SIAM J. NUMER. ANAL.
Vol. 54, No. 4, pp. 2302–2322

ULTRACONVERGENCE OF FINITE ELEMENT METHOD BY RICHARDSON EXTRAPOLATION FOR ELLIPTIC PROBLEMS WITH CONSTANT COEFFICIENTS*

WEN-MING HE†, RUNCHANG LIN‡, AND ZHIMIN ZHANG§

Abstract. In this article, two novel Richardson extrapolation operators P_1^k and P_2^k are proposed to investigate local $2k$ order ultraconvergence properties of the kth order Lagrange finite element method for the second order elliptic problem with constant coefficients. Assume that $\mathbf{x}_0$ is an interior mesh node of the underlying mesh which is away from the boundary for a fixed distance unchanging with further mesh refinement. We show that, for both tensor product $\mathbb{Q}_k$ element and simplicial $\mathbb{P}_k$ element, it holds $|(u - P_1^k u^h)(\mathbf{x}_0)| \le ch^{2k}|\ln h|^{\tilde{k}+1}$ and $|(\nabla u - P_2^k(\overline{\nabla} u^h))(\mathbf{x}_0)| \le ch^{2k}|\ln h|^{\tilde{k}+1}$, where u^h is the finite element approximation of u, $\overline{\nabla}$ is defined in section 1.1, and $\tilde{k} = 1$ if $k = 1$ and $\tilde{k} = 0$ if $k > 1$. Numerical results are provided to demonstrate the theoretic findings.

Key words. finite element method, second order elliptic equation, ultraconvergence, Richardson extrapolation

AMS subject classifications. 65N15, 65N30

DOI. 10.1137/15M1031710

1. Introduction.

In this article, we will investigate local ultraconvergence of kth order Lagrangian finite element (FE) solutions to the following elliptic problem with Dirichlet boundary condition:

$$\begin{cases} \mathcal{L}u \equiv -\nabla \cdot (A\nabla u) = f & \text{in} \quad \Omega, \\ u = 0 & \text{on} \quad \partial\Omega, \end{cases} \tag{1.1}$$

where $\Omega \subset \mathbb{R}^n$ is a bounded open domain with boundary $\partial\Omega$, $n = 2$ or 3, $A = (a_{ij})_{n\times n}$ is a symmetric matrix, and f is sufficiently smooth. To fix the idea, we consider a homogeneous boundary condition in (1.1) for convenience. Since this work is for interior analysis, influence of different boundary conditions can be eliminated by cut-off functions as demonstrated in the subsequent analysis. We assume that A is uniformly elliptic in $\mathbb{R}^n$, i.e., there exists a constant $M > 0$ such that

$$\mathbf{x}^T A\mathbf{x} \ge M\mathbf{x}^T\mathbf{x} \quad \forall \mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n,$$

where $\mathbf{x}^T$ represents the transpose of $\mathbf{x}$.

*Received by the editors July 22, 2015; accepted for publication (in revised form) May 17, 2016; published electronically July 19, 2016.

http://www.siam.org/journals/sinum/54-4/M103171.html

†Department of Mathematics, Wenzhou University, Wenzhou, Zhejiang 320035, China (he_wenming@aliyun.com). The research of this author was supported by the National Natural Science Foundation of China under grant 11171257 and the Zhejiang Provincial Natural Science Foundation, China (LY15A010015).

‡Department of Mathematics and Physics, Texas A&M International University, Laredo, TX 78041 (rlin@tamiu.edu). The research of this author was supported in part by the US National Science Foundation grant DMS 1217268, the National Natural Science Foundation of China under grant 11428103, and a University Research Grant of Texas A&M International University.

§Corresponding author. Beijing Computational Science Research Center, Beijing 100193, China and Department of Mathematics, Wayne State University, Detroit, MI 48202 (zmzhang@csrc.ac.cn, zzhang@math.wayne.edu). The research of this author was supported in part by the US National Science Foundation through grant DMS 1419040 and by the National Natural Science Foundation of China under grants 91430216, 11471031.

We shall remark that the Richardson extrapolation formulation developed in this article relies on local symmetry of the errors in FE solutions over translation invariant meshes, which is well known in the superconvergence literature; see, e.g., [20, 27, 29]. In fact, Schatz and Wahlbin [22] and Wahlbin [27] developed the *symmetry theory* for the FE method (FEM) and proved natural superconvergence at *local mesh symmetry* points. In particular, they proved that the FE solution converges to the exact solution with order $\mathcal{O}(h^{k+2-\varepsilon})$ when k is even, and the averaged gradient of the FE solution converges in order $\mathcal{O}(h^{k+1-\varepsilon})$ when k is odd, for some $\varepsilon > 0$. Note that symmetry of errors may be distorted in anisotropic materials. Therefore, anisotropic problems will not be the focus in this work.

Richardson extrapolations for FEMs for elliptic problems have been presented since the 1970s (see, e.g., [6, 7, 13, 14, 18] for an incomplete list of references). Superconvergence properties have been studied by using this approach since then (see, e.g., [1, 2, 15, 19, 23]). On the other hand, since the paper of Douglas and Dupont [10], investigations of $2k$th order ultraconvergence of kth order Galerkin FE approximations have been carried out through numerous techniques. Bramble and Schatz [4] introduced a class of convolution operators and showed that the order of convergence of the displacement can be almost doubled by "averaging" the FE solutions if the exact solution is locally smooth. This method was extended by Thomée [25] to obtain a $2k$th order interior approximation for derivatives. The idea of averaging via convolution has also been applied to superconvergence studies for parabolic problems [26] and hyperbolic problems [9]. Recently, Chen and Hu [8] showed $2k$th order convergence of bi-kth order FE solutions to the Poisson equation under a strong global regularity assumption. He and Zhang [11] proved the same $2k$th rate by using an anisotropic mesh approximation in weighted Sobolev spaces under much weaker and practical regularity assumptions.

In this study, we present two novel Richardson extrapolation formulas to investigate local ultraconvergence of kth order FEM for (1.1). Unlike the classic extrapolation formulas defined in [15], our Richardson extrapolation formulas involve FE solutions on several levels of refined meshes generated by regular refinement processes (cf. section 1.1 for details). These local ultraconvergence results hold for not only tensor product elements, but also for simplicial elements. In particular, in section 2, we will consider the following auxiliary problem

$$\begin{cases} \mathcal{L}v = f & \text{in} \quad \widehat{\Omega}, \\ v = 0 & \text{on} \quad \partial\widehat{\Omega}, \end{cases} \tag{1.2}$$

where $\widehat{\Omega} \subset \mathbb{R}^n$ is a bounded sufficiently large domain with boundary $\partial\widehat{\Omega}$ such that $\Omega \subset\subset \widehat{\Omega}$. Here we assume that the extension of f in $\widehat{\Omega}$ is sufficiently smooth so that $v \in W^{2k+1,\infty}(\widehat{\Omega})$.

The rest of the article is organized as follows. In the rest of this section, notations and definitions are introduced. In section 2, the ultraconvergence of the Richardson extrapolation for problem (1.2) is investigated. In section 3, ultraconvergence properties of problem (1.1) are studied. Numerical experiments are presented in section 4.

1.1. Preliminary. In this paper, standard notations for Sobolev spaces and their norms are used. $\rho(\mathbf{x}, \partial\Omega)$ denotes the distance between $\mathbf{x}$ and $\partial\Omega$. $B(\mathbf{x}, r)$ is the open ball centered at $\mathbf{x}$ with radius r. $\overline{\nabla} v(\mathbf{x})$ is the arithmetic average gradient of v at $\mathbf{x}$. For a multi-index $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{N}_0^n$, denote $D^\beta v(\mathbf{x})$ as the partial derivative $\partial_{x_1}^{\beta_1} \cdots \partial_{x_n}^{\beta_n} v(\mathbf{x})$ of order $|\beta| = \beta_1 + \cdots + \beta_n$. Moreover, we denote $D^i v(\mathbf{x}) = \sum_{|\beta|=i} D^\beta v(\mathbf{x})$ for any natural number i, and $Dv(\mathbf{x})$ for $D^1 v(\mathbf{x})$ for simplicity.

We consider the usual variational form of (1.1) and (1.2). For a domain $K \subset \mathbb{R}^n$, define the bilinear form and the linear functional, respectively, as

$$\mathcal{B}_K(\psi,\phi) = \int_K A\nabla\psi\cdot\nabla\phi\, d\mathbf{x} \quad \text{and} \quad \mathcal{F}_K(\phi) = \int_K f\phi\, d\mathbf{x}.$$

Let $H_0^1(K) = \{v \in H^1(K) \ : \ v = 0 \text{ on the boundary of } K\}$. The weak problem of (1.2) is to find $v \in H_0^1(\widehat{\Omega})$ such that

$$\mathcal{B}_{\widehat{\Omega}}(v,\phi) = \mathcal{F}_{\widehat{\Omega}}(\phi) \quad \forall\phi \in H_0^1(\widehat{\Omega}).$$

The weak problem of (1.1) can be proposed analogously by replacing $\widehat{\Omega}$ with Ω.

For $\mathbf{z} \in K$, define the Green's function $G_{\mathbf{z}}^K \in W^{1,p}(K)$, $1 \le p < 2$, by

$$\mathcal{B}_K(G_{\mathbf{z}}^K,\psi) = \psi(\mathbf{z}) \quad \forall\psi \in W_0^{1,q}(K) \tag{1.3}$$

for some $q > 2$. In particular, we write $G_{\mathbf{z}}$ for $G_{\mathbf{z}}^\Omega$ for simplicity.

Let $\widehat{\mathcal{T}}_h$ be a certain conforming translation invariant mesh partition of $\widehat{\Omega}$ with mesh size h, which is aligned with Ω (i.e., the interior of any element in $\widehat{\mathcal{T}}_h$ does not intersect $\partial\Omega$). The initial mesh $\widehat{\mathcal{T}}_h$ shall satisfy the condition that nested meshes can be constructed by *regular refinement*, but not *bisection refinement* (cf., e.g., [5, 28]), so that every element of $\widehat{\mathcal{T}}_{h_{s-1}}$ is split into 2^n congruent elements in $\widehat{\mathcal{T}}_{h_s}$ for $1 \le s \le k$, and the refined mesh is similar to the initial mesh. Here, we denote $h_0 = h$, and thus $h_s = h_{s-1}/2$ for $1 \le s \le k$. Clearly, the regular rectangular and cubic meshes satisfy the aforementioned mesh conditions. For commonly used triangular mesh patterns shown in Figure 1, only the *regular pattern* and the *equilateral pattern* satisfy the conditions. As for tetrahedral meshes, only the pattern by Kuhn triangulation shown in Figure 2 satisfies these conditions (cf., e.g., [17]). For $0 \le s \le k$, let $\widehat{\mathcal{N}}_{h_s}$ be the set of all interior nodes of $\widehat{\mathcal{T}}_{h_s}$. Denote by $\mathcal{T}_{h_s}$ and $\mathcal{N}_{h_s}$ the restrictions of $\widehat{\mathcal{T}}_{h_s}$ and $\widehat{\mathcal{N}}_{h_s}$ to Ω, respectively, for $0 \le s \le k$.

