East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.240313.280513a

Vol. **3**, No. 2, pp. 138-153
May 2013

Superconvergence of Fully Discrete Finite Elements for Parabolic Control Problems with Integral Constraints

Y. Tang* and Y. Hua

Department of Mathematics and Computational Science, Hunan University of Science and Engineering, Yongzhou 425100, Hunan, China.

Received 24 March 2013; Accepted (in revised version) 28 May 2013

Available online 31 May 2013

Abstract. A quadratic optimal control problem governed by parabolic equations with integral constraints is considered. A fully discrete finite element scheme is constructed for the optimal control problem, with finite elements for the spatial but the backward Euler method for the time discretisation. Some superconvergence results of the control, the state and the adjoint state are proved. Some numerical examples are performed to confirm theoretical results.

AMS subject classifications: 35B37, 49J20, 65N30

Key words: Superconvergence, finite element method, optimal control problems, parabolic equations, integral constraint.

1. Introduction

The Zienkiewicz-Zhu (ZZ) gradient patch recovery method based on local discrete least-squares fitting [21, 22] is now widely used in engineering practice, due to its robustness in a posteriori error estimates and efficiency in computer implementation. Superconvergence properties of the ZZ patch recovery method have been proven for both linear elements under strongly regular triangular meshes and all popular elements under a rectangular mesh [8, 19].

There has been extensive research on the superconvergence of finite element methods for optimal control problems, mostly focused on the elliptic case. The superconvergence properties of linear and semi-linear elliptic optimal control problems was established in Refs. [15] and [2] respectively, and for finite element approximations of bilinear elliptic optimal control problems [18]. Some superconvergence results for mixed finite element methods applied to elliptic optimal control problems have also been obtained [1, 3, 20].

*Corresponding author. *Email addresses:* tangyuelonga@163.com (Y. Tang), huayuchun306@sina.com (Y. Hua)

In recent years, there has been considerable related research for finite element approximations of parabolic optimal control problems that are frequently met in applications but are much more difficult to handle. A priori and a posteriori error estimates of finite element approximations for parabolic optimal control problems were derived in Refs. [5] and [17], respectively. A priori error estimates for the space-time finite element discretisation of parabolic optimal control problems have been obtained [13, 14], and a posteriori error estimation of spectral methods for parabolic optimal control problems were also investigated [4]. A variational discretisation method for optimal control involving the convection dominated diffusion equation has been considered [6], and superconvergence of a semi-discrete finite element method for parabolic optimal control problems was established [7], although this result has not been implemented in numerical computation. We have previously derived the superconvergence of finite element method for parabolic optimal control problems [16], and to the best of our knowledge there has been little work done on the superconvergence of fully discrete finite element methods for parabolic control problems. The purpose of this article is to investigate the superconvergence of fully discrete finite element approximation for parabolic optimal control problems with integral constraints.

We are interested in the following quadratic parabolic optimal control problem:

$$\begin{cases} \min\limits_{u\in K} \dfrac{1}{2}\displaystyle\int_0^T \left(\|y-y_d\|^2+\|u\|^2\right) dt\,, \\ y_t-\operatorname{div}(A\nabla y)=f+u\,, \qquad x\in\Omega\,,\ t\in J\,, \\ y|_{\partial\Omega}=0\,, \qquad t\in J\,, \\ y(0)=y_0, \qquad x\in\Omega\,, \end{cases} \tag{1.1}$$

where Ω is a bounded domain in $\mathbb{R}^2$ with a Lipschitz boundary $\partial\Omega$, and $J=[0,T]$ $(T>0)$. The coefficient $A=(a_{ij}(x))_{2\times 2}\in\left(W^{1,\infty}(\bar{\Omega})\right)^{2\times 2}$ is such that for any $\xi\in\mathbb{R}^2$ we have $(A(x)\xi)\cdot\xi\geq c\mid\xi\mid^2$ with $c>0$.

Let $f\in C(J;L^2(\Omega))$ and $y_0\in H_0^1(\Omega)$, and assume that K is a nonempty closed convex subset in $L^2(J;L^2(\Omega))$, defined by

$$K=\{\, v\in L^2(J;L^2(\Omega)):\int_0^T\int_\Omega v\,dx\,dt\geq 0\,\}\,.$$

We adopt the standard notation $W^{m,q}(\Omega)$ for Sobolev spaces on Ω with norm $\|\cdot\|_{W^{m,q}(\Omega)}$ and seminorm $|\cdot|_{W^{m,q}(\Omega)}$, set $H_0^1(\Omega)\equiv\left\{v\in H^1(\Omega): v|_{\partial\Omega}=0\right\}$, and denote $W^{m,2}(\Omega)$ by $H^m(\Omega)$. We also denote by $L^s(J;W^{m,q}(\Omega))$ the Banach space of all L^s integrable functions from J into $W^{m,q}(\Omega)$ with norm $\|v\|_{L^s(J;W^{m,q}(\Omega))}=(\int_0^T\|v\|^s_{W^{m,q}(\Omega)}dt)^{1/s}$ for $s\in[1,\infty)$ and the standard modification for $s=\infty$. Similarly, one can define the space $H^l(J;W^{m,q}(\Omega))$ (cf. Ref. [11]). In addition, c or C denotes a generic positive constant.

In Section 2, we define a fully discrete finite element approximation for the model problem, and introduce some intermediate variables and some useful error estimates in Section 3. We derive superconvergence properties in Section 4, and then present some numerical examples in the last section.

