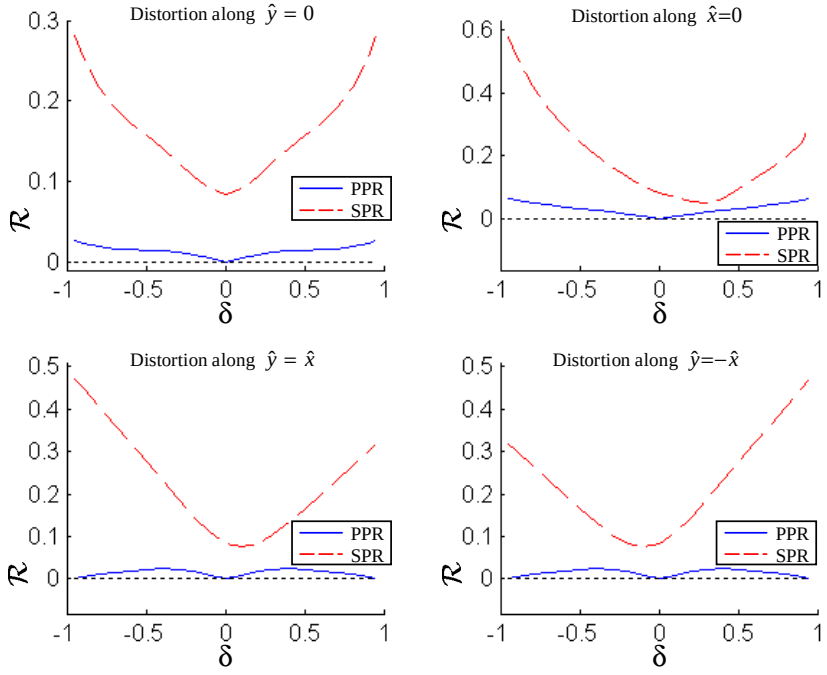
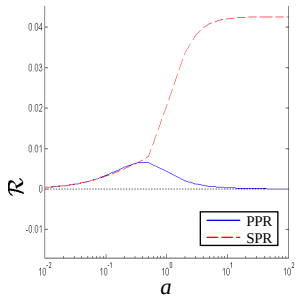
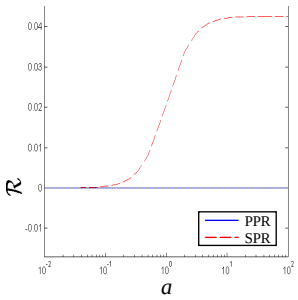


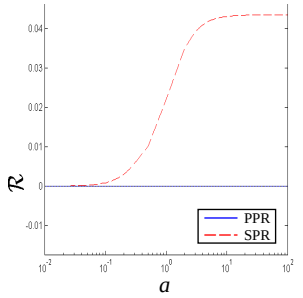
Fig. 23. The change in the robustness index as the mesh on $\mathcal{C}$ is distorted form Criss Cross pattern for the class of general polynomials. Boundary condition on $\hat{\Gamma}$ is of Dirichlet type.



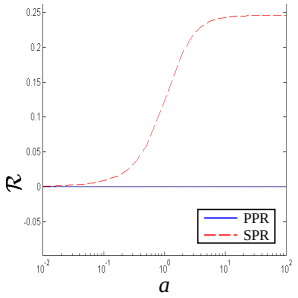
(a) Union Jack, Neumann



(b) Union Jack, Dirichlet

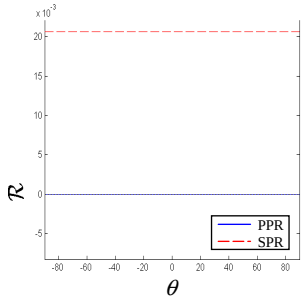


(c) Criss Cross, Neumann

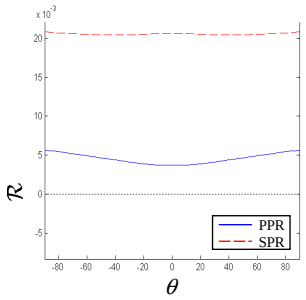


(d) Criss Cross, Dirichlet

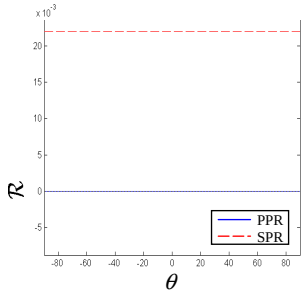
Fig. 24. The change in the robustness index as the aspect ratio of $\mathcal{C}$ changes when the mesh pattern on $\mathcal{C}$ is Union Jack or Criss Cross. For the other two cases, see Table 2