We use standard Lagrange nodal FEMs. Define FE spaces

$$S_{h_s}(\widehat{\Omega}) = \{v \in C(\widehat{\Omega}) \ : \ v|_e \in \mathbb{P}_k \quad \forall e \in \widehat{\mathcal{T}}_{h_s}\}$$

for simplicial elements and FE spaces

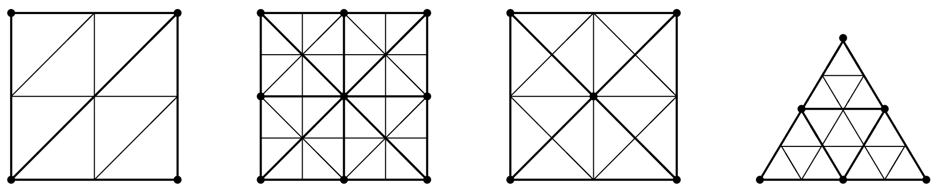
$$S_{h_s}(\widehat{\Omega}) = \{v \in C(\widehat{\Omega}) \ : \ v|_e \in \mathbb{Q}_k \quad \forall e \in \widehat{\mathcal{T}}_{h_s}\}$$

for tensor product elements, $0 \le s \le k$. Here $\mathbb{P}_k$ is the space of polynomials of total degree up to k, and $\mathbb{Q}_k$ is the space of polynomials of degree up to k in each variable. Clearly $S_{h_{s-1}}(\widehat{\Omega}) \subset S_{h_s}(\widehat{\Omega})$ for $1 \le s \le k$. Let $S_{h_s}^0(\widehat{\Omega}) = S_{h_s}(\widehat{\Omega}) \cap H_0^1(\widehat{\Omega})$. Define the FE projectors $R_{h_s,\widehat{\Omega}} : H_0^1(\widehat{\Omega}) \to S_{h_s}^0(\widehat{\Omega})$ by

$$\mathcal{B}_{\widehat{\Omega}}(\psi - R_{h_s,\widehat{\Omega}}\psi, \phi) = 0 \quad \forall\phi \in S_{h_s}^0(\widehat{\Omega})$$

for $0 \le s \le k$. Similarly, for an appropriate domain K (in particular when $K = \Omega$), FE spaces $S_{h_s}(K)$ and $S_{h_s}^0(K)$ and FE projectors $R_{h_s,K}$ can be defined for $0 \le s \le k$.

We assume that each interior mesh vertex $\mathbf{x}_0$ in $\widehat{\mathcal{T}}_{h_s}$ is a *local center of symmetry* of the mesh (see, e.g., [27]); i.e. for a sufficiently large r, if $\phi \in S_{h_s}(B(\mathbf{x}_0, r))$, then $\phi(\mathbf{x}_0 - (\mathbf{x} - \mathbf{x}_0)) \in S_{h_s}(B(\mathbf{x}_0, r))$. In the rest of the paper, we denote $I_{h_s}^k$ as the degree k



Regular Pattern Union Jack Pattern Crisscross Pattern Equilateral Pattern

FIG. 1. *Commonly used translation invariant triangular mesh patterns. Here the thick lines and dots are respectively edges and vertices of elements in coarse meshes, and the thin lines are only for refined meshes. A similar fine mesh in the regular and equilateral patterns is obtained by regular refinement. A similar fine mesh in the Union Jack and Crisscross patterns is obtained by bisection refinement.*

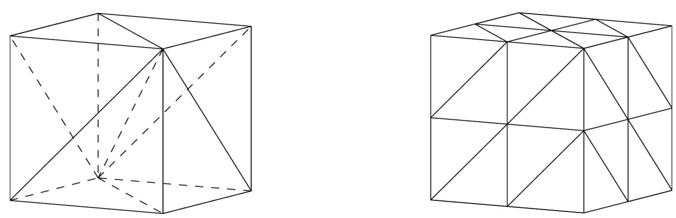


FIG. 2. *Translation invariant tetrahedral mesh pattern. Left: the* Kuhn triangulation *of a cube into six congruent tetrahedra. Right: a uniform tetrahedral mesh on Kuhn triangulation.*

Lagrange interpolation operator over $\widehat{\mathcal{T}}_{h_s}$. c, with or without a subscript, is a constant independent of u and h, which is not necessarily the same at each occurrence. Define $\tilde{k} = 1$ if $k = 1$, and $\tilde{k} = 0$ if $k > 1$. We sometimes write h for h_0 in the subindexes of all aforementioned notations for simplicity.

2. Ultraconvergence of Richardson extrapolation for the problem (1.2). We will propose two novel Richardson extrapolation formulas for the kth order FEM in section 2.1. Local ultraconvergence of the numerical solutions and gradients for problem (1.2) will be investigated by extrapolation in sections 2.2 and 2.3, respectively.

2.1. Two Richardson extrapolation operators. First, define

$$I_1(k) = \begin{cases} k+2 & \text{if } k \text{ is even}, \\ k+1 & \text{if } k \text{ is odd}, \end{cases} \qquad J_1(k) = \begin{cases} \frac{k-2}{2} & \text{if } k \text{ is even}, \\ \frac{k-1}{2} & \text{if } k \text{ is odd}, \end{cases} \tag{2.1}$$

and $F_m = (1, 0, \ldots, 0)^T_{m\times 1}$ for $m \geq 1$. Then let

$$A_1 = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2^{-I_1(k)} & \cdots & 2^{-J_1(k)I_1(k)} \\ 1 & 2^{-(I_1(k)+2)} & \cdots & 2^{-J_1(k)(I_1(k)+2)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 2^{-(I_1(k)+2J_1(k)-2)} & \cdots & 2^{-J_1(k)(I_1(k)+2J_1(k)-2)} \end{pmatrix}_{(J_1(k)+1)\times(J_1(k)+1)}.$$

The first extrapolation operator $\widehat{P}_1^k : S_h^0(\widehat{\Omega}) \to S_{h_k}^0(\widehat{\Omega})$ is defined by

$$\widehat{P}_1^k(R_{h,\widehat{\Omega}}\psi)(\mathbf{x}) = \sum_{s=0}^{J_1(k)} \alpha_s^{(1)} R_{h_s,\widehat{\Omega}}\psi(\mathbf{x}), \tag{2.2}$$

where $\alpha^{(1)} = (\alpha_0^{(1)}, \dots, \alpha_{J_1(k)}^{(1)})^T$ solves the problem

$$A_1\alpha^{(1)} = F_{J_1(k)+1}. \tag{2.3}$$

Similarly, define

$$I_2(k) = \begin{cases} k & \text{if } k \text{ is even,} \\ k+1 & \text{if } k \text{ is odd,} \end{cases} \qquad J_2(k) = \begin{cases} \frac{k}{2} & \text{if } k \text{ is even,} \\ \frac{k-1}{2} & \text{if } k \text{ is odd,} \end{cases}$$

and

$$A_2 = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2^{-I_2(k)} & \cdots & 2^{-J_2(k)I_2(k)} \\ 1 & 2^{-(I_2(k)+2)} & \cdots & 2^{-J_2(k)(I_2(k)+2)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 2^{-(I_2(k)+2J_2(k)-2)} & \cdots & 2^{-J_2(k)(I_2(k)+2J_2(k)-2)} \end{pmatrix}_{(J_2(k)+1)\times(J_2(k)+1)}.$$

The second extrapolation operator $\widehat{P}_2^k$ is defined by

$$\widehat{P}_2^k(D(R_{h,\widehat{\Omega}}\psi))(\mathbf{x}) = \sum_{s=0}^{J_2(k)} \alpha_s^{(2)} D(R_{h_s,\widehat{\Omega}}\psi)(\mathbf{x}), \tag{2.4}$$

where $\alpha^{(2)} = (\alpha_0^{(2)}, \dots, \alpha_{J_2(k)}^{(2)})$ solves the problem

$$A_2\alpha^{(2)} = F_{J_2(k)+1}. \tag{2.5}$$

Remark 2.1. Unlike the postprocessing approaches in [4, 8, 11, 25], the extrapolation operators developed above are independent of the dimension or shape of the element, provided regularly refined translation invariant meshes are used.

2.2. Local ultraconvergence of FE solutions for problem (1.2). The main local ultraconvergence result is summarized in Theorem 2.1.

THEOREM 2.1. *Let υ be the solution to* (1.2) *and $\mathbf{x}_0 \in \widehat{\mathcal{N}}_h$. Assume that there exists a constant r such that $B(\mathbf{x}_0, 2^{J_1(k)}r) \subset \widehat{\Omega}$. Assume also that $\upsilon \in W_0^{2k,\infty}(\widehat{\Omega})$ and $\upsilon(\mathbf{x}) = 0$ for all $\mathbf{x} \in \widehat{\Omega} \setminus B(\mathbf{x}_0, r)$. Then*

$$|(\widehat{P}_1^k(R_{h,\widehat{\Omega}}\upsilon) - \upsilon)(\mathbf{x}_0)| \le ch^{2k}|\ln h|^{\tilde{k}+1}\|\upsilon\|_{W^{2k,\infty}(\widehat{\Omega})}. \tag{2.6}$$

In the process of proving Theorem 2.1, we will use an auxiliary function. Let υ be the solution to (1.2). For $0 \le q \le J_1(k)$, define

$$\upsilon_q(\mathbf{x}) = \begin{cases} \upsilon\left(\frac{\mathbf{x}-\mathbf{x}_0}{2^q} + \mathbf{x}_0\right) & \text{if } \mathbf{x} \in B(\mathbf{x}_0, 2^q r), \\ 0 & \text{if } \mathbf{x} \in \widehat{\Omega} \setminus B(\mathbf{x}_0, 2^q r), \end{cases} \tag{2.7}$$

where $\mathbf{x}_0 \in \widehat{\mathcal{N}}_h$ and r is the constant satisfying the hypotheses of Theorem 2.1. We need the following lemma.

LEMMA 2.2. *Let υ be the solution to* (1.2) *and $\mathbf{x}_0 \in \widehat{\mathcal{N}}_h$, so that the hypotheses of Theorem* 2.1 *hold. Set*

$$\omega(\mathbf{x}) = \sum_{q=0}^{J_1(k)} \alpha_q^{(1)} \overline{\upsilon}_q(\mathbf{x}), \tag{2.8}$$

where

$$\overline{\upsilon}_q(\mathbf{x}) = \frac{1}{2}\big(\upsilon_q(\mathbf{x}) + \upsilon_q(\mathbf{x}_0 - (\mathbf{x} - \mathbf{x}_0))\big), \tag{2.9}$$

and υ_q *is defined in* (2.7). *Then there holds*

$$|(R_{h,\widehat{\Omega}}\omega - \omega)(\mathbf{x}_0)| \le ch^{2k}|\ln h|^{\tilde{k}+1}\|\upsilon\|_{W^{2k,\infty}(\widehat{\Omega})}. \tag{2.10}$$

Proof. We first estimate $D^{k+1}\omega(\mathbf{x})$ for $\mathbf{x} \in B(\mathbf{x}_0, r) \subset\subset \widehat{\Omega}$ for some sufficiently small r. Note that (2.7) and (2.9) imply that, for any positive even integer $i \le 2k$ and $0 \le q \le J_1(k)$,

$$D^i\overline{\upsilon}_q(\mathbf{x}_0) = 2^{-qi}D^i\upsilon(\mathbf{x}_0). \tag{2.11}$$

Hence, by (2.7), (2.8), (2.9), and (2.3), it follows that, for all $0 \le s \le J_1(k) - 1$,

$$\begin{aligned} D^{I_1(k)+2s}\omega(\mathbf{x}_0) &= \sum_{q=0}^{J_1(k)} \alpha_q^{(1)} D^{I_1(k)+2s}\overline{\upsilon}_q(\mathbf{x}_0) \\ &= \sum_{q=0}^{J_1(k)} \alpha_q^{(1)} 2^{-q(I_1(k)+2s)} D^{I_1(k)+2s}\upsilon(\mathbf{x}_0) = 0. \end{aligned}$$

We conclude that, if $I_1(k) \le i \le I_1(k) + 2(J_1(k) - 1) = 2k - 2$ is even, then

$$D^i\omega(\mathbf{x}_0) = 0. \tag{2.12}$$

On the other hand, for i odd and $0 \le q \le J_1(k)$, it follows from (2.9) that

$$D^i\overline{\upsilon}_q(\mathbf{x}_0) = \frac{1}{2}(2^{-qi}D^i\upsilon(\mathbf{x}_0) - 2^{-qi}D^i\upsilon(\mathbf{x}_0)) = 0,$$

and hence

$$D^i\omega(\mathbf{x}_0) = 0 \tag{2.13}$$

by (2.8). Using (2.12) and (2.13), we arrive at $D^{k+l}\omega(\mathbf{x}_0) = 0$ for $1 \le l \le k - 1$. Therefore, from Taylor's expansions of the $(k+1)$st order derivatives of ω at $\mathbf{x}_0$, one obtains

$$|D^{k+1}\omega(\mathbf{x})| \le \sum_{|\beta|=k+1} c_\beta|D^{2k}\omega(\mathbf{y}_\beta)||\mathbf{x} - \mathbf{x}_0|^{k-1} \le c\|\omega\|_{W^{2k,\infty}(\widehat{\Omega})}|\mathbf{x} - \mathbf{x}_0|^{k-1}, \tag{2.14}$$

where $\mathbf{y}_\beta \in B(\mathbf{x}_0, r)$.

