East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.010314.110115a

Vol. **5**, No. 1, pp. 85-108
February 2015

A Posteriori Error Estimates of Semidiscrete Mixed Finite Element Methods for Parabolic Optimal Control Problems

Yanping Chen* and Zhuoqing Lin

School of Mathematical Sciences, South China Normal University, Guangzhou 510631, Guangdong, P.R. China.

Received 1 March 2014; Accepted (in revised version) 11 January 2015

Abstract. *A posteriori* error estimates of semidiscrete mixed finite element methods for quadratic optimal control problems involving linear parabolic equations are developed. The state and co-state are discretised by Raviart-Thomas mixed finite element spaces of order k, and the control is approximated by piecewise polynomials of order k ($k \geq 0$). We derive our *a posteriori* error estimates for the state and the control approximations via a mixed elliptic reconstruction method. These estimates seem to be unavailable elsewhere in the literature, although they represent an important step towards developing reliable adaptive mixed finite element approximation schemes for the control problem.

AMS subject classifications: 49J20, 65N30

Key words: A posteriori error estimates, optimal control problems, parabolic equations, elliptic reconstruction, semidiscrete mixed finite element methods.

1. Introduction

Optimal control problems (OCP) involving partial differential equations (PDE) arise in various fields such as fluid dynamics, environmental modelling and engineering. Efficient numerical methods are usually required to solve such OCP. Finite element approximations have proven suitable in engineering design work [4, 18–20, 23, 26, 35, 37], and very many authors have considered OCP governed by elliptic or parabolic state equations previously [1, 17, 22, 25, 27–30, 32, 36]. There has been a growing demand for reliable and efficient space-time algorithms to solve both linear and nonlinear time-dependent PDE numerically, and most of the algorithms are based on *a posteriori* error estimators in order to provide appropriate tools for adaptive mesh refinements. The theory for the *a posteriori* analysis of finite element methods for elliptic problems is well developed, but it is yet to be as complete for time-dependent linear and nonlinear problems. For parabolic problems, there are schemes dealing with space-time adaptivity [14, 15, 38, 39] or with time

*Corresponding author. *Email addresses:* yanpingchen@scnu.edu.cn (Y. Chen), 0626lzq@sina.com (Z. Lin)

adaptivity alone [21], or with spatial adaptivity on keeping the temporal variable continuous [3, 4]. For space-time adaptivity, typically the finite element discretization depends upon a space-time variational formulation and the error indicators include both space and time errors. Makridakis & Nochetto [33] introduced an elliptic reconstruction operator that plays a role in *a posteriori* estimates quite similar to the role played by an elliptic projection for recovering optimal *a priori* error estimates for parabolic problems [41]. This elliptic construction method was developed for a completely discrete scheme based on the backward Euler method [31], for maximum norm estimates [13], and for discontinuous Galerkin methods for parabolic problems [16].

In many control problems, the objective functional contains the gradient of the state variables. For example, in flow control problems the gradient representing the Darcy velocity is an important variable, and in temperature control problems large temperature gradients during cooling or heating are important as they may be quite destructive. The accuracy of the gradient is therefore important in the numerical discretization of the coupled state equations. Mixed finite element methods are appropriate for the state equations in such cases, since both the scalar variable and its flux variable can be approximated to the same accuracy — e.g. see Ref. [6]. Indeed, when the objective functional contains the gradient of the state variable, mixed finite element methods can be used for the state equation such that both the scalar variable and its flux variable can be approximated to the same accuracy. Recently, we have worked on both *a priori* superconvergence and *a posteriori* error estimates in the application of mixed finite element methods to linear elliptic OCP [7, 8, 10].