2. A Fully Discrete Finite Element Approximation

A fully discrete finite element approximation for the model problem (1.1) is now considered. For ease of exposition, we denote $L^p(J;W^{m,q}(\Omega))$ by $L^p(W^{m,q})$, and let $W = H_0^1(\Omega)$ and $U = L^2(\Omega)$. Moreover, we denote $\|\cdot\|_{H^m(\Omega)}$ and $\|\cdot\|_{L^2(\Omega)}$ by $\|\cdot\|_m$ and $\|\cdot\|$, respectively. Let

$$a(v,w) = \int_\Omega (A\nabla v)\cdot\nabla w\,, \qquad \forall\, v\,,\, w \in W\,,$$
$$(f_1,f_2) = \int_\Omega f_1\cdot f_2\,, \qquad \forall\, f_1, f_2 \in U\,.$$

It follows from the assumptions on A that

$$a(v,v) \ge c\|v\|_1^2\,, \quad |a(v,w)| \le C\|v\|_1\|w\|_1\,, \qquad \forall\, v,w \in W\,.$$

Thus a possible weak formulation for the model problem (1.1) is

$$\begin{cases} \min\limits_{u\in K} \dfrac{1}{2}\displaystyle\int_0^T \left(\|y-y_d\|^2 + \|u\|^2\right) dt\,, \\ (y_t,w) + a(y,w) = (f+u,w)\,, \qquad \forall\, w\in W,\ t\in J\,, \\ y(0) = y_0\,, \qquad x\in\Omega\,. \end{cases} \tag{2.1}$$

It is well known (e.g. see [10]) that the problem (2.1) has a unique solution (y,u), and the pair $(y,u)\in\left(H^1(L^2)\cap L^2(H^1)\right)\times K$ is the solution of the formulation (2.1) if and only if there is a adjoint state $p\in H^1(L^2)\cap L^2(H^1)$ such that the triplet (y,p,u) satisfies the following optimality conditions:

$$\begin{aligned} (y_t,w) + a(y,w) = (f+u,w)\,, &\qquad \forall\, w\in W,\ t\in J\,, \\ y(0) = y_0\,, &\qquad x\in\Omega\,, \end{aligned} \tag{2.2}$$

$$\begin{aligned} -(p_t,q) + a(q,p) = (y-y_d,q)\,, &\qquad \forall\, q\in W,\ t\in J\,, \\ p(T) = 0\,, &\qquad x\in\Omega\,, \end{aligned} \tag{2.3}$$

$$(u+p,v-u) \ge 0\,, \qquad \forall\, v\in K,\ t\in J\,. \tag{2.4}$$

We have the following Lemma.

Lemma 2.1 ([4]). *Let (y,p,u) be the solution of* (2.2)-(2.4). *Then* $u = \max(0,\overline{p}) - p$, *where*

$$\overline{p} = \frac{\int_0^T\int_\Omega p\,dx\,dt}{\int_0^T\int_\Omega 1\,dx\,dt}$$

denotes the integral average on $\Omega\times J$ of the function p.

Now let $\mathscr{T}^h$ be regular triangulations of Ω such that $\bar{\Omega} = \bigcup_{\tau\in\mathscr{T}^h}\bar{\tau}$ and $h = \max_{\tau\in\mathscr{T}^h}\{h_\tau\}$, where h_τ denotes the diameter of the element τ. Furthermore, set

$$\begin{aligned} U^h &= \left\{ v_h \in L^2(\Omega) : v_h|_\tau = \text{constant}\,,\ \forall\, \tau \in \mathscr{T}^h \right\}\,, \\ W^h &= \left\{ v_h \in C(\bar{\Omega}) : v_h|_\tau \in \mathbb{P}_1,\ \forall\, \tau \in \mathscr{T}^h,\, w_h|_{\partial\Omega} = 0 \right\}\,, \end{aligned}$$

where $\mathbb{P}_1$ is the space of polynomials up to order 1 and $K^h = \{v_h \in U^h : \int_\Omega v_h dx \geq 0\}$. Let $\Delta t > 0$, $N = T/\Delta t \in \mathbb{Z}^+$, and $t_n = n\Delta t$ for $n = 0, 1, \cdots, N$. Set $\varphi^n = \varphi(x, t_n)$ and

$$d_t\varphi^n = \frac{\varphi^n - \varphi^{n-1}}{\Delta t}\,,\ \text{for } n = 1, 2, \cdots, N\,.$$

Moreover, for $1 \leq p < \infty$ let us define the discrete time-dependent norms

$$|||\varphi|||_{l^p(J;W^{m,q}(\Omega))} := \left(\Delta t \sum_{n=1-l}^{N-l} \left\|\varphi^n\right\|^p_{W^{m,q}(\Omega)} \right)^{\frac{1}{p}}\,,$$

where $l = 0$ for the control u and state y, and $l = 1$ for the adjoint state p, with the standard modification for $p = \infty$. For convenience, we denote $|||\cdot|||_{l^s(J;W^{m,q}(\Omega))}$ by $|||\cdot|||_{l^s(W^{m,q})}$ and let

$$l^p_D(H^s) := \left\{ f : |||f|||_{l^p(H^s)} < \infty \right\}, \qquad 1 \leq p \leq \infty\,.$$

We also define the elliptic projection operator $R_h :\ W \to W^h$, such that for any $\phi \in W$ we have

$$a\left(\phi - R_h\phi, w_h\right) = 0\,, \qquad \forall\, w_h \in W^h\,, \tag{2.5}$$

and the approximation property

$$\|\phi - R_h\phi\| \leq Ch^2\|\phi\|_2\,, \qquad \forall\, \phi \in H^2(\Omega)\,. \tag{2.6}$$

Then a possible fully discrete finite element approximation of the weak formulation (2.1) is

$$\begin{cases} \min\limits_{u_h^n\in K^h} \dfrac{1}{2}\displaystyle\sum_{n=1}^{N} \Delta t \left(\|y_h^n - y_d^n\|^2 + \|u_h^n\|^2\right), \\ \left(d_t y_h^n, w_h\right) + a\left(y_h^n, w_h\right) = \left(f^n + u_h^n, w_h\right), \qquad \forall\, w_h \in W^h,\, n = 1, 2, \cdots, N\,, \\ y_h^0(x) = y_0^h(x)\,, \qquad x \in \Omega\,, \end{cases} \tag{2.7}$$

where $y_0^h(x) = R_h\left(y_0(x)\right)$ and R_h is the elliptic projection operator defined in Eq. (2.5).