Next, we prove (2.10). Let $R_{h,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}$ be the discrete Green's function for problem (1.2) at $\mathbf{x}_0$ (cf. the definition (1.3) for $G_{\mathbf{x}_0}^{\widehat{\Omega}}$). Then it holds that

$$\|G_{\mathbf{x}_0}^{\widehat{\Omega}} - R_{h,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}\|_{W^{1,1}(\widehat{\Omega})} \le ch|\ln h|^{\tilde{k}}, \tag{2.15}$$

and for any $r > 0$,

$$\|G_{\mathbf{x}_0}^{\widehat{\Omega}} - R_{h,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}\|_{W^{1,\infty}(\widehat{\Omega}\setminus B(\mathbf{x}_0,r))} \le ch^k r^{1-k-n}|\ln h|^{\tilde{k}}, \tag{2.16}$$

which are obtained from [24, Theorem 2] and [21, Theorem 6.1], respectively.

Since υ has a compact support, by definition (2.8), there exists a constant $d_0 > 0$ such that $\omega(\mathbf{x}) = 0$ for all $\mathbf{x} \in \widehat{\Omega} \setminus B(\mathbf{x}_0, d_0)$. Using (2.14)–(2.16), the triangle inequality, Hölder's inequality, and standard approximation theory, we have

$$\begin{aligned}
&|(R_{h,\widehat{\Omega}}\omega - \omega)(\mathbf{x}_0)| = |\mathcal{B}_{B(\mathbf{x}_0,d_0)}(\omega - I_h^k\omega, G_{\mathbf{x}_0}^{\widehat{\Omega}} - R_{h,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}})| \\
\le& |\mathcal{B}_{B(\mathbf{x}_0,h)}(\omega - I_h^k\omega, G_{\mathbf{x}_0}^{\widehat{\Omega}} - R_{h,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}})| \\
&+ |\mathcal{B}_{B(\mathbf{x}_0,d_0)\setminus B(\mathbf{x}_0,h)}(\omega - I_h^k\omega, G_{\mathbf{x}_0}^{\widehat{\Omega}} - R_{h,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}})| \\
\le& ch^{2k-1}\|\omega\|_{W^{2k,\infty}(\widehat{\Omega})}\|G_{\mathbf{x}_0}^{\widehat{\Omega}} - R_{h,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}\|_{W^{1,1}(\widehat{\Omega})} \\
&+ c\int_{B(\mathbf{x}_0,d_0)\setminus B(\mathbf{x}_0,h)} ch^k|D^{k+1}\omega(\mathbf{x})|\,|D(G_{\mathbf{x}_0}^{\widehat{\Omega}} - R_{h,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}})(\mathbf{x})|d\mathbf{x} \\
\le& ch^{2k-1}\|\omega\|_{W^{2k,\infty}(\widehat{\Omega})}\|G_{\mathbf{x}_0}^{\widehat{\Omega}} - R_{h,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}\|_{W^{1,1}(\widehat{\Omega})} \\
&+ c\|\omega\|_{W^{2k,\infty}(\widehat{\Omega})}\int_{B(\mathbf{x}_0,d_0)\setminus B(\mathbf{x}_0,h)} h^k|\mathbf{x}-\mathbf{x}_0|^{k-1}h^k|\mathbf{x}-\mathbf{x}_0|^{1-k-n}|\ln h|^{\tilde{k}}d\mathbf{x} \\
\le& ch^{2k}|\ln h|^{\tilde{k}}\,\|\omega\|_{W^{2k,\infty}(\widehat{\Omega})} + ch^{2k}|\ln h|^{\tilde{k}+1}\|\omega\|_{W^{2k,\infty}(\widehat{\Omega})} \\
\le& ch^{2k}|\ln h|^{\tilde{k}+1}\|\upsilon\|_{W^{2k,\infty}(\widehat{\Omega})}.
\end{aligned}$$

Here we assume that h is small. This ends the proof. ∎

We are now in a position to prove Theorem 2.1.

Proof of Theorem 2.1. Note that $\psi(\mathbf{x}) \in S_h(\widehat{\Omega})$ implies $\psi(\frac{\mathbf{x}-\mathbf{x}_0}{2^q} + \mathbf{x}_0) \in S_{h_q}(\widehat{\Omega})$ for $0 \le q \le k$. Therefore, by definition (2.7), we have, for all $0 \le q \le J_1(k)$,

$$R_{h,\widehat{\Omega}}\upsilon_q(\mathbf{x}_0) = R_{h_q,\widehat{\Omega}}\upsilon(\mathbf{x}_0). \tag{2.17}$$

Let $\mathbf{y} = \mathbf{x}_0 - (\mathbf{x} - \mathbf{x}_0) = 2\mathbf{x}_0 - \mathbf{x}$. Set $\varpi_q(\mathbf{x}) = \upsilon_q(\mathbf{y})$ for $0 \le q \le J_1(k)$, then

$$\begin{aligned}
R_{h,\widehat{\Omega}}\varpi_q(\mathbf{x}_0) &= \sum_{i,j=1}^n \int_{\widehat{\Omega}} a_{ij}\frac{\partial R_{h,\widehat{\Omega}}\varpi_q(\mathbf{x})}{\partial x_i}\frac{\partial R_{h,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}(\mathbf{x})}{\partial x_j}d\mathbf{x} \\
&= -\sum_{i,j=1}^n \int_{\widehat{\Omega}} a_{ij}\frac{\partial R_{h,\widehat{\Omega}}\upsilon_q(\mathbf{y})}{\partial x_i}\frac{\partial R_{h,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}(\mathbf{x})}{\partial x_j}d\mathbf{x} \\
&= \sum_{i,j=1}^n \int_{\widehat{\Omega}} a_{ij}\frac{\partial R_{h,\widehat{\Omega}}\upsilon_q(\mathbf{y})}{\partial x_i}\frac{\partial R_{h,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}(\mathbf{y})}{\partial x_j}d\mathbf{x} \\
&= \sum_{i,j=1}^n \int_{\widehat{\Omega}} a_{ij}\frac{\partial R_{h,\widehat{\Omega}}\upsilon_q(\mathbf{y})}{\partial x_i}\frac{\partial R_{h,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}(\mathbf{y})}{\partial x_j}d\mathbf{y} = R_{h,\widehat{\Omega}}\upsilon_q(\mathbf{x}_0) = R_{h_q,\widehat{\Omega}}\upsilon(\mathbf{x}_0).
\end{aligned}$$

Using this together with (2.9) and (2.17), we have

$$R_{h,\widehat{\Omega}}\overline{\upsilon}_q(\mathbf{x}_0) = R_{h_q,\widehat{\Omega}}\upsilon(\mathbf{x}_0). \tag{2.18}$$

It follows that

$$(R_{h,\widehat{\Omega}}\overline{\upsilon}_q - \overline{\upsilon}_q)(\mathbf{x}_0) = (R_{h_q,\widehat{\Omega}}\upsilon - \upsilon)(\mathbf{x}_0).$$

By (2.3) and (2.8), we arrive at

$$(\widehat{P}_1^k(R_{h,\widehat{\Omega}}\upsilon)-\upsilon)(\mathbf{x}_0)=\sum_{q=0}^{J_1(k)}\alpha_q^{(1)}(R_{h_q,\widehat{\Omega}}\upsilon-\upsilon)(\mathbf{x}_0)$$

$$=\sum_{q=0}^{J_1(k)}\alpha_q^{(1)}(R_{h,\widehat{\Omega}}\overline{\upsilon}_q-\overline{\upsilon}_q)(\mathbf{x}_0)=(R_{h,\widehat{\Omega}}\omega-\omega)(\mathbf{x}_0). \tag{2.19}$$

Then (2.6) follows from (2.10) and (2.19). ∎

2.3. Local ultraconvergence of FE gradients for problem (1.2). In this section, we consider local ultraconvergence of the derivatives of FE solutions.

THEOREM 2.3. *Let υ be the solution to* (1.2) *and $\mathbf{x}_0\in\widehat{\mathcal{N}}_h$. Assume that there exists a constant r such that $B(\mathbf{x}_0,2^{J_2(k)}r)\subset\widehat{\Omega}$. Assume also that $\upsilon\in W_0^{2k+1,\infty}(\widehat{\Omega})$ and $\upsilon(\mathbf{x})=0$ for all $\mathbf{x}\in\widehat{\Omega}\setminus B(\mathbf{x}_0,r)$. Then*

$$|(\widehat{P}_2^k(\overline{\nabla}R_{h,\widehat{\Omega}}\upsilon)-\nabla\upsilon)(\mathbf{x}_0)|\le ch^{2k}|\ln h|^{\tilde{k}+1}\|\upsilon\|_{W^{2k+1,\infty}(\widehat{\Omega})}. \tag{2.20}$$

Before proving Theorem 2.3, we introduce the following lemma.

LEMMA 2.4. *Let υ be the solution to* (1.2) *and $\mathbf{x}_0\in\widehat{\mathcal{N}}_h$, so that the hypotheses of Theorem* 2.3 *hold true. Let $\upsilon_q(\mathbf{x})$ be defined in* (2.7) *for $0\le q\le J_2(k)$. Set*

$$\gamma(\mathbf{x})=\sum_{q=0}^{J_2(k)}\alpha_q^{(2)}\hat{\upsilon}_q(\mathbf{x}), \tag{2.21}$$

where

$$\hat{\upsilon}_q(\mathbf{x})=2^{q-1}\big(\upsilon_q(\mathbf{x})-\upsilon_q(\mathbf{x}_0-(\mathbf{x}-\mathbf{x}_0))\big). \tag{2.22}$$

Then there holds

$$|D(R_{h,\widehat{\Omega}}\gamma-\gamma)(\mathbf{x}_0)|\le ch^{2k}|\ln h|^{\tilde{k}+1}\|\upsilon\|_{W^{2k+1,\infty}(\widehat{\Omega})}. \tag{2.23}$$

Proof. We first estimate $D^{k+1}\gamma(\mathbf{x})$ for $\mathbf{x}\in B(\mathbf{x}_0,r)\subset\subset\widehat{\Omega}$ for some sufficiently small r, which is similar to the counterpart in the proof of Lemma 2.4. Combining (2.4), (2.5), (2.7), (2.21), and (2.22) gives, for all $0\le s\le J_2(k)-1$,

$$D^{I_2(k)+2s+1}\gamma(\mathbf{x}_0)=D^{I_2(k)+2s}D\gamma(\mathbf{x}_0)=\sum_{q=0}^{J_2(k)}\alpha_q^{(2)}D^{I_2(k)+2s}D\hat{\upsilon}_q(\mathbf{x}_0)$$

$$=\sum_{q=0}^{J_2(k)}\alpha_q^{(2)}D^{I_2(k)+2s}2^qD\upsilon_q(\mathbf{x}_0)=\sum_{q=0}^{J_2(k)}\alpha_q^{(2)}2^{-q(I_2(k)+2s)}D^{I_2(k)+2s}D\upsilon(\mathbf{x}_0)$$

$$=\sum_{q=0}^{J_2(k)}\alpha_q^{(2)}2^{-q(I_2(k)+2s)}D^{I_2(k)+2s+1}\upsilon(\mathbf{x}_0)=0. \tag{2.24}$$