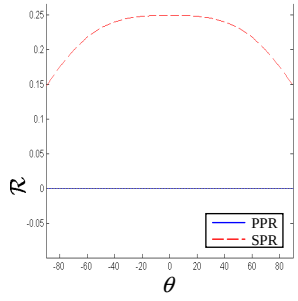
(a) Union Jack, Neuman



(b) Union Jack, Dirichlet



(c) Criss Cross, Neuman



(d) Criss Cross, Dirichlet

Fig. 25. The change in the robustness index as the material orthotropy orientation angle θ changes from $-90°$ to $90°$ when $d = 100$. The mesh pattern on $\mathcal{C}$ is Union Jack or Criss Cross. For the other two cases, see Table 3

.

References

[1] Ainsworth M, Oden JT. *A Poseriori Error Estimation in Finite Element Analysis*, Wiley Interscience: New York, 2000; 145-187.

[2] Babuška I, Aziz AK. On the angle condition in the finite element method. *SIAM Journal on Numerical Analysis* 1976; **13**:214-226.

[3] Babuška I, Strouboulis T. *The Finite Element Method and Its Reliability*, Oxford University Press: London, 2001.

[4] Babuška I, Strouboulis T, Upadhyay CS. A model study of the quality of *a posteriori* error estimators for linear elliptic problems. Error Estimation in the interior of patchwise uniform grids of triangles. *Computer Methods in Applied Mechanical Engineering* 1994; **114**:307-378.

[5] Babuška I, Strouboulis T, Upadhyay CS. A model study of the quality of *a posteriori* error estimators for finite element solutions of linear elliptic problems, with particular refrence to the behavior near the boundary. *International Journal for Numerical Methods in Engineering* 1997; **40**:2521-2577.

[6] Bank RE, Weiser A. Some *a posteriori* error estimators for elliptic partial differential equations. *Mathematics of Computation* 1985; **44**:283-301.

[7] Bank RE, Xu J. Asymptotically exact *a posteriori* error estimators, part I: grids with superconvergence. submitted

[8] Oden JT, Demkowicz L, Rachowicz W, Westermann TA. Toward a universal h-p adaptive finite element strategy: Part 2, *a posteriori* error estimates. *Computer Methods in Applied Mechanical Engineering* 1989; **77**:113-180.

[9] Verfürth R. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner: New York, 1996.

[10] Li XD, Wiberg N-E. *A posteriori* error estimate by element patch postprocessing, adaptive analysis in energy norm and L_2 norms. *Computers & Structures* 1994; **53**(4):907-919.

[11] Wiberg N-E, Li XD. Superconvergent patch recovery of finite element solution and *a posteriori* error L_2 norm estimate. *Communications in Numerical Methods in Engineering* 1994; **10**(4):313-320.

[12] Xu J, Zhang Z. Analysis of recovery type *a posteriori* error estimators for mildly structured grids. *Mathematics of Computations* 2003, accepted.

[13] Zhang Z, Naga A. A meshless gradient recovery method, part I: superconvergence recovery. Research Report #2 (2002), Department of Mathematics, Wayne State University.

[14] Zhang Z, Naga A. *A posteriori* error estimates based on polynomial preserving recovery. Research Report #9 (2002), Department of Mathematics, Wayne State University.

[15] Zienkiewicz OC, Zhu JZ. A simple error estimator and adaptive procedure for practical engineering analysis. *International Journal for Numerical Methods in Engineering* 1987; **24**:337-357.

[16] Zienkiewicz OC, Zhu JZ. The superconvergence patch recovery and *a posteriori* error estimates, part I: the recovery technique. *International Journal for Numerical Methods in Engineering* 1992; **33**:1331-1364.

[17] Zienkiewicz OC, Zhu JZ. The superconvergence patch recovery and *a posteriori* error estimates, part II: error estimates and adaptivity. *International Journal for Numerical Methods in Engineering* 1992; **33**:1365-1382.