It follows (e.g. see [12]) that the control problem (2.7) has a unique solution (y_h^n, u_h^n) for $n = 1, 2, \cdots, N$, and $(y_h^n, u_h^n) \in W^h \times K^h$ for $n = 1, 2, \cdots, N$ is the solution of (2.7) if and only if there is a adjoint state $p_h^{n-1} \in W^h$, $n = 1,2,\cdots,N$ such that the triplet $(y_h^n, p_h^{n-1}, u_h^n) \in W^h \times W^h \times K^h$, $n = 1,2,\cdots,N$ satisfies the optimality conditions

$$\begin{aligned} \left(d_t y_h^n, w_h\right) + a\left(y_h^n, w_h\right) = \left(f^n + u_h^n, w_h\right), \qquad &\forall w_h \in W^h, n = 1,2,\cdots,N, \\ y_h^0(x) = y_0^h(x), \qquad &x \in \Omega, \end{aligned} \tag{2.8}$$

$$\begin{aligned} -\left(d_t p_h^n, q_h\right) + a\left(q_h, p_h^{n-1}\right) = \left(y_h^n - y_d^n, q_h\right), \qquad &\forall q_h \in W^h, n = N,\cdots,2,1, \\ p_h^N(x) = 0, \qquad &x \in \Omega, \end{aligned} \tag{2.9}$$

$$\left(u_h^n + p_h^{n-1}, v - u_h^n\right) \geq 0, \qquad \forall v \in K^h, n = 1,2,\cdots,N. \tag{2.10}$$

3. Error Estimates of Intermediate Variables

We now give some error estimates of intermediate variables. For any $v \in K$, let $(y(v), p(v)) \in \left(H^1(L^2) \cap L^2(H^1)\right) \times \left(H^1(L^2) \cap L^2(H^1)\right)$ be the solution of the following equations:

$$\begin{aligned} \left(y_t(v), w\right) + a\left(y(v), w\right) = \left(f + v, w\right), \qquad &\forall w \in W, t \in J, \\ y(v)(0) = y_0, \qquad &x \in \Omega, \end{aligned} \tag{3.1}$$

$$\begin{aligned} -\left(p_t(v), q\right) + a\left(q, p(v)\right) = \left(y(v) - y_d, q\right), \quad &\forall q \in W, t \in J, \\ p(v)(T) = 0, \qquad &x \in \Omega. \end{aligned} \tag{3.2}$$

For any $v \in K$, any pair $\left(y_h^n(v), p_h^{n-1}(v)\right) \in W^h \times W^h$ for $n = 1,2,\cdots,N$ satisfies the following system:

$$\begin{aligned} \left(d_t y_h^n(v), w_h\right) + a\left(y_h^n(v), w_h\right) = \left(f^n + v^n, w_h\right), \quad &\forall w_h \in W^h, n = 1,2,\cdots,N, \\ y_h^0(v) = y_0^h, \quad &x \in \Omega, \end{aligned} \tag{3.3}$$

$$\begin{aligned} -\left(d_t p_h^n(v), q_h\right) + a\left(q_h, p_h^{n-1}(v)\right) = \left(y_h^n(v) - y_d^n, q_h\right), \; &\forall q_h \in W^h, n = N,\cdots,2,1, \\ p_h^N(v) = 0, \quad &x \in \Omega. \end{aligned} \tag{3.4}$$

Thus we have $(y, p) = (y(u), p(u))$ and $(y_h, p_h) = (y_h(u_h), p_h(u_h))$.

As in Ref. [15], we define the interpolation function $u_I(x,t) \in U^h$ for any $t \in J$ such that

$$u_I(x,t) = u(S_i, t), \qquad \forall x, S_i \in \tau_i, \tau_i \in \mathscr{T}^h, \tag{3.5}$$

where S_i is the centroid of the triangle τ_i.

Lemma 3.1 ([15]). *If $f \in H^2(\Omega)$, then*

$$\left|\int_{\tau_i} \left(f(x) - f(S_i)\right) dx\right| \le Ch^2 \sqrt{|\tau_i|} |f|_{H^2(\tau_i)} , \tag{3.6}$$

and

$$\sum_{\tau_i \in \mathcal{T}^h} \left|\int_{\tau_i} \left(f(x) - f(S_i)\right) dx\right| \le Ch^2 \left(\sum_{\tau_i \in \mathcal{T}^h} |f|^2_{H^2(\tau_i)}\right)^{\frac{1}{2}} . \tag{3.7}$$

Lemma 3.2. *Let $\left(y_h(u), p_h(u)\right)$ and $(y_h(u_I), p_h(u_I))$ be the discrete solutions of* (3.3)-(3.4) *with $v = u$ and $v = u_I$, respectively. Suppose that $u \in l^2_D(H^2)$. Then*

$$|||y_h(u_I) - y_h(u)|||_{l^2(H^1)} + |||p_h(u_I) - p_h(u)|||_{l^2(H^1)} \le Ch^2 |||u|||_{l^2(H^2)} . \tag{3.8}$$

Proof. Set $v = u_I$ and $v = u$ in (3.3), respectively. For $n = 1, 2, \cdots, N$ we obtain the error equation

$$\left(d_t y_h^n(u_I) - d_t y_h^n(u), w_h\right) + a\left(y_h^n(u_I) - y_h^n(u), w_h\right) = \left(u_I^n - u^n, w_h\right) , \ \forall w_h \in W^h. \tag{3.9}$$

We note that

$$\begin{aligned} &\left(d_t y_h^n(u_I) - d_t y_h^n(u), y_h^n(u_I) - y_h^n(u)\right) \\ \ge &\frac{1}{2\Delta t}\left(\left\|y_h^n(u_I) - y_h^n(u)\right\|^2 - \left\|y_h^{n-1}(u_I) - y_h^{n-1}(u)\right\|^2\right), \end{aligned} \tag{3.10}$$

and

$$\begin{aligned} \left(u_I^n - u^n, y_h^n(u_I) - y_h^n(u)\right) &\le C\left\|u_I^n - u^n\right\| \left\|y_h^n(u_I) - y_h^n(u)\right\| \\ &\le Ch^2 \|u^n\|_2 \left\|y_h^n(u_I) - y_h^n(u)\right\| \\ &\le C(\delta) h^4 \|u^n\|_2^2 + \delta \left\|y_h^n(u_I) - y_h^n(u)\right\|^2 . \end{aligned} \tag{3.11}$$

By choosing $w_h = y_h^n(u_I) - y_h^n(u)$ in (3.9) and multiplying both sides of (3.9) by $2\Delta t$, and then summing n from 1 to N, we get

$$\begin{aligned} &\left\|y_h^N(u_I) - y_h^N(u)\right\|^2 + c\sum_{n=1}^N \Delta t \left\|y_h^n(u_I) - y_h^n(u)\right\|_1^2 \\ \le &C(\delta) h^4 \sum_{n=1}^N \Delta t \|u^n\|_2^2 + \delta \sum_{n=1}^N \Delta t \left\|y_h^n(u_I) - y_h^n(u)\right\|_1^2 , \end{aligned} \tag{3.12}$$

and hence

$$|||y_h(u_I) - y_h(u)|||_{l^2(H^1)} \le Ch^2 |||u|||_{l^2(H^2)} . \tag{3.13}$$

On setting $v = u_I$ and $v = u$ in (3.4) respectively, we obtain the error equation

$$-\left(d_t p_h^n(u_I) - d_t p_h^n(u), q_h\right) + a\left(q_h, p_h^{n-1}(u_I) - p_h^{n-1}(u)\right) = \left(y_h^n(u_I) - y_h^n(u), q_h\right), \quad \forall q_h \in W^h, n = N, \cdots, 2, 1. \tag{3.14}$$