By (2.24), we obtain that, if $I_2(k)\le i\le I_2(k)+2(J_2(k)-1)+1=2k-1$ is odd, then

$$D^i\gamma(\mathbf{x}_0)=0.$$

On the other hand, it follows from (2.22) that, if i is even, then

$$D^i\gamma(\mathbf{x}_0) = 0.$$

Therefore, for all $1 \le l \le k$,

$$D^{k+l}\gamma(\mathbf{x}_0) = 0. \tag{2.25}$$

Similarly to (2.14), by Taylor's theorem and (2.25), we obtain

$$|D^{k+1}\gamma(\mathbf{x})| \le c\|\gamma\|_{W^{2k+1,\infty}(\widehat{\Omega})}|\mathbf{x}-\mathbf{x}_0|^k. \tag{2.26}$$

Next, we prove (2.23). Without loss of generality, assume that $\gamma(\mathbf{x}) = 0$ for all $\mathbf{x} \in \widehat{\Omega} \setminus B(\mathbf{x}_0, d_0)$ for some constant $d_0 > 0$. Let $\mu \in \mathbb{P}_k$ satisfying

$$D^l(\mu-\gamma)(\mathbf{x}_0) = 0, \quad 0 \le l \le k.$$

Let χ be the cutoff function such that $0 \le \chi \le 1$, $\chi = 1$ in $B(\mathbf{x}_0, d_0/2)$, $\chi = 0$ in $\widehat{\Omega} \setminus B(\mathbf{x}_0, d_0)$, and $\|\chi\|_{W^{p,\infty}(\widehat{\Omega})} \le c$ for all $p \le 2k+1$. Decompose γ into

$$\gamma(\mathbf{x}) = \hat{\gamma}(\mathbf{x}) + \overline{\gamma}(\mathbf{x}),$$

where $\hat{\gamma}(\mathbf{x}) = \chi(\mathbf{x})\mu(\mathbf{x})$. By the triangle inequality,

$$|D(R_{h,\widehat{\Omega}}\gamma - \gamma)(\mathbf{x}_0)| \le |D(R_{h,\widehat{\Omega}}\hat{\gamma} - \hat{\gamma})(\mathbf{x}_0)| + |D(R_{h,\widehat{\Omega}}\overline{\gamma} - \overline{\gamma})(\mathbf{x}_0)|. \tag{2.27}$$

First we investigate $D(R_{h,\widehat{\Omega}}\hat{\gamma} - \hat{\gamma})(\mathbf{x}_0)$. Note that for any $\mathbf{x} \in B(\mathbf{x}_0, d_0/2)$, $\hat{\gamma}(\mathbf{x}) = \mu(\mathbf{x}) = I_h^k\mu(\mathbf{x}) = I_h^k\hat{\gamma}(\mathbf{x})$. It follows that, for all $\mathbf{x} \in B(\mathbf{x}_0, d_0/4)$, by (2.16)

$$\begin{aligned}
|(R_{h,\widehat{\Omega}}\hat{\gamma} - \hat{\gamma})(\mathbf{x})| &= |\mathcal{B}_{B(\mathbf{x}_0,d_0)}(I_h^k\hat{\gamma} - \hat{\gamma}, G_{\mathbf{x}}^{\widehat{\Omega}} - R_{h,\widehat{\Omega}}G_{\mathbf{x}}^{\widehat{\Omega}})| \\
&\le c\|\hat{\gamma} - I_h^k\hat{\gamma}\|_{W^{1,\infty}(\widehat{\Omega}\setminus B(\mathbf{x}_0,d_0/2))}\|G_{\mathbf{x}}^{\widehat{\Omega}} - R_{h,\widehat{\Omega}}G_{\mathbf{x}}^{\widehat{\Omega}}\|_{W^{1,1}(\widehat{\Omega}\setminus B(\mathbf{x}_0,d_0/2))} \\
&\le c\|\hat{\gamma} - I_h^k\hat{\gamma}\|_{W^{1,\infty}(\widehat{\Omega}\setminus B(\mathbf{x}_0,d_0/2))}\|G_{\mathbf{x}}^{\widehat{\Omega}} - R_{h,\widehat{\Omega}}G_{\mathbf{x}}^{\widehat{\Omega}}\|_{W^{1,1}(\widehat{\Omega}\setminus B(\mathbf{x},d_0/4))} \\
&\le ch^k\|\hat{\gamma}\|_{W^{k+1,\infty}(\widehat{\Omega})}h^k d_0^{1-k}|\ln h|^{\tilde{k}} \le ch^{2k}|\ln h|^{\tilde{k}}\ \|v\|_{W^{k+1,\infty}(\widehat{\Omega})},
\end{aligned} \tag{2.28}$$

where $G_{\mathbf{x}}^{\widehat{\Omega}}(\mathbf{y})$ and $R_{h,\widehat{\Omega}}G_{\mathbf{x}}^{\widehat{\Omega}}(\mathbf{y})$ are the Green's function and the discrete Green's function for problem (1.2) at $\mathbf{x}$, respectively. Applying the interior maximum norm estimates in [27] and (2.28), we obtain

$$\begin{aligned}
|D(R_{h,\widehat{\Omega}}\hat{\gamma} - \hat{\gamma})(\mathbf{x}_0)| &\le c\|\hat{\gamma} - I_h^k\hat{\gamma}\|_{W^{1,\infty}(B(\mathbf{x}_0,d_0/4))} + cd_0^{-1}\|R_{h,\widehat{\Omega}}\hat{\gamma} - \hat{\gamma}\|_{L^\infty(B(\mathbf{x}_0,d_0/4))} \\
&\le c\|R_{h,\widehat{\Omega}}\hat{\gamma} - \hat{\gamma}\|_{L^\infty(B(\mathbf{x}_0,d_0/4))} \\
&\le ch^{2k}|\ln h|^{\tilde{k}}\|v\|_{W^{2k+1,\infty}(\widehat{\Omega})}.
\end{aligned} \tag{2.29}$$

Next we turn to the estimation of $D(R_{h,\widehat{\Omega}}\overline{\gamma} - \overline{\gamma})(\mathbf{x}_0)$. Let T be the positive integer satisfying $2^{T-1}h < 2d_0 \le 2^T h$. Let $r_{-1} = 0$ and $r_j = 2^j h$ for $0 \le j \le T$. For $1 \le j \le T$, let θ_j be the cutoff function such that $0 \le \theta_j \le 1$, $\theta_j = 1$ in $B(\mathbf{x}_0, r_{j-1}) \setminus B(\mathbf{x}_0, r_{j-2})$, $\theta_j = 0$ in $(\widehat{\Omega} \setminus B(\mathbf{x}_0, r_j)) \cup B(\mathbf{x}_0, r_{j-3})$, and $\|\theta_j\|_{W^{p,\infty}(\widehat{\Omega})} \le cr_j^{-p}$ for all positive integers p. Set

$$\overline{\theta}_j(\mathbf{x}) = \frac{\theta_j(\mathbf{x})}{\sum\limits_{i=1}^{T}\theta_i(\mathbf{x})} \quad \forall \mathbf{x} \in B(\mathbf{x}_0, d_0).$$

Then split $\overline{\gamma}(\mathbf{x})$ into

$$\overline{\gamma}(\mathbf{x}) = \sum_{j=1}^{T} \overline{\gamma}_j(\mathbf{x}),$$

where $\overline{\gamma}_j(\mathbf{x}) = \overline{\theta}_j(\mathbf{x})\overline{\gamma}(\mathbf{x})$. Hence $D(R_{h,\widehat{\Omega}}\overline{\gamma} - \overline{\gamma})(\mathbf{x}_0)$ can be decomposed into

$$D(R_{h,\widehat{\Omega}}\overline{\gamma} - \overline{\gamma})(\mathbf{x}_0) = \sum_{j=1}^{T} D(R_{h,\widehat{\Omega}}\overline{\gamma}_j - \overline{\gamma}_j)(\mathbf{x}_0). \tag{2.30}$$

We shall estimate $D(R_{h,\widehat{\Omega}}\overline{\gamma}_j - \overline{\gamma}_j)(\mathbf{x}_0)$ for each $1 \le j \le T$. First consider the case that $1 \le j \le 3$. Let $\widehat{\mathcal{T}}_h^* = \{\tau^* \in \widehat{\mathcal{T}}_h \ : \ \tau^* \subset B(\mathbf{x}_0, r_4)\}$. By inverse estimate and (2.26), we have, for all $\tau^* \in \widehat{\mathcal{T}}_h^*$,

$$\begin{aligned} \|\overline{\gamma}_j\|_{W^{k+1,\infty}(\tau^*)} &\le c\sum_{i=1}^{j} h^{i-k-1}\|\overline{\gamma}\|_{W^{i,\infty}(\tau^*)} \le c\sum_{i=1}^{j} h^{i-k-1}\|\gamma - \mu\|_{W^{i,\infty}(\tau^*)} \\ &\le c\sum_{i=1}^{j} h^{i-k-1}h^{k+1-i}\|\gamma\|_{W^{k+1,\infty}(\tau^*)} \le ch^k\|\upsilon\|_{W^{2k+1,\infty}(\widehat{\Omega})}. \end{aligned}$$

Therefore,

$$\begin{aligned} \|\overline{\gamma}_j - R_{h,\widehat{\Omega}}\overline{\gamma}_j\|_{W^{1,\infty}(\widehat{\Omega})} &\le c\|\overline{\gamma}_j - I_h^k\overline{\gamma}_j\|_{W^{1,\infty}(\widehat{\Omega})} = c\|\overline{\gamma}_j - I_h^k\overline{\gamma}_j\|_{W^{1,\infty}(B(\mathbf{x}_0,r_4))} \\ &\le ch^k \max_{\tau^*\in\widehat{\mathcal{T}}_h^*} \|\overline{\gamma}_j\|_{W^{k+1,\infty}(\tau^*)} \le ch^{2k}\|\upsilon\|_{W^{2k+1,\infty}(\widehat{\Omega})}. \end{aligned} \tag{2.31}$$

Then consider the case for $4 \le j \le T$. Note that, for any $\mathbf{x} \in (\widehat{\Omega} \setminus B(\mathbf{x}_0, r_j)) \cup B(\mathbf{x}_0, r_{j-3})$, $\overline{\gamma}_j(\mathbf{x}) = 0$. One observes that, if $\mathbf{x} \in B(\mathbf{x}_0, r_{j-4})$, then by (2.16),

$$\begin{aligned} |(\overline{\gamma}_j - R_{h,\widehat{\Omega}}\overline{\gamma}_j)(\mathbf{x})| &= |\mathcal{B}_{\widehat{\Omega}}(\overline{\gamma}_j - I_h^k\overline{\gamma}_j, G_{\mathbf{x}}^{\widehat{\Omega}} - R_{h,\widehat{\Omega}}G_{\mathbf{x}}^{\widehat{\Omega}})| \\ &\le c\|\overline{\gamma}_j - I_h^k\overline{\gamma}_j\|_{W^{1,\infty}(B(\mathbf{x}_0,r_j))}\|G_{\mathbf{x}}^{\widehat{\Omega}} - R_{h,\widehat{\Omega}}G_{\mathbf{x}}^{\widehat{\Omega}}\|_{W^{1,1}(\widehat{\Omega}\setminus B(\mathbf{x}_0,r_{j-3}))} \\ &\le ch^k\|\overline{\gamma}_j\|_{W^{k+1,\infty}(B(\mathbf{x}_0,r_j))}h^k r_j^{1-k}|\ln h|^{\tilde{k}} \\ &\le ch^k r_j^k\|\overline{\gamma}_j\|_{W^{2k+1,\infty}(B(\mathbf{x}_0,r_j))}h^k r_j^{1-k}|\ln h|^{\tilde{k}} \\ &\le ch^{2k} r_j|\ln h|^{\tilde{k}}\ \|\upsilon\|_{W^{2k+1,\infty}(\widehat{\Omega})}. \end{aligned}$$

Subsequently, using the interior maximum norm estimates for FEM, we have

$$\begin{aligned} |D(\overline{\gamma}_j - R_{h,\widehat{\Omega}}\overline{\gamma}_j)(\mathbf{x}_0)| \le& cr_j^{-1}\|R_{h,\widehat{\Omega}}\overline{\gamma} - \overline{\gamma}\|_{L^\infty(B(\mathbf{x}_0,r_{j-4}))} \\ \le& ch^{2k}|\ln h|^{\tilde{k}}\|\upsilon\|_{W^{2k+1,\infty}(\widehat{\Omega})}. \end{aligned} \tag{2.32}$$

By estimates (2.31) and (2.32), we have from (2.30) that

$$|D(\overline{\gamma} - R_{h,\widehat{\Omega}}\overline{\gamma})(\mathbf{x}_0)| \le ch^{2k}|\ln h|^{\tilde{k}+1}\|\upsilon\|_{W^{2k+1,\infty}(\widehat{\Omega})}. \tag{2.33}$$

Then (2.23) follows from (2.27), (2.29), and (2.33). ∎

We next use the result of Lemma 2.4 to show Theorem 2.3.