In this article, we develop *a posteriori* error estimates of a semidiscrete mixed finite element approximation for parabolic OCP. Combining the elliptic construction idea of Ref. [33] with the parabolic OCP, we define a corresponding mixed elliptic construction for the state and co-state variables, and then use this mixed elliptic construction to derive *a posteriori* error estimates for both the state and the control approximation. The OCP of interest are of the form

$$\min_{u\in K\subset U}\left\{\frac{1}{2}\int_0^T(\|\boldsymbol{p}-\boldsymbol{p}_d\|^2+\|y-y_d\|^2+\|u\|^2)\,dt\right\}, \tag{1.1}$$

$$y_t(x,t)+\mathrm{div}\boldsymbol{p}(x,t)=f(x,t)+u(x,t), \qquad x\in\Omega,\ t\in J, \tag{1.2}$$

$$\boldsymbol{p}(x,t)=-A(x)\nabla y(x,t), \qquad x\in\Omega,\ t\in J, \tag{1.3}$$

$$y(x,t)=0, \qquad x\in\partial\Omega,\ t\in J, \tag{1.4}$$

$$y(x,0)=y_0(x), \qquad x\in\Omega, \tag{1.5}$$

where the bounded open set $\Omega\subset R^2$ is a convex polygon with boundary $\partial\Omega$ and $J=[0,T]$.

Let K be a closed convex set in control space $U=L^2(0,T;L^2(\Omega))$, $f, y_d\in L^2(0,T;L^2(\Omega))$, $\boldsymbol{p}_d\in(L^2(0,T;L^2(\Omega)))^2$ and $y_0\in H_0^1(\Omega)$. We assume the coefficient matrix

$$A(x)=\big(a_{ij}(x)\big)_{2\times2}\in W^{1,\infty}(\bar\Omega;R^{2\times2})$$

is symmetric 2×2, that there are constants $c_1,c_2>0$ for any vector $X\in R^2$ such that

$c_1\|X\|_{R^2}^2 \le X^tAX \le c_2\|X\|_{R^2}^2$, and K is a set defined by

$$K = \left\{ u \in U : \int_0^T \int_\Omega u\,dx\,dt \ge 0 \right\} . \tag{1.6}$$

Here we also adopt the standard notation $W^{m,p}(\Omega)$ for Sobolev spaces on Ω with a norm $\|\cdot\|_{m,p}$ given by

$$\|v\|_{m,p}^p = \sum_{|\alpha|\le m} \|D^\alpha v\|_{L^p(\Omega)}^p ,$$

and a semi-norm $|\cdot|_{m,p}$ given by $\|v\|_{m,p}^p = \sum_{|\alpha|=m} \|D^\alpha v\|_{L^p(\Omega)}^p$. We set

$$W_0^{m,p}(\Omega) = \left\{ v \in W^{m,p}(\Omega) : v|_{\partial\Omega} = 0 \right\} ,$$

and for $p = 2$ write

$$H^m(\Omega) = W^{m,2}(\Omega), \qquad H_0^m(\Omega) = W_0^{m,2}$$

and

$$\|\cdot\|_m = \|\cdot\|_{m,2}, \qquad \|\cdot\| = \|\cdot\|_{0,2} .$$

Furthermore, $L^s(0,T;W^{m,p}(\Omega))$ denotes the Banach space of all L^s integrable functions from J into $W^{m,p}(\Omega)$ with norm

$$\|v\|_{L^s(J;W^{m,p}(\Omega))} = \left(\int_0^T \|v\|_{W^{m,p}(\Omega)}^s \, dt \right)^{1/s}$$

for $s \in [1,\infty)$ and the standard modification for $s = \infty$. Similarly, one can define the spaces $H^1(J;W^{m,p}(\Omega))$ and $C^k(J;W^{m,p}(\Omega))$ — cf. Ref. [24] for more detail. In addition, C denotes a general positive constant independent of h.

After the semidiscrete mixed finite element approximation for the parabolic OCP (1.1)-(1.5) is presented in Section 2, we introduce some projection operators and mixed elliptic constructions. Using these mixed elliptic reconstructions, we derive our *a posteriori* error estimates for both the state and the control approximation in Section 3. Our conclusions and comment on some possible future work appear in Section 4.