Similarly, we derive

$$|||p_h(u_I) - p_h(u)|||_{l^2(H^1)} \leq C|||y_h(u_I) - y_h(u)|||_{l^2(L^2)}, \tag{3.15}$$

so inequality (3.8) follows from (3.13) and (3.15). □

Lemma 3.3. *For any $v \in K$, let $(y(v), p(v))$ and $(y_h(v), p_h(v))$ be the solutions of* (3.1)-(3.2) *and* (3.3)-(3.4)*, respectively. Assume that $y(v), p(v) \in l_D^2(H^2) \cap H^1(H^2) \cap H^2(L^2)$ and $y_d \in H^1(L^2)$. Then*

$$|||R_h y(v) - y_h(v)|||_{l^2(H^1)} + |||R_h p(v) - p_h(v)|||_{l^2(H^1)} \leq C\left(h^2 + \Delta t\right). \tag{3.16}$$

Proof. From (3.1) and (3.3), we obtain

$$\left(y_t^n(v) - d_t y_h^n(v), w_h\right) + a\left(y^n(v) - y_h^n(v), w_h\right) = 0, \ \forall w_h \in W^h, \ n = 1, 2, \cdots, N. \tag{3.17}$$

From the definition of R_h,

$$\begin{aligned} &\left(d_t R_h y^n(v) - d_t y_h^n(v), w_h\right) + a\left(R_h y^n(v) - y_h^n(v), w_h\right) \\ =&\left(d_t R_h y^n(v) - d_t y^n(v) + d_t y^n(v) - y_t^n(v), w_h\right), \end{aligned} \tag{3.18}$$

and we note that

$$\begin{aligned} &\left(d_t R_h y^n(v) - d_t y^n(v), R_h y^n(v) - y_h^n(v)\right) \\ \leq& \left\|d_t R_h y^n(v) - d_t y^n(v)\right\| \left\|R_h y^n(v) - y_h^n(v)\right\| \\ \leq& Ch^2 \left\|d_t y^n(v)\right\|_2 \left\|R_h y^n(v) - y_h^n(v)\right\| \\ \leq& Ch^2(\Delta t)^{-1} \int_{t_{n-1}}^{t_n} \left\|y_t(v)\right\|_2 dt \left\|R_h y^n(v) - y_h^n(v)\right\| \\ \leq& Ch^2(\Delta t)^{-\frac{1}{2}} ||y_t(v)||_{L^2(t_{n-1}, t_n; H^2(\Omega))} \left\|R_h y^n(v) - y_h^n(v)\right\|, \end{aligned} \tag{3.19}$$

and

$$\begin{aligned} &\left(d_t y^n(v) - y_t^n(v), R_h y^n(v) - y_h^n(v)\right) \\ =& (\Delta t)^{-1}\left(y^n(v) - y^{n-1}(v) - \Delta t y_t^n(v), R_h y^n(v) - y_h^n(v)\right) \\ \leq& (\Delta t)^{-1}\left\|y^n(v) - y^{n-1}(v) - \Delta t y_t^n(v)\right\| \left\|R_h y^n(v) - y_h^n(v)\right\| \\ =& (\Delta t)^{-1}\left\|\int_{t_{n-1}}^{t_n} (t_{n-1} - s)(y_{tt}(v))(s) ds\right\| \left\|R_h y^n(v) - y_h^n(v)\right\| \\ \leq& C(\Delta t)^{\frac{1}{2}} ||y_{tt}(v)||_{L^2(t_{n-1}, t_n; L^2(\Omega))} \left\|R_h y^n(v) - y_h^n(v)\right\|. \end{aligned} \tag{3.20}$$

Similarly to Lemma 3.2, from (3.18)-(3.20) and Young's inequality we have

$$\begin{aligned}&\left\|R_h y^N(v)-y_h^N(v)\right\|^2+c\sum_{n=1}^{N}\Delta t\left\|R_h y^n(v)-y_h^n(v)\right\|_1^2\\ \le\, & C(\delta)\left(h^4\|y_t(v)\|^2_{L^2(H^2)}+(\Delta t)^2\|y_{tt}(v)\|^2_{L^2(L^2)}\right)\\ &+\delta\sum_{n=1}^{N}\Delta t\left\|R_h y^n(v)-y_h^n(v)\right\|^2\ ,\end{aligned}\tag{3.21}$$

whence

$$|||R_h y(v)-y_h(v)|||_{l^2(H^1)}\le C\left(h^2\|y_t(v)\|_{L^2(H^2)}+\Delta t\|y_{tt}(v)\|_{L^2(L^2)}\right).\tag{3.22}$$

From (3.2) and (3.4), for any $q_h\in W^h$, $n=N,\cdots,2,1$ we obtain

$$\begin{aligned}&-\left(p_t^{n-1}(v)-d_t p_h^n(v),q_h\right)+a\left(q_h,p^{n-1}(v)-p_h^{n-1}(v)\right)\\ =&\left(y^{n-1}(v)-y_h^n(v)-y_d^{n-1}+y_d^n,q_h\right)\ ,\end{aligned}\tag{3.23}$$

and from the definition of R_h we derive

$$\begin{aligned}&-\left(d_t R_h p^n(v)-d_t p_h^n(v),q_h\right)+a\left(q_h,R_h p^{n-1}(v)-p_h^{n-1}(v)\right)\\ =&\left(-d_t R_h p^n(v)+p_t^{n-1}(v)+y^{n-1}(v)-y_h^n(v)+y_d^n-y_d^{n-1},w_h\right)\ .\end{aligned}\tag{3.24}$$

Similarly, we can prove that

$$\begin{aligned}&|||R_h p(v)-p_h(v)|||^2_{l^2(H^1)}\\ \le\, & C(\delta)\left(|||R_h y(v)-y_h(v)|||^2_{l^2(H^1)}+h^4|||y(v)|||^2_{l^2(H^2)}+h^4\|p_t(v)\|^2_{L^2(H^2)}\right)\\ &+C(\delta)(\Delta t)^2\left(|||p(v)|||^2_{l^2(H^1)}+\|p_{tt}(v)\|^2_{L^2(L^2)}\right)\\ &+C(\delta)(\Delta t)^2\left(\|y_t(v)\|^2_{L^2(L^2)}+\|(y_d)_t\|^2_{L^2(L^2)}\right)\ .\end{aligned}\tag{3.25}$$

From (3.22) and (3.25), we then obtain (3.16). □

4. Superconvergence Analysis

We now discuss superconvergence properties between the finite element solution and projections of the exact solution, and begin by deriving the superconvergence of the control variable.