Proof of Theorem 2.3. By (2.7), we have, for any $\mathbf{y} \in \widehat{\Omega}$ and $0 \le q \le J_2(k)$,

$$R_{h_q,\widehat{\Omega}}\upsilon(\mathbf{x}_0 + \mathbf{y}) = R_{h,\widehat{\Omega}}\upsilon_q(\mathbf{x}_0 + 2^q\mathbf{y}).$$

Hence

$$DR_{h_q,\widehat{\Omega}}\upsilon(\mathbf{x}_0) = 2^q DR_{h,\widehat{\Omega}}\upsilon_q(\mathbf{x}_0).$$

By the same arguments in the proof for (2.18), we have

$$\overline{\nabla} R_{h,\widehat{\Omega}}\hat{\upsilon}_q(\mathbf{x}_0) = 2^q \overline{\nabla} R_{h,\widehat{\Omega}}\upsilon_q(\mathbf{x}_0).$$

From the above two identities, we have

$$\overline{\nabla} R_{h,\widehat{\Omega}}\hat{\upsilon}_q(\mathbf{x}_0) = \overline{\nabla} R_{h_q,\widehat{\Omega}}\upsilon(\mathbf{x}_0).$$

Similarly, we have

$$\overline{\nabla}\hat{\upsilon}_q(\mathbf{x}_0) = \overline{\nabla}\upsilon(\mathbf{x}_0).$$

Therefore

$$\overline{\nabla}(\hat{\upsilon}_q - R_{h,\widehat{\Omega}}\hat{\upsilon}_q)(\mathbf{x}_0) = \overline{\nabla}(\upsilon - R_{h_q,\widehat{\Omega}}\upsilon)(\mathbf{x}_0).$$

Furthermore, we derive

$$\begin{aligned}\overline{\nabla}(\upsilon - \widehat{P}_2^k(R_{h,\widehat{\Omega}}\upsilon))(\mathbf{x}_0) &= \sum_{q=0}^{J_2(k)} \alpha_q^{(2)} \overline{\nabla}(\upsilon - R_{h_q,\widehat{\Omega}}\upsilon)(\mathbf{x}_0) \\ &= \sum_{q=0}^{J_2(k)} \alpha_q^{(2)} \overline{\nabla}(\hat{\upsilon}_q - R_{h,\widehat{\Omega}}\hat{\upsilon}_q)(\mathbf{x}_0) = \overline{\nabla}(\gamma - R_{h,\widehat{\Omega}}\gamma)(\mathbf{x}_0), \end{aligned} \tag{2.34}$$

Then (2.20) follows by combining (2.23) and (2.34). ∎

3. Ultraconvergence of Richardson extrapolation of FEM for the problem (1.1). In this section, we consider local ultraconvergence of FE solutions to the problem (1.1). Let $\mathcal{T}_h$ be a quasi-uniform conforming partition of Ω, such that there exists a parallelogram or parallelepiped $\tau \subset \Omega$ which satisfies the following properties: (1) τ is the union of some elements in $\mathcal{T}_h$; (2) τ has a size diam $(\tau) \simeq 1$; and (3) $\bigcup_{\tau' \in \mathcal{T}_h, \tau' \subset \tau} \tau' = \bigcup_{\tau' \in \widehat{\mathcal{T}}_h, \tau' \subset \tau} \tau'$. Recall the Richardson extrapolation operators defined in section 2.1. We define analog extrapolation operators P_1^k and P_2^k by

$$P_1^k(R_{h,\Omega}\psi)(\mathbf{x}) = \sum_{s=0}^{J_1(k)} \alpha_s^{(1)} R_{h_s,\Omega}\psi(\mathbf{x}),$$

$$P_2^k(D(R_{h,\Omega}\psi))(\mathbf{x}) = \sum_{s=0}^{J_2(k)} \alpha_s^{(2)} D(R_{h_s,\Omega}\psi)(\mathbf{x}).$$

Similarly to Theorems 2.1 and 2.3, we have the following results.

THEOREM 3.1. *Let u be the solution to* (1.1). *Let $\mathcal{N}_h^\tau$ be the set of all vertices in $\mathcal{T}_h \cap \tau$. Assume that $\mathbf{x}_0 \in \mathcal{N}_h^\tau$ is away from the boundary of τ for a fixed distance, which is not changing when the mesh is refined.*

(i) *If* $u \in W^{2k,\infty}(\tau) \cap H^{k+1}(\Omega)$, *then*

$$|(u - P_1^k R_{h,\Omega} u)(\mathbf{x}_0)| \leq ch^{2k} |\ln h|^{\tilde{k}+1} (\|u\|_{W^{2k,\infty}(\tau)} + \|u\|_{H^{k+1}(\Omega)}). \tag{3.1}$$

(ii) *If* $u \in W^{2k+1,\infty}(\tau) \cap H^{k+1}(\Omega)$, *then*

$$|(\nabla u - P_2^k \overline{\nabla} R_{h,\Omega} u)(\mathbf{x}_0)| \leq ch^{2k} |\ln h|^{\tilde{k}+1} (\|u\|_{W^{2k+1,\infty}(\tau)} + \|u\|_{H^{k+1}(\Omega)}). \tag{3.2}$$

We need the following lemmas in order to prove Theorem 3.1.

LEMMA 3.2. *Assume that* $\rho(\mathbf{x}_0, \partial\Omega) \simeq 1$. *Then, for all real numbers* $r > 0$ *and integers* $m \geq 1$, *there holds the following estimate*

$$\|G_{\mathbf{x}_0}\|_{W^{m,\infty}(\Omega \setminus B(\mathbf{x}_0, r))} \leq cr^{2-m-n}. \tag{3.3}$$

Proof. Assume that there exists $d \simeq 1$ such that $B(\mathbf{x}_0, 2d) \subset \Omega$. We make the following claims.

(i) For $r > 0$ and $m \geq 1$, it holds that

$$\|G_{\mathbf{x}_0}\|_{W^{m,\infty}(B(\mathbf{x}_0,d) \setminus B(\mathbf{x}_0,r))} \leq cr^{2-m-n}. \tag{3.4}$$

(ii) For $m \geq 1$,

$$\|G_{\mathbf{x}_0}\|_{W^{m,\infty}(\Omega \setminus B(\mathbf{x}_0,d))} \leq c. \tag{3.5}$$

The result (3.3) follows immediately from (3.4) and (3.5). We shall show these two claims in the rest of the proof. To this end, we first prove that, for $\mathbf{x} \in \Omega$,

$$|G_{\mathbf{x}_0}(\mathbf{x})| \leq \begin{cases} c\big|\ln|\mathbf{x} - \mathbf{x}_0|\big| + c_1, & n = 2, \\ c|\mathbf{x} - \mathbf{x}_0|^{-1}, & n = 3. \end{cases} \tag{3.6}$$

Let R be sufficiently big such that $B(\mathbf{x}_0, R) \supset \Omega$. Set $E_{\mathbf{x}_0}(\mathbf{x}) = G_{\mathbf{x}_0}(\mathbf{x}) - G_{\mathbf{x}_0}^{B(\mathbf{x}_0,R)}(\mathbf{x})$. Note that

$$\mathcal{L}E_{\mathbf{x}_0}(\mathbf{x}) = 0 \quad \forall \mathbf{x} \in \Omega \quad \text{and} \quad E_{\mathbf{x}_0}(\mathbf{x}) \leq 0 \quad \forall \mathbf{x} \in \partial\Omega.$$

By the maximum principle, it follows that $E_{\mathbf{x}_0}(\mathbf{x}) \leq 0$ for all $x \in \Omega$. Furthermore, we obtain

$$|G_{\mathbf{x}_0}(\mathbf{x})| = G_{\mathbf{x}_0}(\mathbf{x}) \leq G_{\mathbf{x}_0}^{B(\mathbf{x}_0,d)}(\mathbf{x}),$$

which gives (3.6).

Now we consider (i). Let $\psi \in C^\infty(B(\mathbf{x}_0, R))$ be the cutoff function satisfying $\psi(\mathbf{x}) = 1$ if $\mathbf{x} \in B(\mathbf{x}_0, 3d/2)$, $\psi(\mathbf{x}) = 0$ if $\mathbf{x} \in B(\mathbf{x}_0, R) \setminus B(\mathbf{x}_0, 2d)$, and $|D^j \psi(\mathbf{x})| \leq c$ for all positive integers j. Consider the split $G_{\mathbf{x}_0} = G_1 + G_2$, where

$$G_1(\mathbf{x}) = \psi(\mathbf{x}) G_{\mathbf{x}_0}^{B(\mathbf{x}_0,R)}(\mathbf{x}). \tag{3.7}$$

Krasovskiĭ [12] showed that, for $\mathbf{x} \in B(\mathbf{x}_0, R)$ and $m \geq 0$,

$$|D^m G_{\mathbf{x}_0}^{B(\mathbf{x}_0,R)}(\mathbf{x})| \leq \begin{cases} c|\mathbf{x} - \mathbf{x}_0|^{2-m-n} & \text{if } m + n \geq 3, \\ c(|\ln|\mathbf{x} - \mathbf{x}_0|| + 1) & \text{if } m = 0 \text{ and } n = 2. \end{cases}$$

It follows that, for $\mathbf{x} \in B(\mathbf{x}_0, d) \setminus B(\mathbf{x}_0, r)$ and $m \geq 0$,

$$|D^m G_1(\mathbf{x})| \leq \begin{cases} c|\mathbf{x}-\mathbf{x}_0|^{2-m-n} & \text{if } m+n \geq 3, \\ c(|\ln|\mathbf{x}-\mathbf{x}_0||+1) & \text{if } m=0 \text{ and } n=2. \end{cases} \tag{3.8}$$

On the other hand, by (3.7), we have

$$\mathcal{L}G_2(\mathbf{x}) = \mathcal{L}(G_{\mathbf{x}_0} - G_1)(\mathbf{x}) = 0 \quad \forall \mathbf{x} \in B(\mathbf{x}_0, 3d/2). \tag{3.9}$$

Since $d \simeq 1$, by (3.6), (3.8), (3.9), and [3, Theorem 3.7], we get, for all integers $j \geq 1$,

$$\begin{aligned} &\|G_2\|_{H^j(B(\mathbf{x}_0,d))} \leq c\|G_2\|_{L^2(\partial B(\mathbf{x}_0,3d/2))} \\ &\leq c(\|G_1\|_{L^2(\partial B(\mathbf{x}_0,3d/2))} + \|G_{\mathbf{x}_0}\|_{L^2(\partial B(\mathbf{x}_0,3d/2))}) \leq c. \end{aligned}$$

It follows that, for all $m \geq 1$,

$$\|G_2\|_{W^{m,\infty}(B(\mathbf{x}_0,d))} \leq c. \tag{3.10}$$

By (3.8) and (3.10), we obtain (3.4).