2. Mixed Methods for Optimal Control Problems

The state spaces assumed for the OCP are $\boldsymbol{L} = L^2(J;\boldsymbol{V})$ and $Q = H^1(J;W)$, where

$$\boldsymbol{V} = H(\text{div};\Omega) = \left\{ \boldsymbol{v} \in (L^2(\Omega))^2, \text{div}\boldsymbol{v} \in L^2(\Omega) \right\} \quad \text{and} \quad W = L^2(\Omega) ,$$

and the Hilbert space $\boldsymbol{V}$ is equipped with the norm

$$\|\boldsymbol{v}\|_{H(\text{div};\Omega)} = \left(\|\boldsymbol{v}\|_{0,\Omega}^2 + \|\text{div}\boldsymbol{v}\|_{0,\Omega}^2 \right)^{\frac{1}{2}} .$$

Letting $\alpha = A^{-1}$, we recast (1.1)-(1.5) in the following weak form:

Find $(\boldsymbol{p}, y, u) \in \boldsymbol{L} \times Q \times K$ such that

$$\min_{u \in K \subset U} \left\{ \frac{1}{2} \int_{o}^{T} (\|\boldsymbol{p} - \boldsymbol{p}_d\|^2 + \|y - y_d\|^2 + \|u\|^2)\, dt \right\}, \tag{2.1}$$

$$(\alpha \boldsymbol{p}, \boldsymbol{v}) - (y, \mathrm{div}\boldsymbol{v}) = 0 \qquad \forall\ \boldsymbol{v} \in \boldsymbol{V},\ t \in J\,, \tag{2.2}$$

$$(y_t, w) + (\mathrm{div}\boldsymbol{p}, w) = (f + u, w) \qquad \forall\ w \in W,\ t \in J\,, \tag{2.3}$$

$$y(x, 0) = y_0(x) \qquad \forall\ x \in \Omega\,. \tag{2.4}$$

The OCP (2.1)-(2.4) has a unique triplet solution $(\boldsymbol{p}, y, u)$ if and only if there is a co-state $(\boldsymbol{q}, z) \in \boldsymbol{L} \times Q$ such that $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ satisfies the following optimality conditions [23,29]:

$$(\alpha \boldsymbol{p}, \boldsymbol{v}) - (y, \mathrm{div}\boldsymbol{v}) = 0 \qquad \forall\ \boldsymbol{v} \in \boldsymbol{V},\ t \in J\,, \tag{2.5}$$

$$(y_t, w) + (\mathrm{div}\boldsymbol{p}, w) = (f + u, w) \qquad \forall\ w \in W,\ t \in J\,, \tag{2.6}$$

$$y(x, 0) = y_0(x) \qquad \forall\ x \in \Omega\,, \tag{2.7}$$

$$(\alpha \boldsymbol{q}, \boldsymbol{v}) - (z, \mathrm{div}\boldsymbol{v}) = -(\boldsymbol{p} - \boldsymbol{p}_d, \boldsymbol{v}) \qquad \forall\ \boldsymbol{v} \in \boldsymbol{V},\ t \in J\,, \tag{2.8}$$

$$-(z_t, w) + (\mathrm{div}\boldsymbol{q}, w) = (y - y_d, w) \qquad \forall\ w \in W,\ t \in J\,, \tag{2.9}$$

$$z(x, T) = 0 \qquad \forall\ x \in \Omega\,, \tag{2.10}$$

$$\int_0^T (u + z, \tilde{u} - u)\, dt \geq 0 \qquad \forall\ \tilde{u} \in K\,, \tag{2.11}$$

where $(\cdot,\cdot)$ is the inner product in $L^2(\Omega)$. Due to the special structure of the control constraint K, on using the same technique as in Ref. [9] we derive an important relationship between the optimal control u and the optimal co-state z that is a key to our analysis.

Lemma 2.1. *Let* $(y, \boldsymbol{p}, z, \boldsymbol{q}, u)$ *be the solution of* (2.5)-(2.11). *Then* $u = \max\{0, \bar{z}\} - z$, *where*

$$\bar{z} = \frac{\int_0^T \int_\Omega z\, dx\, dt}{\int_0^T \int_\Omega 1\, dx\, dt}$$

is the integral average of the function z *on* $\Omega \times [0, T]$.