Theorem 4.1. *Let u and u_h be the solutions of* (2.2)-(2.4) *and* (2.8)-(2.10), *respectively. Assume that all of the conditions in* Lemmas 3.1-3.3 *are valid. Then*

$$|||u_I-u_h|||_{l^2(L^2)}\le C\left(h^2+\Delta t\right)\ .\tag{4.1}$$

Proof. As in Ref. [15], since $u_h \in K$ it follows from (2.4) that we have point-wise almost everywhere

$$(u(x,t)+p(x,t))(u_h(x,t)-u(x,t)) \geq 0\,, \qquad \forall\, t \in J\,. \tag{4.2}$$

On applying this formula for $x = S_i$, it follows that

$$(u(S_i,t)+p(S_i,t))(u_h(S_i,t)-u(S_i,t)) \geq 0\,, \quad \forall\, S_i \in \tau_i, \tau_i \in \mathcal{T}^h \text{ and } \forall\, t \in J$$

i.e.,

$$(u_I^n + p_I^n)(u_h^n - u_I^n) \geq 0\,, \quad \forall\, S_i \in \tau_i, \tau_i \in \mathcal{T}^h \text{ and } \forall\, n = 1,2,\cdots,N\,, \tag{4.3}$$

because of the continuity of u^n, p^n, and u_h^n at these points, and hence

$$(u_I^n + p_I^n, u_h^n - u_I^n) \geq 0\,, \qquad n = 1,2,\cdots,N\,. \tag{4.4}$$

On choosing $v = u_I$ in (2.10),

$$(u_h^n + p_h^{n-1}, u_I^n - u_h^n) \geq 0\,, \qquad n = 1,2,\cdots,N\,, \tag{4.5}$$

hence

$$(u_h^n - u_I^n + p_h^{n-1} - p_I^n, u_I^n - u_h^n) \geq 0\,, \qquad n = 1,2,\cdots,N\,. \tag{4.6}$$

Moreover, we obtain

$$\begin{aligned}
&|||u_I - u_h|||^2_{l^2(L^2)} \\
\leq\;& \Delta t \sum_{n=1}^{N} \left(p_h^{n-1} - p_I^n, u_I^n - u_h^n\right) \\
=\;& \Delta t \sum_{n=1}^{N} \left(p_h^{n-1}(u_h) - p_h^{n-1}(u_I), u_I^n - u_h^n\right) + \Delta t \sum_{n=1}^{N} \left(p_h^{n-1}(u_I) - p_h^{n-1}(u), u_I^n - u_h^n\right) \\
&+ \Delta t \sum_{n=1}^{N} \left(p_h^{n-1}(u) - p^{n-1}(u), u_I^n - u_h^n\right) + \Delta t \sum_{n=1}^{N} \left(p^{n-1}(u) - p_I^n, u_I^n - u_h^n\right) \\
:=\;& I_1 + I_2 + I_3 + I_4\,.
\end{aligned} \tag{4.7}$$

According to (3.3)-(3.4), for the first term we have

$$\begin{aligned}
&\left(d_t y_h^n(u_h) - d_t y_h^n(u_I), w_h\right) + a\left(y_h^n(u_h) - y_h^n(u_I), w_h\right) = \left(u_h^n - u_I^n, w_h\right), \\
&\qquad \forall\, w_h \in W^h, \quad n = 1,2,\cdots,N\,, \\
&\qquad y_h^0(u_h) = y_h^0(u_I) = y_0^h\,, \quad x \in \Omega\,,
\end{aligned} \tag{4.8}$$

and

$$\begin{aligned}
&-\left(d_t p_h^n(u_h) - d_t p_h^n(u_I), q_h\right) + a\left(q_h, p_h^{n-1}(u_h) - p_h^{n-1}(u_I)\right) = \left(y_h^n(u_h) - y_h^n(u_I), q_h\right), \\
&\qquad \forall\, q_h \in W^h, \quad n = N,\cdots,2,1\,, \\
&\qquad p_h^N(u_h) = p_h^N(u_I) = 0\,, \quad x \in \Omega\,.
\end{aligned} \tag{4.9}$$

On selecting $w_h = p_h^{n-1}(u_h) - p_h^{n-1}(u_I)$ in (4.8) and $q_h = y_h^n(u_h) - y_h^n(u_I)$ in (4.9),

$$I_1 = -|||y_h(u_I) - y_h(u_h)|||^2_{l^2(L^2)} \le 0\,, \tag{4.10}$$

and from Lemma 3.2 we have

$$I_2 = \Delta t \sum_{i=1}^{N} \left(p_h^{n-1}(u_I) - p_h^{n-1}(u), u_I^n - u_h^n\right) \le C(\delta)h^4 + \delta|||u_I - u_h|||^2_{l^2(L^2)}\,. \tag{4.11}$$

From (2.6) and Lemma 3.3, we derive

$$\begin{aligned} I_3 &= \Delta t \sum_{i=1}^{N} \left(p_h^{n-1}(u) - p^{n-1}(u), u_I^n - u_h^n\right) \\ &\le \Delta t \sum_{i=1}^{N} \left(p_h^{n-1}(u) - R_h p^{n-1}(u), u_I^n - u_h^n\right) + \Delta t \sum_{i=1}^{N} \left(R_h p^{n-1}(u) - p^{n-1}(u), u_I^n - u_h^n\right) \\ &\le C(\delta)\left(h^4 + (\Delta t)^2\right) + \delta|||u_I - u_h|||^2_{l^2(L^2)}\,. \end{aligned} \tag{4.12}$$

For the last term, from Lemma 3.1 we have

$$\begin{aligned} I_4 &= \Delta t \sum_{i=1}^{N} \left(p^{n-1}(u) - p_I^n, u_I^n - u_h^n\right) \\ &= \Delta t \sum_{i=1}^{N} \left(p^{n-1}(u) - p^n(u), u_I^n - u_h^n\right) + \Delta t \sum_{i=1}^{N} \left(p^n(u) - p_I^n, u_I^n - u_h^n\right) \\ &\le C(\delta)\left(h^4 + (\Delta t)^2\right) + \delta|||u_I - u_h|||^2_{l^2(L^2)}\,, \end{aligned} \tag{4.13}$$

and finally inequality (4.1) follows from (4.7)-(4.13). □

Secondly, we consider the superconvergence of the state and the adjoint state.