Using a similar approach, we can prove (ii). Let $\phi \in C^\infty(B(\mathbf{x}_0, R))$ be the cutoff function such that $\phi(\mathbf{x}) = 1$ if $\mathbf{x} \in B(\mathbf{x}_0, d/2)$, $\phi(\mathbf{x}) = 0$ if $\mathbf{x} \in B(\mathbf{x}_0, R) \setminus B(\mathbf{x}_0, 3d/4)$, and $|D^j \phi(\mathbf{x})| \leq c$ for all positive integers j. Put $G_{\mathbf{x}_0} = \overline{G}_1 + \overline{G}_2$, where

$$\overline{G}_1(\mathbf{x}) = \phi(\mathbf{x}) G_{\mathbf{x}_0}^{B(\mathbf{x}_0,R)}(\mathbf{x}). \tag{3.11}$$

For $\mathbf{x} \in \Omega \setminus B(\mathbf{x}_0, d)$, a similar estimate to (3.8) is obtained for $\overline{G}_1$, which reads

$$|D^m \overline{G}_1(\mathbf{x})| \leq c, \tag{3.12}$$

since the distance $|\mathbf{x} - \mathbf{x}_0|$ is big. On the other hand, by (3.11), we have

$$\mathcal{L}\overline{G}_2(\mathbf{x}) = \mathcal{L}(G_{\mathbf{x}_0} - \overline{G}_1)(\mathbf{x}) = 0 \quad \forall \mathbf{x} \in \Omega \setminus B(\mathbf{x}_0, 3d/4). \tag{3.13}$$

Note also that

$$\overline{G}_2(\mathbf{x}) = 0, \quad x \in \partial\Omega. \tag{3.14}$$

Combining (3.6), (3.13), (3.14), and [3, Theorem 3.7], for all integers $j \geq 1$, it holds

$$\begin{aligned} &\|\overline{G}_2\|_{H^j(\Omega\setminus B(\mathbf{x}_0,d))} \leq c\|\overline{G}_2\|_{L^2(\partial B(\mathbf{x}_0,3d/4))} \\ &\leq c\|\overline{G}_1\|_{L^2(\partial B(\mathbf{x}_0,3d/4))} + c\|G_{\mathbf{x}_0}\|_{L^2(\partial B(\mathbf{x}_0,3d/4))} \leq c. \end{aligned}$$

This implies that, for all integers $m \geq 1$,

$$\|\overline{G}_2\|_{W^{m,\infty}(\Omega\setminus B(\mathbf{x}_0,d))} \leq c. \tag{3.15}$$

The second claim follows from (3.12) and (3.15). ∎

Based on Lemma 3.2, applying the same arguments in the proof of [21, Theorem 6.1], we have the following result.

LEMMA 3.3. *For $r > 0$, there exists the following estimate*

$$\|G_{\mathbf{x}_0} - R_{h_s,\Omega} G_{\mathbf{x}_0}\|_{W^{1,\infty}(\Omega\setminus B(\mathbf{x}_0,r))} \leq ch^k r^{1-n-k} |\ln h|^{\tilde{k}}. \tag{3.16}$$

We shall next prove Theorem 3.1.

Proof of Theorem 3.1. (i) First, show (3.1). Let $d \simeq 1$ such that $B(\mathbf{x}_0, 2d) \subset \tau$ and $B(\mathbf{x}_0, 2^{J_1(k)+1}d) \subset \widehat{\Omega}$. Let χ be the cutoff function satisfying $0 \le \chi \le 1$, $\chi = 1$ in $B(\mathbf{x}_0, d)$, $\chi = 0$ in $\widehat{\Omega} \setminus B(\mathbf{x}_0, 2d)$, and $\|\chi\|_{W^{2k,\infty}(\widehat{\Omega})} \le c$. Set $u_1(\mathbf{x}) = \chi(\mathbf{x})u(\mathbf{x})$. Write

$$(u - P_1^k R_{h,\Omega}u)(\mathbf{x}_0) = (u_1 - \widehat{P}_1^k R_{h,\widehat{\Omega}}u_1)(\mathbf{x}_0) + (\widehat{P}_1^k R_{h,\widehat{\Omega}}u_1 - P_1^k R_{h,\Omega}u)(\mathbf{x}_0). \tag{3.17}$$

By (2.6), one has

$$\begin{aligned} |(u_1 - \widehat{P}_1^k R_{h,\widehat{\Omega}}u_1)(\mathbf{x}_0)| &\le ch^{2k}|\ln h|^{\tilde{k}+1}\|u_1\|_{W^{2k,\infty}(\widehat{\Omega})} \\ &\le ch^{2k}|\ln h|^{\tilde{k}+1}\|u\|_{W^{2k,\infty}(\tau)}. \end{aligned} \tag{3.18}$$

We claim that

$$|(\widehat{P}_1^k R_{h,\widehat{\Omega}}u_1 - P_1^k R_{h,\Omega}u)(\mathbf{x}_0)| \le ch^{2k}|\ln h|^{\tilde{k}}\|u\|_{H^{k+1}(\Omega)}. \tag{3.19}$$

Then (3.1) follows from (3.17), (3.18), and (3.19).

We prove the claim below. Consider $(R_{h_s,\widehat{\Omega}}u_1 - R_{h_s,\Omega}u)(\mathbf{x}_0)$ for all $0 \le s \le J_2(k)$. Note that

$$\begin{aligned} (R_{h_s,\widehat{\Omega}}u_1 - u_1)(\mathbf{x}_0) &= \mathcal{B}_{\widehat{\Omega}}(I_{h_s}^k u_1 - u_1, G_{\mathbf{x}_0}^{\widehat{\Omega}} - R_{h_s,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}), \\ (R_{h_s,\Omega}u - u)(\mathbf{x}_0) &= \mathcal{B}_{\Omega}(I_{h_s}^k u - u, G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0}), \end{aligned}$$

and $u_1(\mathbf{x}_0) = u(\mathbf{x}_0)$. It follows immediately that

$$\begin{aligned} &R_{h_s,\widehat{\Omega}}u_1(\mathbf{x}_0) - R_{h_s,\Omega}u(\mathbf{x}_0) = (R_{h_s,\widehat{\Omega}}u_1 - u_1)(\mathbf{x}_0) - (R_{h_s,\Omega}u - u)(\mathbf{x}_0) \\ &= \mathcal{B}_{\widehat{\Omega}}(I_{h_s}^k u_1 - u_1, G_{\mathbf{x}_0}^{\widehat{\Omega}} - R_{h_s,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}) - \mathcal{B}_{\Omega}(I_{h_s}^k u - u, G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0}) \\ &= \mathcal{B}_{\Omega\setminus\tau}(I_{h_s}^k u - u, G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0}) + \mathcal{B}_{\widehat{\Omega}\setminus\tau}(I_{h_s}^k u_1 - u_1, G_{\mathbf{x}_0}^{\widehat{\Omega}} - R_{h_s,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}) \\ &\quad + \mathcal{B}_{\tau}(I_{h_s}^k u_1 - u_1, G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0} - G_{\mathbf{x}_0}^{\widehat{\Omega}} + R_{h_s,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}) \\ &= K_1 + K_2 + K_3. \end{aligned} \tag{3.20}$$

We shall estimate the three terms in the right-hand side of (3.20). For K_1, from standard interpolation error estimates and Lemma 3.3, it follows that

$$\begin{aligned} |K_1| &\le c\|I_{h_s}^k u - u\|_{H^1(\Omega\setminus\tau)}\|G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0}\|_{H^1(\Omega\setminus\tau)} \\ &\le ch^k\|u\|_{H^{k+1}(\Omega)}ch^k|\ln h|^{\tilde{k}} \le ch^{2k}|\ln h|^{\tilde{k}}\|u\|_{H^{k+1}(\Omega)}. \end{aligned} \tag{3.21}$$

For K_2, by [21, Theorem 6.1], we have

$$|K_2| \le ch^{2k}|\ln h|^{\tilde{k}}\|u_1\|_{H^{k+1}(\Omega)} \le ch^{2k}|\ln h|^{\tilde{k}}\|u\|_{H^{k+1}(\Omega)}. \tag{3.22}$$

We next evaluate K_3. Denote by $\mathcal{T}_{h_s}^{\tau}$ the restriction of $\widehat{\mathcal{T}}_{h_s}$ to τ. Put

$$\begin{aligned} &G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0} - G_{\mathbf{x}_0}^{\widehat{\Omega}} + R_{h_s,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}} \\ =&(E_G - R_{h_s,\tau}E_G) + (R_{h_s,\tau}G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0}) + (R_{h_s,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}} - R_{h_s,\tau}G_{\mathbf{x}_0}^{\widehat{\Omega}}), \end{aligned} \tag{3.23}$$

where $E_G = G_{\mathbf{x}_0} - G_{\mathbf{x}_0}^{\widehat{\Omega}}$. To measure $E_G - R_{h_s,\tau}E_G$, by the same arguments as in the proof of Lemma 3.2, we have, for all integers $m \geq 1$,

$$\|G_{\mathbf{x}_0}^{\widehat{\Omega}}\|_{W^{m,\infty}(\widehat{\Omega}\setminus\mathcal{B}(\mathbf{x}_0,d))} \leq c.$$

This, together with an application of Lemma 3.2, gives

$$\|E_G\|_{W^{m,\infty}(\tau\setminus\mathcal{B}(\mathbf{x}_0,d))} \leq c. \tag{3.24}$$

Note that

$$\mathcal{L}E_G(\mathbf{x}) = 0 \quad \forall \mathbf{x} \in \tau. \tag{3.25}$$

Using (3.24), (3.25), and [3, Theorem 3.7], one has

$$\|E_G\|_{W^{m,\infty}(\mathcal{B}(\mathbf{x}_0,d))} \leq c\|E_G\|_{L^\infty(\partial\tau)} \leq c.$$

This, together with (3.24), gives

$$\|E_G\|_{W^{m,\infty}(\tau)} \leq c. \tag{3.26}$$

It follows that

$$\|E_G - R_{h_s,\tau}E_G\|_{H^1(\tau)} \leq ch^k. \tag{3.27}$$

Next, estimate $R_{h_s,\tau}G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0}$. By Lemma 3.3, we have

$$\begin{aligned} &\|R_{h_s,\tau}G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0}\|_{H^1(\tau\setminus B(\mathbf{x}_0,d))} \\ &\leq \|R_{h_s,\tau}G_{\mathbf{x}_0} - G_{\mathbf{x}_0}\|_{H^1(\tau\setminus B(\mathbf{x}_0,d))} + \|G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0}\|_{H^1(\tau\setminus B(\mathbf{x}_0,d))} \\ &\leq ch^k|\ln h|^{\tilde{k}} + ch^k|\ln h|^{\tilde{k}} = ch^k|\ln h|^{\tilde{k}}, \end{aligned} \tag{3.28}$$

where the same reasoning to show (2.16) has been used (cf. [21]). For all $v \in S_0^{h_s}(\tau)$,

$$\begin{aligned} &\mathcal{B}_\tau(R_{h_s,\tau}G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0}, v) \\ =&\mathcal{B}_\tau(R_{h_s,\tau}G_{\mathbf{x}_0} - G_{\mathbf{x}_0}, v) + \mathcal{B}_\tau(G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0}, v) = 0. \end{aligned}$$

Then applying the arguments in [22, 20], we obtain

$$\begin{aligned} &\|R_{h_s,\tau}G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0}\|_{H^1(B(\mathbf{x}_0,d))} \\ &\leq c\|R_{h_s,\tau}G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0}\|_{L^2(\tau\setminus B(\mathbf{x}_0,3d/2))} \leq ch^k|\ln h|^{\tilde{k}}. \end{aligned} \tag{3.29}$$

Combining (3.28) and (3.29) gives

$$\|R_{h_s,\tau}G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0}\|_{H^1(\tau)} \leq ch^k|\ln h|^{\tilde{k}}. \tag{3.30}$$

Similarly, we obtain

$$\|R_{h_s,\tau}G_{\mathbf{x}_0}^{\widehat{\Omega}} - R_{h_s,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}\|_{H^1(\tau)} \leq ch^k|\ln h|^{\tilde{k}}. \tag{3.31}$$

Inserting (3.27), (3.30), and (3.31) into (3.23) gives

$$\|G_{\mathbf{x}_0} - R_{h_s,\Omega}G_{\mathbf{x}_0} - G_{\mathbf{x}_0}^{\widehat{\Omega}} + R_{h_s,\widehat{\Omega}}G_{\mathbf{x}_0}^{\widehat{\Omega}}\|_{H^1(\tau)} \leq ch^k|\ln h|^{\tilde{k}}. \tag{3.32}$$

By (3.20) and (3.32), we have

$$|K_3| \leq ch^{2k}|\ln h|^{\tilde{k}}\|u\|_{H^{k+1}(\Omega)}. \tag{3.33}$$

Therefore, by (3.20), (3.21), (3.22), and (3.33), we have

$$|(R_{h_s,\widehat{\Omega}}u_1 - R_{h_s,\Omega}u)(\mathbf{x}_0)| \leq ch^{2k}|\ln h|^{\tilde{k}}\|u\|_{H^{k+1}(\Omega)}. \tag{3.34}$$

The claim (3.19) follows.