Let $\mathcal{T}_h$ be regular triangulations of Ω, h_τ the diameter of τ, and $h = \max h_\tau$. Further, let $\mathcal{E}_h$ be the set of element sides of the triangulation $\mathcal{T}_h$ with $\Gamma_h = \cup \mathcal{E}_h$. Let $\boldsymbol{V}_h \times W_h \subset \boldsymbol{V} \times W$ denote the Raviart-Thomas space [12] associated with the triangulations $\mathcal{T}_h$ of Ω, P_k the space of polynomials of total degree at most $k \geq 0$, $\boldsymbol{V}(\tau) = \{\boldsymbol{v} \in P_k^2(\tau) + x \cdot P_k(\tau)\}$, and $W(\tau) = P_k(\tau)$. We define

$$\boldsymbol{V}_h := \left\{ \boldsymbol{v}_h \in \boldsymbol{V} : \forall \tau \in \mathcal{T}_h, \boldsymbol{v}_h|_\tau \in \boldsymbol{V}(\tau) \right\},$$

$$W_h := \left\{ w_h \in W : \forall \tau \in \mathcal{T}_h, w_h|_\tau \in W(\tau) \right\},$$

$$K_h := L^2(J; W_h) \cap K\,.$$

The mixed finite element discretisation of (2.1)-(2.4) is as follows.

Compute $(\boldsymbol{p}_h, y_h, u_h) \in L^2(J;\boldsymbol{V}_h) \times H^1(J;W_h) \times K_h$ such that

$$\min_{u_h(t)\in K_h} \left\{ \frac{1}{2}\int_0^T (\|\boldsymbol{p}_h - \boldsymbol{p}_d\|^2 + \|y_h - y_d\|^2 + \|u_h\|^2)\,dt \right\}, \tag{2.12}$$

$$(\alpha \boldsymbol{p}_h, \boldsymbol{v}_h) - (y_h, \mathrm{div}\boldsymbol{v}_h) = 0 \ \ \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, \qquad t\in J, \tag{2.13}$$

$$(y_{h,t}, w_h) + (\mathrm{div}\boldsymbol{p}_h, w_h) = (f + u_h, w_h) \ \ \forall\ w_h \in W_h, \qquad t\in J, \tag{2.14}$$

$$y_h(x,0) = y_0^h(x) \qquad \forall x\in\Omega, \tag{2.15}$$

where $y_0^h(x) \in W_h$ is an approximation of y_0. The OCP (2.12)-(2.15) again has a unique triplet solution $(\boldsymbol{p}_h, y_h, u_h)$ if and only if there is a co-state $(\boldsymbol{q}_h, z_h) \in L^2(J;\boldsymbol{V}_h) \times H^1(J;W_h)$ such that $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h)$ satisfies the optimality conditions

$$(\alpha \boldsymbol{p}_h, \boldsymbol{v}_h) - (y_h, \mathrm{div}\boldsymbol{v}_h) = 0 \qquad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h,\ t\in J, \tag{2.16}$$

$$(y_{h,t}, w_h) + (\mathrm{div}\boldsymbol{p}_h, w_h) = (f + u_h, w_h) \qquad \forall\ w_h \in W_h,\ t\in J, \tag{2.17}$$

$$y_h(x,0) = y_0^h(x) \qquad \forall\ x\in\Omega, \tag{2.18}$$

$$(\alpha \boldsymbol{q}_h, \boldsymbol{v}_h) - (z_h, \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h - \boldsymbol{p}_d, \boldsymbol{v}_h) \qquad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h,\ t\in J, \tag{2.19}$$

$$-(z_{h,t}, w_h) + (\mathrm{div}\boldsymbol{q}_h, w_h) = (y_h - y_d, w_h) \qquad \forall\ w_h \in W_h,\ t\in J, \tag{2.20}$$

$$z_h(x,T) = 0 \qquad \forall\ x\in\Omega, \tag{2.21}$$

$$\int_0^T (u_h + z_h, \tilde{u}_h - u_h)\,dt \ge 0 \qquad \forall\ \tilde{u}_h \in K_h. \tag{2.22}$$