Theorem 4.2. *Let* (y,p,u) *and* (y_h,p_h,u_h) *be the solutions* (2.2)-(2.4) *and* (2.8)-(2.10) *respectively, and assume that all of the conditions in* Theorem 4.1 *are valid. Then*

$$|||R_h y - y_h|||_{l^2(H^1)} + |||R_h p - p_h|||_{l^2(H^1)} \le C\left(h^2 + \Delta t\right). \tag{4.14}$$

Proof. From (2.2) and (2.8), we have the error equation

$$\left(y_t^n - d_t y_h^n, w_h\right) + a\left(y^n - y_h^n, w_h\right) = \left(u^n - u_h^n, w_h\right), \forall\, w_h \in W^h,\ n = 1,2,\cdots,N. \tag{4.15}$$

From the definition of R_h,

$$\begin{aligned} &\left(d_t R_h y^n - d_t y_h^n, w_h\right) + a\left(R_h y^n - y_h^n, w_h\right) \\ =&\left(d_t R_h y^n - d_t y^n + d_t y^n - y_t^n + u^n - u_I^n + u_I^n - u_h^n, w_h\right),\ \forall\, w_h \in W^h, n = 1,2,\cdots,N. \end{aligned} \tag{4.16}$$

Let us note that

$$\begin{aligned}\left(u^n - u_I^n, R_h y^n - y_h^n\right) &\le C \left\|u^n - u_I^n\right\| \left\|R_h y^n - y_h^n\right\| \\ &\le C(\delta) \left\|u^n - u_I^n\right\|^2 + \delta \left\|R_h y^n - y_h^n\right\|^2\end{aligned} \tag{4.17}$$

and

$$\begin{aligned}\left(u_I^n - u_h^n, R_h y^n - y_h^n\right) &\le C \left\|u_I^n - u_h^n\right\| \left\|R_h y^n - y_h^n\right\| \\ &\le C(\delta) \left\|u_I^n - u_h^n\right\|^2 + \delta \left\|R_h y^n - y_h^n\right\|^2 .\end{aligned} \tag{4.18}$$

Similarly to Lemma 3.2, from (4.16)-(4.18), Lemma 3.1 and Theorem 4.1 we have

$$\begin{aligned}&|||R_h y - y_h|||_{l^2(H^1)} \\ &\le C(\delta) h^2 \left(\left\|y_t\right\|_{L^2(H^2)} + |||p_h|||_{l^2(H^1)}\right) + C(\delta)\Delta t \left(\left\|y_{tt}\right\|_{L^2(L^2)} + \left\|p_t\right\|_{L^2(L^2)}\right) .\end{aligned} \tag{4.19}$$

From (2.3) and (2.9), for $n = N, \cdots, 2, 1$ we obtain

$$\begin{aligned}-\left(p_t^{n-1} - d_t p_h^n, q_h\right) + a\left(q_h, p^{n-1} - p_h^{n-1}\right) = \left(y^{n-1} - y_h^n - y_d^{n-1} + y_d^n, q_h\right) , \\ \forall\, q_h \in W^h.\end{aligned} \tag{4.20}$$

By using the definition of R_h, for any $q_h \in W^h$, $n = N, \cdots, 2, 1$ we get

$$\begin{aligned}&-\left(d_t R_h p^n - d_t p_h^n, q_h\right) + a\left(q_h, R_h p^{n-1} - p_h^{n-1}\right) \\ &= \left(-d_t R_h p^n + p_t^n + y^{n-1} - y_h^n - y_d^{n-1} + y_d^n, q_h\right).\end{aligned} \tag{4.21}$$

Similarly, we can prove that

$$\begin{aligned}&|||R_h p - p_h|||_{l^2(H^1)}^2 \\ &\le C(\delta)\left(|||R_h y - y_h|||_{l^2(H^1)}^2 + h^4 |||y|||_{l^2(H^2)}^2 + h^4 ||p_t||_{L^2(H^2)}^2\right) \\ &\quad + C(\delta)(\Delta t)^2 \left(|||p|||_{l^2(H^1)}^2 + ||p_{tt}||_{L^2(L^2)}^2 + ||y_t||_{L^2(L^2)}^2 + ||(y_d)_t||_{L^2(L^2)}^2\right) ,\end{aligned} \tag{4.22}$$

whence (4.14) from (4.19) and (4.22). □

5. Numerical Experiments

For a constrained parabolic optimization problem:

$$\min_{u \in K} J(u),$$

where $J(u)$ is a convex functional on X and K is a close convex subset of X, the iterative scheme reads ($n = 0, 1, 2, \cdots$):

$$\begin{cases} b(u_{n+\frac{1}{2}}, v) = b(u_n, v) - \rho_n \left(J'(u_n), v\right), \quad \forall\, v \in X , \\ u_{n+1} = P_K^b(u_{n+\frac{1}{2}}) , \end{cases} \tag{5.1}$$

where $b(\cdot,\cdot)=\int_0^T(\cdot,\cdot)$ is a symmetric positive definite bilinear form, ρ_n is the step size of the iteration, and the projection operator P_K^b can be computed as in Ref. [9].

Similar to Ref. [9], for an acceptable error Tol we present the following projection gradient algorithm, by applying (5.1) to the discretised parabolic optimal control problem (2.7). For ease of exposition, we have omitted the subscript h.