(ii) We turn now to the proof of (3.2). Recast

$$\begin{aligned}&(\nabla u - P_2^k\overline{\nabla}R_{h,\Omega}u)(\mathbf{x}_0)\\ &= (\nabla u_1 - \widehat{P}_2^k\overline{\nabla}R_{h,\widehat{\Omega}}u_1)(\mathbf{x}_0) + (\widehat{P}_2^k\overline{\nabla}R_{h,\widehat{\Omega}}u_1 - P_2^k\overline{\nabla}R_{h,\Omega}u)(\mathbf{x}_0).\end{aligned} \tag{3.35}$$

We need to estimate the two terms of the right-hand side. For $(\nabla u_1 - P_2^k\overline{\nabla}R_{h,\widehat{\Omega}}u_1)(\mathbf{x}_0)$, from Theorem 2.3, it follows that

$$|(\nabla u_1 - \widehat{P}_2^k\overline{\nabla}R_{h,\widehat{\Omega}}u_1)(\mathbf{x}_0)| \leq ch^{2k}|\ln h|^{\tilde{k}+1}\|u\|_{W^{2k+1,\infty}(\tau)}. \tag{3.36}$$

For the other term, we first estimate $(\overline{\nabla}R_{h,\widehat{\Omega}}u_1 - \overline{\nabla}R_{h,\Omega}u)(\mathbf{x}_0)$ for all $0 \leq s \leq J_1(k)$. One observes that, similarly as in (3.34), for all $\mathbf{x} \in B(\mathbf{x}_0, d)$,

$$|(R_{h_s,\widehat{\Omega}}u_1 - R_{h_s,\Omega}u)(\mathbf{x})| \leq ch^{2k}|\ln h|^{\tilde{k}}\|u\|_{H^{k+1}(\Omega)}. \tag{3.37}$$

On the other hand, for all $v \in S_0^{h_s}(\tau)$,

$$\begin{aligned}&\mathcal{B}_\tau(R_{h_s,\widehat{\Omega}}u_1 - R_{h_s,\Omega}u, v)\\ &= \mathcal{B}_\tau(R_{h_s,\widehat{\Omega}}u_1 - u_1, v) + \mathcal{B}_{B(\mathbf{x}_0,d)}(u - R_{h_s,\Omega}u, v) = 0.\end{aligned} \tag{3.38}$$

Applying (3.37), (3.38), and the arguments of Schatz and Wahlbin [22] and Schatz, Sloan, and Wahlbin [20], we obtain

$$\begin{aligned}|(\overline{\nabla}R_{h_s,\widehat{\Omega}}u_1 - \overline{\nabla}R_{h_s,\Omega}u)(\mathbf{x}_0)| &\leq c\|R_{h_s,\widehat{\Omega}}u_1 - R_{h_s,\Omega}u\|_{L^\infty(B(\mathbf{x}_0,d))}\\ &\leq ch^{2k}|\ln h|^{\tilde{k}}\|u\|_{H^{k+1}(\Omega)}.\end{aligned}$$

Furthermore, we have

$$|(\widehat{P}_2^k\overline{\nabla}R_{h,\widehat{\Omega}}u_1 - P_2^k\overline{\nabla}R_{h,\Omega}u)(\mathbf{x}_0)| \leq ch^{2k}|\ln h|^{\tilde{k}}\|u\|_{H^{k+1}(\Omega)}. \tag{3.39}$$

The desired result (3.2) follows from (3.36) and (3.39). ∎

Remark 3.1. When $k = 1$, $P_1^1R_{h,\Omega}u = u^h$ and $P_2^1\overline{\nabla}R_{h,\Omega}u = \overline{\nabla}u^h$. The results of Theorem 3.1 coincide with the optimal error estimate for displacement, and the superconvergence in gradient by simple average recovery (see, e.g., [30]).

When $k = 2$, $P_1^2R_{h,\Omega}u = u^h$. The result by (3.1) aligns with the natural superconvergence results in the literature (cf., e.g., [16, 17, 20, 27]). But (3.2) is an ultraconvergence result.

Table 1
Example 4.1. *Displacement and gradient errors and rates of convergence for $\mathbb{P}_2$ and $\mathbb{Q}_2$ elements.*

		Triangular regular mesh		Triangular equilateral mesh		Rectangular mesh	
$A = \widehat{A}$		$\|u - P_1^2 u^h\|_{\infty,h}$ and convergence rates					
	16	1.0954e-6	–	1.7488e-7	–	3.0590e-7	–
	20	4.4112e-7	4.1	7.1533e-8	4.0	1.2584e-7	4.0
	24	2.1075e-7	4.1	3.4471e-8	4.0	6.0749e-8	4.0
	28	1.1311e-7	4.0	1.8599e-8	4.0	3.2785e-8	4.0
		$\|\nabla u - P_2^2(\overline{\nabla} u^h)\|_{\infty,h}$ and convergence rates					
	16	3.0629e-5	–	1.7779e-6	–	1.8559e-6	–
	20	1.2509e-5	4.0	7.2902e-7	4.0	7.5765e-7	4.0
	24	6.0231e-6	4.0	3.5178e-7	4.0	3.6652e-7	4.0
	28	3.2481e-6	4.0	1.8995e-7	4.0	1.9751e-7	4.0
$A = \widetilde{A}$		$\|u - P_1^2 u^h\|_{\infty,h}$ and convergence rates					
	16	1.3956e-6	–	2.6456e-7	–	2.5331e-7	–
	20	5.7420e-7	4.0	1.0837e-7	4.0	1.0310e-7	4.0
	24	2.7807e-7	4.0	5.2261e-8	4.0	4.9778e-8	4.0
	28	1.5026e-7	4.0	2.8209e-8	4.0	2.6816e-8	4.0
		$\|\nabla u - P_2^2(\overline{\nabla} u^h)\|_{\infty,h}$ and convergence rates					
	16	2.3475e-5	–	1.8548e-6	–	2.5366e-6	–
	20	9.6030e-6	4.0	7.6104e-7	4.0	1.0388e-6	4.0
	24	4.6284e-6	4.0	3.6736e-7	4.0	5.0098e-7	4.0
	28	2.4976e-6	4.0	1.9840e-7	4.0	2.7043e-7	4.0

4. Numerical Examples. In this section, we test numerically the ultraconvergence results in section 3. Since the errors of the FE solutions converge very fast, to avoid error pollution due to boundary effects, we consider numerical errors in a subdomain $\widetilde{\Omega}$ of Ω whose boundary is 1/4 away from $\partial\Omega$ (cf. Remark 4.1). Moreover, to avoid errors from quadrature rules, stiffness matrices and load vectors are all obtained exactly by using Maple, a computer algebra system produced by Maplesoft$^{\text{TM}}$. Let $\widetilde{\mathcal{N}}_h$ be the restriction of $\mathcal{N}_h$ to $\widetilde{\Omega}$. Define a discrete maximum norm $\|v\|_{\infty,h} = \max_{\mathbf{x}\in\widetilde{\mathcal{N}}_h} |v(\mathbf{x})|$, which will be used to measure numerical errors.

Example 4.1 (two-dimensional case). We first consider problem (1.1) in two dimensions. We shall report convergence rates of errors in displacement and gradient for $k = 2$, 3, and 4; the case for $k = 1$ is not reported due to Remark 3.1. Let the coefficient matrix A be

$$\widehat{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \widetilde{A} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix},$$

so that the impact of the coefficients can be illustrated. Three mesh patterns are tested. For rectangular mesh and regular triangular mesh, set $\Omega = [0,1]^2$. For equilateral triangular mesh, let Ω be the triangular region with vertices $(0,0)$, $(1,0)$, and $(1/2, \sqrt{3}/2)$. Let f be the function that fits (1.1) with the exact solution $u(x,y) = x(1-x^3)y^4(1-y)$ when $k = 2, 3$, and $u(x,y) = x^5(1-x^6)y^4(1-y^7)$ when $k = 4$.

In Table 1, the errors and convergence rates for second order FEs are provided for $A = \widehat{A}$ and $\widetilde{A}$. The errors measured in $\|\cdot\|_{\infty,h}$ appear in the left columns and the associated rates of convergence are on the right. Each convergence rate r is computed under the presumption that the error converges in order h^r, which is superior to the rate obtained in Theorem 3.1. Forth order convergence rates are observed in all cases, as predicted in section 3.

TABLE 2

Example 4.1. *Displacement and gradient errors and rates of convergence for* $\mathbb{P}_3$ *and* $\mathbb{Q}_3$ *elements.*

		Triangular regular mesh		Triangular equilateral mesh		Rectangular mesh	
$A = \widehat{A}$		$\|u - P_1^3 u^h\|_{\infty,h}$ and convergence rates					
	16	9.0820e-10	–	2.5496e-12	–	7.4144e-12	–
	20	2.3791e-10	6.0	6.6605e-13	6.0	1.9341e-12	6.0
	24	7.9645e-11	6.0	2.2270e-13	6.0	6.3723e-13	6.1
	28	3.1576e-11	6.0	8.7797e-14	6.0	2.3955e-13	6.3
		$\|\nabla u - P_2^3(\overline{\nabla} u^h)\|_{\infty,h}$ and convergence rates					
	16	1.1902e-8	–	5.5029e-10	–	7.3326e-11	–
	20	3.1212e-9	6.0	1.4426e-10	6.0	1.9223e-11	6.0
	24	1.0455e-9	6.0	4.8309e-11	6.0	6.4387e-12	6.0
	28	4.1463e-10	6.0	1.9154e-11	6.0	2.5610e-12	6.0
$A = \widetilde{A}$		$\|u - P_1^3 u^h\|_{\infty,h}$ and convergence rates					
	16	1.0692e-9	–	4.4839e-12	–	3.2392e-12	–
	20	2.9728e-10	5.7	1.1763e-12	6.0	8.4636e-13	6.0
	24	1.0467e-10	5.7	3.9484e-13	6.0	2.8827e-13	5.9
	28	4.3370e-11	5.7	1.5768e-13	6.0	1.2170e-13	5.6
		$\|\nabla u - P_2^3(\overline{\nabla} u^h)\|_{\infty,h}$ and convergence rates					
	16	1.0941e-8	–	5.4928e-10	–	6.1653e-11	–
	20	2.7938e-9	6.1	1.4410e-10	6.0	1.5871e-11	6.1
	24	9.1312e-10	6.1	4.8255e-11	6.0	5.2429e-12	6.1
	28	3.5395e-10	6.1	1.9133e-11	6.0	2.0549e-12	6.1

TABLE 3

Example 4.1. *Displacement and gradient errors and rates of convergence for* $\mathbb{P}_4$ *and* $\mathbb{Q}_4$ *elements.*

		Triangular regular mesh		Triangular equilateral mesh		Rectangular mesh	
$A = \widehat{A}$		$\|u - P_1^4 u^h\|_{\infty,h}$ and convergence rates					
	4	7.3502e-6	–	6.1941e-9	–	1.7351e-8	–
	8	1.9021e-8	8.6	1.6425e-11	8.6	1.7442e-11	10.0
	12	6.8783e-10	8.2	5.7631e-13	8.3	5.0058e-13	8.8
	16	6.7035e-11	8.1	6.3498e-14	7.7	3.8011e-14	9.0
		$\|\nabla u - P_2^4(\overline{\nabla} u^h)\|_{\infty,h}$ and convergence rates					
	4	3.9963e-5	–	1.5392e-8	–	4.6570e-8	–
	8	1.9729e-7	7.7	8.1943e-11	7.6	3.2362e-10	7.2
	12	6.8919e-9	8.3	3.1421e-12	8.0	1.2952e-11	7.9
	16	6.6402e-10	8.1	3.1745e-13	8.0	1.3213e-12	7.9
$A = \widetilde{A}$		$\|u - P_1^4 u^h\|_{\infty,h}$ and convergence rates					
	4	3.7876e-6	–	8.7511e-9	–	1.0290e-7	–
	8	8.5886e-9	8.8	2.4629e-11	8.5	1.5056e-10	9.4
	12	2.1109e-10	9.1	9.7129e-13	8.0	5.2715e-12	8.3
	16	1.8698e-11	8.4	1.0011e-13	7.9	6.4047e-13	7.3
		$\|\nabla u - P_2^4(\overline{\nabla} u^h)\|_{\infty,h}$ and convergence rates					
	4	3.6797e-5	–	1.6238e-8	–	2.5851e-7	–
	8	1.2657e-7	8.2	7.8061e-11	7.7	2.7813e-10	9.9
	12	4.6658e-9	8.1	2.9657e-12	8.1	5.5320e-12	9.7
	16	4.6332e-10	8.0	2.9825e-13	8.0	8.1380e-13	6.7

In Tables 2 and 3, the same numerical data as presented in Table 1 are provided for the cases of $k = 3$ and 4, respectively. The numerical results suggest that errors in both displacement and gradient converge with rate h^{2k} for all three meshes and both coefficient matrices. The ultraconvergence results in Theorem 3.1 are confirmed.