We now introduce some intermediate variables. Thus for any control function $u_h \in K_h$ we first define the state solution $(\boldsymbol{p}(u_h), y(u_h), \boldsymbol{q}(u_h), z(u_h))$ such that

$$\big(\alpha \boldsymbol{p}(u_h), \boldsymbol{v}\big) - \big(y(u_h), \mathrm{div}\boldsymbol{v}\big) = 0 \qquad \forall\ \boldsymbol{v}\in\boldsymbol{V},\ t\in J, \tag{2.23}$$

$$\big(y_t(u_h), w\big) + \big(\mathrm{div}\boldsymbol{p}(u_h), w\big) = (f + u_h, w) \qquad \forall\ w\in W,\ t\in J, \tag{2.24}$$

$$y(u_h)(x,0) = y_0(x) \qquad \forall\ x\in\Omega, \tag{2.25}$$

$$\big(\alpha \boldsymbol{q}(u_h), \boldsymbol{v}\big) - \big(z(u_h), \mathrm{div}\boldsymbol{v}\big) = -\big(\boldsymbol{p}(u_h) - \boldsymbol{p}_d, \boldsymbol{v}\big) \qquad \forall\ \boldsymbol{v}\in\boldsymbol{V}_h,\ t\in J, \tag{2.26}$$

$$-\big(z_t(u_h), w\big) + \big(\mathrm{div}\boldsymbol{q}(u_h), w\big) = \big(y(u_h) - y_d, w\big) \qquad \forall\ w\in W,\ t\in J, \tag{2.27}$$

$$z(u_h)(x,T) = 0 \qquad \forall\ x\in\Omega, \tag{2.28}$$

where the exact solutions $y(u_h)$ and $z(u_h)$ satisfy the zero boundary condition. The errors are defined as

$$\begin{aligned} e_y &= y(u_h) - y_h, \quad & e_{\boldsymbol{p}} &= \boldsymbol{p}(u_h) - \boldsymbol{p}_h, \\ e_z &= z(u_h) - z_h, \quad & e_{\boldsymbol{q}} &= \boldsymbol{q}(u_h) - \boldsymbol{q}_h. \end{aligned}$$

Then from Eqs. (2.5)-(2.6), (2.8)-(2.9), (2.23)-(2.24) and (2.26)-(2.27), the errors satisfy the equations

$$(\alpha e_{\boldsymbol{p}}, \boldsymbol{v}) - (e_y, \mathrm{div}\boldsymbol{v}) = \mathbf{r}_1(\boldsymbol{v}) \qquad \forall\ \boldsymbol{v} \in \boldsymbol{V}, \tag{2.29}$$
$$(e_{y,t}, w) + (\mathrm{div} e_{\boldsymbol{p}}, w) = \mathbf{r}_2(w) \qquad \forall\ w \in W, \tag{2.30}$$
$$(\alpha e_{\boldsymbol{q}}, \boldsymbol{v}) - (e_z, \mathrm{div}\boldsymbol{v}) = -(e_{\boldsymbol{p}}, \boldsymbol{v}) + \mathbf{r}_3(\boldsymbol{v}) \qquad \forall\ \boldsymbol{v} \in \boldsymbol{V}, \tag{2.31}$$
$$-(e_{z,t}, w) + (\mathrm{div} e_{\boldsymbol{q}}, w) = (e_y, w) + \mathbf{r}_4(w) \qquad \forall\ w \in W, \tag{2.32}$$

and the residuals $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\}$ are

$$\mathbf{r}_1(\boldsymbol{v}) := (\alpha \boldsymbol{p}_h, \boldsymbol{v}) - (y_h, \mathrm{div}\boldsymbol{v}), \tag{2.33}$$
$$\mathbf{r}_2(w) := (y_{h,t}, w) + (\mathrm{div}\boldsymbol{p}_h, w) - (f + u_h, w), \tag{2.34}$$
$$\mathbf{r}_3(\boldsymbol{v}) := (\alpha \boldsymbol{q}_h, \boldsymbol{v}) + (\boldsymbol{p}_h - \boldsymbol{p}_d, \boldsymbol{v}) - (z_h, \mathrm{div}\boldsymbol{v}), \tag{2.35}$$
$$\mathbf{r}_4(w) := -(z_{h,t}, w) + (\mathrm{div}\boldsymbol{q}_h, w) - (y_h - y_d, w). \tag{2.36}$$