Algorithm 5.1. Projection gradient algorithm

Step 1. Initialize u_0.
Step 2. Solve the following equations:

$$\begin{cases} b(u_{n+\frac{1}{2}},v)=b(u_n,v)-\rho_n\displaystyle\int_0^T\left(u_n+p_n,v\right), & u_{n+\frac{1}{2}},u_n\in U^h,\forall\, v\in U^h,\\ \left(\dfrac{y_n^i-y_n^{i-1}}{\Delta t},w\right)+a\left(y_n^i,w\right)=\left(f^i+u_n^i,w\right), & y_n^i,y_n^{i-1}\in W^h,\forall\, w\in W^h,\\ \left(\dfrac{p_n^{i-1}-p_n^i}{\Delta t},q\right)+a\left(q,p_n^{i-1}\right)=\left(y_n^i-y_d^i,q\right), & p_n^i,p_n^{i-1}\in W^h,\forall\, q\in W^h\,,\\ u_{n+1}=P_K^b(u_{n+\frac{1}{2}})\,. \end{cases} \tag{5.2}$$

Step 3. Calculate the iterative error: $E_{n+1}=|||u_{n+1}-u_n|||_{l^2(L^2)}$.
Step 4. If $E_{n+1}\le Tol$, stop; else go to Step 2.

Let $\Omega=[0,1]\times[0,1]$, $T=1$, and $A(x)=I$. The optimal control problem was solved numerically with codes developed based on AFEPack, a package that is freely available (cf. Ref. [9]). The discretisation was as described in Section 2. We denote $|||\cdot|||_{l^2(H^1)}$ and $|||\cdot|||_{l^2(L^2)}$ by $|||\cdot|||_1$ and $|||\cdot|||$, respectively. The convergence order rate is computed from the formula

$$Rate=\frac{log(e_{i+1})-log(e_i)}{log(h_{i+1})-log(h_i)}\,,$$

where e_i (e_{i+1}) denotes the error when the spatial partition size is h_i (h_{i+1}). We solved the parabolic optimal control problem

$$\begin{cases} \min\limits_{u\in K}\dfrac{1}{2}\displaystyle\int_0^T\left(\|y(x,t)-y_d(x,t)\|^2+\|u(x,t)\|^2\right)dt\,,\\ y_t(x,t)-\mathrm{div}(A(x)\nabla y(x,t))=f(x,t)+u(x,t)\text{ in }\Omega\times(0,T]\,,\\ y(x,t)=0\text{ on }\partial\Omega\times(0,T]\,,\\ y(x,0)=y_0(x)\text{ in }\Omega\,, \end{cases} \tag{5.3}$$

for the following two examples.

Example 5.1. The data are as follows:

$$\begin{aligned}
p(x,t) &= \sin(2\pi x_1)\sin(2\pi x_2)(1-t)\,,\\
y(x,t) &= \sin(2\pi x_1)\sin(2\pi x_2)t\,,\\
u(x,t) &= \max(0,\overline{p(x,t)}) - p(x,t)\,,\\
f(x,t) &= y_t(x,t) - \operatorname{div}(A(x)\nabla y(x,t)) - u(x,t)\,,\\
y_d(x,t) &= y(x,t) + p_t(x,t) + \operatorname{div}(A^*(x)\nabla p(x,t))\,.
\end{aligned}$$

The errors on a sequence of uniformly refined meshes are shown in Table 1. In Fig. 1, we show the numerical solution u_h at $t=0.5$ when $h=1.25\times10^{-2}$ and $\Delta t = 6.25\times10^{-3}$. It is easy to see $|||u_I-u_h||| = \mathscr{O}\left(h^2+\Delta t\right)$, $|||R_hy-y_h|||_1 = \mathscr{O}\left(h^2+\Delta t\right)$ and $|||R_hp-p_h|||_1 = \mathscr{O}\left(h^2+\Delta t\right)$.

Table 1: Numerical results, Example 5.1.

h	Δt	$\|\|\|u_I-u_h\|\|\|$	Rate	$\|\|\|R_hy-y_h\|\|\|_1$	Rate	$\|\|\|R_hp-p_h\|\|\|_1$	Rate
$\frac{1}{10}$	$\frac{1}{10}$	4.57845e-02	—	3.31962e-02	—	3.31860e-02	—
$\frac{1}{20}$	$\frac{1}{40}$	1.22388e-02	1.90	8.22965e-03	2.01	8.22864e-03	2.01
$\frac{1}{40}$	$\frac{1}{160}$	3.09192e-03	1.98	2.05253e-03	2.00	2.05242e-03	2.00
$\frac{1}{80}$	$\frac{1}{640}$	7.78481e-04	1.98	5.16524e-04	1.99	5.18543e-04	1.98

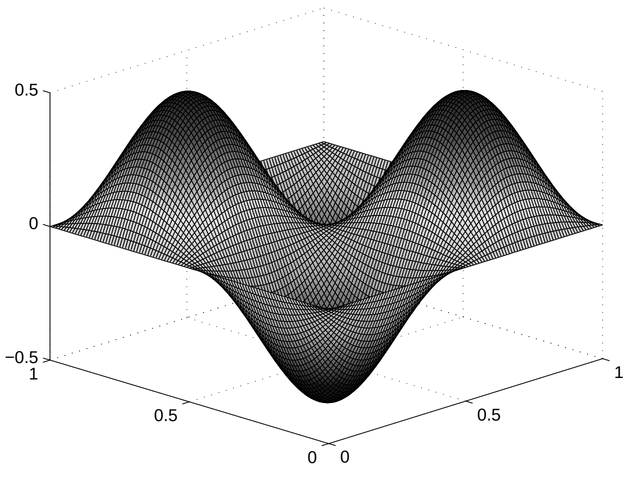

Figure 1: The numerical solution u_h at $t=0.5$ for Example 5.1.

Example 5.2. The data are as follows:

$$\begin{aligned}
p(x,t) &= \sin(2\pi x_1)\sin(2\pi x_2)\sin(\pi t)\,,\\
y(x,t) &= \sin(2\pi x_1)\sin(2\pi x_2)\sin(\pi t)\,,\\
u(x,t) &= \max(0,\overline{p(x,t)}) - p(x,t)\,,\\
f(x,t) &= y_t(x,t) - \mathrm{div}(A(x)\nabla y(x,t)) - u(x,t)\,,\\
y_d(x,t) &= y(x,t) + p_t(x,t) + \mathrm{div}(A^*(x)\nabla p(x,t))\,.
\end{aligned}$$

In Table 2, the errors $|||u_I - u_h|||$, $|||R_h y - y_h|||_1$ and $|||R_h p - p_h|||_1$ based on a sequence of uniformly refined meshes are shown. We plot the profile of the numerical solution u_h at $t = 0.5$ when $h = 1.25 \times 10^{-2}$ and $\Delta t = 6.25 \times 10^{-3}$ in Fig. 2. It is easy to see that $|||u_I - u_h|||$, $|||R_h y - y_h|||_1$ and $|||R_h p - p_h|||_1$ are the second order convergent which is consistent with our theoretical results.