Example 4.2 (three-dimensional case). We next consider problem (1.1) in three dimensions. Due to Remark 3.1, we shall examine convergence rates of displacement

TABLE 4
Example 4.2. Displacement and gradient errors and rates of convergence for $\mathbb{P}_2$ and $\mathbb{Q}_2$ elements.

	Tetrahedral mesh				Hexahedral mesh			
$1/h$	$\widehat{A}$		$\widetilde{A}$		$\widehat{A}$		$\widetilde{A}$	
	$\|u - P_1^2 u^h\|_{\infty,h}$ and convergence rates							
4	8.4732e-5	–	1.2989e-4	–	3.2425e-6	–	8.0866e-6	–
8	5.5467e-6	3.9	9.1321e-6	3.8	2.5041e-7	3.7	5.9265e-7	3.8
12	1.0876e-6	4.0	1.7880e-6	4.0	5.0300e-8	4.0	1.1881e-7	4.0
16	3.4284e-7	4.0	5.6023e-7	4.0	1.5923e-8	4.0	3.7559e-8	4.0
	$\|\nabla u - P_2^2(\overline{\nabla} u^h)\|_{\infty,h}$ and convergence rates							
4	1.1624e-4	–	1.4345e-4	–	1.4697e-5	–	4.1150e-5	–
8	8.2705e-6	3.8	1.1162e-5	3.7	9.4547e-7	4.0	1.9371e-6	4.4
12	1.6380e-6	4.0	2.2732e-6	3.9	1.9160e-7	3.9	3.5817e-7	4.2
16	5.3392e-7	3.9	7.2425e-7	4.0	6.1052e-8	4.0	1.1237e-7	4.0

TABLE 5
Example 4.2. Displacement and gradient errors and rates of convergence for $\mathbb{P}_3$ and $\mathbb{Q}_3$ elements.

	Tetrahedral mesh				Hexahedral mesh			
$1/h$	$\widehat{A}$		$\widetilde{A}$		$\widehat{A}$		$\widetilde{A}$	
	$\|u - P_1^3 u^h\|_{\infty,h}$ and convergence rates							
4	6.2971e-8	–	3.1736e-7	–	1.1304e-8	–	8.8181e-9	–
8	2.3223e-10	8.1	8.2811e-10	8.6	7.8890e-12	10.5	5.9407e-11	7.2
12	2.2723e-11	5.7	6.5300e-11	6.3	8.0739e-13	5.6	4.3343e-12	6.5
	$\|\nabla u - P_2^3(\overline{\nabla} u^h)\|_{\infty,h}$ and convergence rates							
4	1.6915e-6	–	4.1046e-6	–	4.8029e-7	–	3.8147e-7	–
8	5.1484e-9	8.4	2.1415e-8	7.6	1.0982e-9	8.8	6.7174e-10	9.1
12	1.8678e-10	8.2	7.9347e-10	8.1	1.0817e-11	11.4	2.4549e-11	8.2

and gradient errors for $k = 2$ and 3. Let the coefficient matrix A be

$$\widehat{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{or} \quad \widetilde{A} = \begin{pmatrix} 1 & -0.45 & -0.5 \\ -0.45 & 1 & -0.4 \\ -0.5 & -0.4 & 1 \end{pmatrix}.$$

We considered a hexahedral mesh and tetrahedral mesh in Figure 2 on $\Omega = [0,1]^3$. Let f be the function that fits (1.1) with the exact solution $u(x,y,z) = e^y \sin x \cos z$.

In Table 4, the errors and convergence rates for second order FEs are provided. The data are organized in a similar way as those in tables for Example 4.1. Forth order convergence rates are observed in all cases, as predicted in section 3.

In Table 5, we provide the errors and convergence rates for third order FEs. The rates of convergence are about h^6 or even better. The much higher convergence rates are mainly due to the "poor" results for big h, in which cases the round-off errors have high impacts.

Remark 4.1. The superconvergence properties studied in this paper are *local* results, namely, for any $\delta > 0$, there is an $\hbar(\delta) > 0$ such that when $h_0 \leq \hbar(\delta)$ the superconvergence properties hold in an interior region of distance at least δ away from the boundary $\partial\Omega$. For Examples 4.1 and 4.2, instead of using fixed subregions of distance $1/4$ away from $\partial\Omega$, the same asymptotic convergence rates are obtained in subregions of $2h_0$ or $3h_0$ away from $\partial\Omega$ (i.e. $\hbar(\delta) = \delta/2$ or $\delta/3$). But in general, $\hbar(\delta)$ depends on the regularity of the boundary $\partial\Omega$, which is not the focus of this study.

Acknowledgment. The authors acknowledge the valuable inputs of the anonymous referees, whose comments helped to improve the quality of the paper.

REFERENCES

[1] M. Asadzadeh, A. H. Schatz, and W. Wendland, *Asymptotic error expansions for the finite element method for second order elliptic problems in* $\mathbb{R}^N$, $N \geq 2$. I*: Local interior expansions*, SIAM J. Numerical Analysis, 48 (2010), pp. 2000–2017.

[2] M. Asadzadeh, A. H. Schatz, and W. Wendland, *A new approach to Richardson extrapolation in the finite element method for second order elliptic problems*, Math. Comp., 78 (2009), pp. 1951–1973.

[3] C. Bacuta, V. Nistor, and L. T. Zikatanov, *Improving the rate of convergence of high-order finite elements on polyhedra* I*: A priori estimates*, Numer. Funct. Anal. Optim., 26 (2005), pp. 613–639.

[4] J. H. Bramble and A. H. Schatz, *High order local accuracy by averaging in the finite element method*, Math. Comp., 31 (1977), pp. 94–111.

[5] J. Bey, *Simplicial grid refinement: On Freudenthal's algorithm and the optimal number of congruence classes*, Numer. Math., 85 (2000), pp. 1–29.

[6] H. Blum, *On Richardson extrapolation for linear finite elements on domains with reentrant corners*, Z. Angew. Math. Mech., 29 (1987), pp. 351–353.

[7] C. Chen and Q. Lin, *Extrapolation of finite element approximations in a rectanglar domain*, J. Comput. Math., 7 (1989), pp. 227–233.

[8] C. Chen and S. Hu, *The highest order superconvergence for bi-k degree rectangular elements at nodes: a proof of $2k$-conjecture*, Math. Comp., 82 (2013), pp. 1337–1355.

[9] B. Cockburn, M. Luskin, C.-W. Shu, and E. Suli, *Enhanced accuracy by post-processing for finite element methods for hyperbolic equations*, Math. Comp., 72 (2003), pp. 577–606.

[10] J. Douglas, Jr. and T. Dupont, *Some superconvergence results for Galerkin methods for the approximate solution of two-point boundary problems*, in Topics in Numerical Analysis, Academic, London, 1973, pp. 89–92.

[11] W. He and Z. Zhang, $2k$ *superconvergence of Q_k finite element by anisotropic mesh approximation in weighted Sobolev spaces*, Math. Comp., in press, doi:10.1090/mcom/3159.

[12] J. P. Krasovskiĭ, *Isolation of singularities of the Green's function*, Math. USSR-IZV, 1 (1967), pp. 935–966.

[13] Q. Lin and T. Lu, *Asymptotic expansions the finite element approximation of elliptic problem on polygonal domains*, in Proceedings 6th International Symposium on Computing Methods in Applied Sciences and Engineering, Versailles, 1983, North-Holland, Amsterdam, 1985, pp. 317–321.

[14] Q. Lin and J. P. Wang, *Some Expansions for Finite Element Approximation*, Technical report IMS 15, Academia Sinica, Institute of Mathematical Sciences, Chendu, China (1984), pp. 1–11.

[15] Q. Lin and J. Lin, *Finite Element Methods: Accuracy and Improvement*, Science Press, Beijing, 2006.

[16] R. Lin and Z. Zhang, *Natural superconvergent points of triangular finite elements*, Numer. Methods Partial Differential Equations 20 (2004), pp. 864–906.

[17] R. Lin and Z. Zhang, *Natural superconvergence points in three-dimensional finite elements*, SIAM J. Numer. Anal., 46 (2008), pp. 1281–1297.

[18] R. Rannacher, *Extrapolation techniques in the finite element method: A survey*, in Proceedings of the Summer School in Numerical Analysis, Helsinki, 1988, Helsinki University of Technology, Espoo, Finland, 1988, pp. 80–113.

[19] A. H. Schatz, *Pointwise error estimate and asymptotic error expansion inequalities for the finite element method on irregular grids: Part* I*. Global estimates*, Math. Comp., 67 (1998), pp. 877–899.

[20] A. H. Schatz, I. H. Sloan, and L. B. Wahlbin, *Superconvergence in finite element methods and meshes that are locally symmetric with respect to a point*, SIAM J. Numer. Anal., 33 (1996), pp. 505–521.

[21] A. H. Schatz and L. B. Wahlbin, *Interior maximum norm estimates for finite element methods*, Math. Comp., 31 (1977), pp. 414–442.

[22] A. H. Schatz and L. B. Wahlbin, *Interior maximum norm estimates for finite element methods*, Part II, Math. Comp., 211 (1995), pp. 907–928.

[23] A. H. SCHATZ AND L. B. WAHLBIN, *Asymptotically exact a posterior estimators for the pointwise gradient error on each element in irregular meshes*. Part II: The piecewise linear case, Math. Comp., 73 (2004), pp. 517–523.
[24] R. SCOTT, *Optimal L^∞ estimate for the finite element method on irregular mehses*, Math. Comp., 30 (1976), pp. 681–697.
[25] V. THOMÉE, *High order local approximation to derivatives in the finite element method*, Math. Comp., 31 (1977), pp. 652–660.
[26] V. THOMÉE, *Negative norm estimates and superconvergence in Galerkin methods for parabolic problems*, Math. Comp., 34 (1980), pp. 93–113.
[27] L. B. WAHLBIN, *Superconvergence in Galerkin Finite Element Methods*, Lecture Notes in Math 1605, Springer, New York, 1995.
[28] K. WEISS AND L. DE FLORIANI, *Bisection-based triangulations of nested hypercubic meshes*, in Proceedings of the 19th International Meshing Roundtable, S. Shontz ed., Springer, Berlin, 2010, pp. 315–333.
[29] Z. ZHANG, *Ultraconvergence of the patch recovery technique* (II), Math. Comp., 69 (2000), pp. 141–158.
[30] Z. ZHANG, *Recovery technique in finite element methods*, in Adaptive Computations: Theory and Algorithms, T. Tang and J. Xu eds., Math. Monogr. Ser. 6, Science Press, Beijing, 2007, pp. 333–412.