Table 2: Numerical results for Example 5.2.

h	Δt	$\|\|\|u-u_h\|\|\|$	Rate	$\|\|\|R_h y - y_h\|\|\|_1$	Rate	$\|\|\|R_h p - p_h\|\|\|_1$	Rate
$\frac{1}{10}$	$\frac{1}{10}$	4.97429e-02	—	8.36251e-03	—	7.42735e-03	—
$\frac{1}{20}$	$\frac{1}{40}$	1.31232e-02	1.92	2.30218e-03	1.86	2.00054e-03	1.89
$\frac{1}{40}$	$\frac{1}{160}$	3.31625e-03	1.98	5.91429e-04	1.96	5.10980e-04	1.97
$\frac{1}{80}$	$\frac{1}{640}$	8.31263e-04	2.00	1.48953e-04	1.99	1.28457e-04	1.99

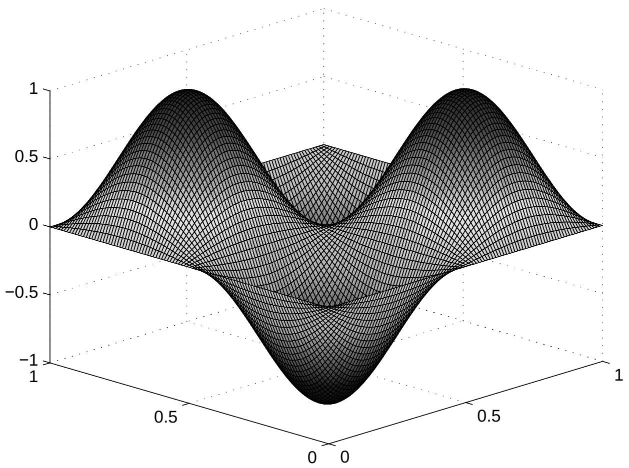

Figure 2: The numerical solution u_h at $t = 0.5$ for Example 5.2.

Acknowledgments

This work is supported by the Hunan Provincial Innovation Foundation for Postgraduate Studies (Grant lx2009B120), and the TianYuan Special Funds of the National Natural Science Foundation of China (Grant No. 11226313). The authors would like to thank the referees for their helpful suggestions, which led to improvements in the presentation.

References

[1] Y. Chen, Superconvergence of mixed finite element methods for optimal control problems, Math. Comp. 77 (2008) 1269-1291.

[2] Y. Chen and Y. Dai, Superconvergence for optimal control problems governed by semi-linear elliptic equations, J. Sci. Comput. 39 (2009) 206-221.

[3] Y. Chen, Y. Huang, W. Liu and N. Yan, Error estimates and superconvergence of mixed finite element methods for convex optimal control problems, J. Sci. Comput. 42 (2010) 382-403.

[4] Y. Chen, Y. Huang and N. Yi, A posteriori error estimates of spectral method for optimal control problems governed by parabolic equations, Sci. China Seri. A: Math. 51(8) (2008) 1376-1390.

[5] H. Fu and H. Rui, A priori error estimates for optimal control problems governed by transient advection-diffusion equations, J. Sci. Comput. 38 (2009) 290-315.

[6] M. Hinze, N. Yan and Z. Zhou, Variational discretisation for optimal control governed by convection dominated diffusion equations, J. Comput. Math. 27 (2009) 237-253.

[7] C. Hou, Y. Chen and Z. Lu, Superconvergece property of finite element methods for parabolic optimal control problems, J. Ind. Manag. Optim. 7(4) (2011) 927-945.

[8] B. Li and Z. Zhang, Analysis of a class of superconvergence patch recovery techniques for linear and bilinear finite elements, Numer. Meth. PDE. 15 (1999) 151-167.

[9] R. Li, W. Liu and N. Yan, A posteriori error estimates of recovery type for distributed convex optimal control problems, J. Sci. Comput. 33 (2007) 155-182.

[10] J. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer-Verlag, Berlin, 1971.

[11] J. Lions and E. Magenes, Non Homogeneous Boundary Value Problems and Applications, Springer-Verlag, Berlin, 1972.

[12] W. Liu and N. Yan, A posteriori error estimates for optimal control problems governed by parabolic equations, Numer. Math. 93 (2003) 497-521.

[13] D. Meidner and B. Vexler, A priori error estimates for space-time finite element discretisation of parabolic optimal control problems Part I: Problems without control constraints, SIAM J. Control Optim. 47(3) (2008) 1150-1177.

[14] D. Meidner and B. Vexler, A priori error estimates for space-time finite element discretisation of parabolic optimal control problems Part II: Problems with control constraints, SIAM J. Control Optim. 47(3) (2008) 1301-1329.

[15] C. Meyer and A. Rösch, Superconvergence properties of optimal control problems, SIAM J. Control Optim. 43(3) (2004) 970-985.

[16] Y. Tang and Y. Chen, Superconvergence analysis of fully discrete finite element methods for semilinear parabolic optimal control problems, Front. Math. China 8(2) (2013) 443-464.

[17] C. Xiong and Y. Li, A posteriori error estimates for optimal distributed control governed by the evolution equations, Appl. Numer. Math. 61 (2011) 181-200.

[18] D. Yang, Y. Chang and W. Liu, A priori error estimate and superconvergence annalysis for an opitmal control problem of bilinear type, J. Comput. Math. 26(4) (2008) 471-487.
[19] Z. Zhang, Ultraconvergence of the patch recovery technique II, Math. Comp. 69 (2000) 141-158.
[20] J. Zhou, Y. Chen and Y. Dai, Superconvergence of triangular mixed finite elements for optimal control problems with an integral constraint, Appl. Math. Comput. 217 (2010) 2057-2066.
[21] O.C. Zienkiwicz and J.Z. Zhu, The superconvergence patch recovery and a posteriori error estimates, Int. J. Numer. Meth. Eng. 33 (1992) 1331-1382.
[22] O.C. Zienkiwicz and J.Z. Zhu, The superconvergence patch recovery (SPR) and adaptive finite element refinement, Comput. Meth. Appl. Math. 101 (1992) 207-224.