We also introduce the mixed elliptic reconstructions $\tilde{y}(t)$, $\tilde{z}(t) \in H_0^1(\Omega)$ and $\tilde{\boldsymbol{p}}(t)$, $\tilde{\boldsymbol{q}}(t) \in V$ of y_h, z_h and $\boldsymbol{p}_h$, $\boldsymbol{q}_h$ for $t \in [0, T]$, respectively. For given y_h, z_h, $\boldsymbol{p}_h$ and $\boldsymbol{q}_h$, the mixed elliptic reconstructions $\tilde{y}(t)$, $\tilde{z}(t) \in H_0^1(\Omega)$ and $\tilde{\boldsymbol{p}}(t)$, $\tilde{\boldsymbol{q}}(t) \in \boldsymbol{V}$ are assumed to satisfy

$$\big(\alpha(\tilde{\boldsymbol{p}} - \boldsymbol{p}_h), \boldsymbol{v}\big) - (\tilde{y} - y_h, \mathrm{div}\boldsymbol{v}) = \mathbf{r}_1(\boldsymbol{v}) \qquad \forall\ \boldsymbol{v} \in \boldsymbol{V}, \tag{2.37}$$
$$\big(\mathrm{div}(\tilde{\boldsymbol{p}} - \boldsymbol{p}_h), w\big) = \mathbf{r}_2(w) \qquad \forall\ w \in W, \tag{2.38}$$
$$\big(\alpha(\tilde{\boldsymbol{q}} - \boldsymbol{q}_h), \boldsymbol{v}\big) - (\tilde{z} - z_h, \mathrm{div}\boldsymbol{v}) = -(\tilde{\boldsymbol{p}} - \boldsymbol{p}_h, \boldsymbol{v}) + \mathbf{r}_3(\boldsymbol{v}) \qquad \forall\ \boldsymbol{v} \in \boldsymbol{V}, \tag{2.39}$$
$$\big(\mathrm{div}(\tilde{\boldsymbol{q}} - \boldsymbol{q}_h), w\big) = (\tilde{y} - y_h, w) + \mathbf{r}_4(w) \qquad \forall\ w \in W. \tag{2.40}$$

Since $\mathbf{r}_1(\boldsymbol{v}_h) = \mathbf{r}_3(\boldsymbol{v}_h) = 0$, $\forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h$, $\mathbf{r}_2(w_h) = \mathbf{r}_4(w_h) = 0$, $\forall\ w_h \in W_h$, we note that y_h and $\boldsymbol{p}_h$ are standard mixed elliptic projections of $\tilde{y}$ and $\tilde{\boldsymbol{p}}$, and z_h and $\boldsymbol{q}_h$ are nonstandard mixed elliptic projections of $\tilde{z}$ and $\tilde{\boldsymbol{q}}$, respectively. Using the mixed elliptic reconstructions, we now rewrite

$$e_{\boldsymbol{p}} = (\tilde{\boldsymbol{p}} - \boldsymbol{p}_h) - (\tilde{\boldsymbol{p}} - \boldsymbol{p}(u_h)) := \eta_{\boldsymbol{p}} - \xi_{\boldsymbol{p}},$$
$$e_y = (\tilde{y} - y_h) - (\tilde{y} - y(u_h)) := \eta_y - \xi_y,$$
$$e_{\boldsymbol{q}} = (\tilde{\boldsymbol{q}} - \boldsymbol{q}_h) - (\tilde{\boldsymbol{q}} - \boldsymbol{q}(u_h)) := \eta_{\boldsymbol{q}} - \xi_{\boldsymbol{q}},$$
$$e_z = (\tilde{z} - z_h) - (\tilde{z} - z(u_h)) := \eta_z - \xi_z.$$

Let $P_h : W \to W_h$ be the orthogonal $L^2(\Omega)$-projection into W_h [2] that satisfies

$$(P_h w - w, \chi) = 0, \qquad w \in W,\ \chi \in W_h, \tag{2.41}$$
$$\|P_h w - w\|_{0,q} \le \|w\|_{t,q} h^t, \qquad 0 \le t \le k+1 \ \text{ if } w \in W \cap W^{t,q}(\Omega), \tag{2.42}$$
$$\|P_h w - w\|_{-r} \le C\|w\|_t h^{r+t}, \qquad 0 \le r, t \le k+1 \ \text{ if } w \in H^t(\Omega). \tag{2.43}$$

Next, we recall the Fortin projection [6, 12] $\Pi_h : \boldsymbol{V} \to \boldsymbol{V}_h$ such that for any $\boldsymbol{q} \in \boldsymbol{V}$

$$\big(\mathrm{div}(\Pi_h \boldsymbol{q} - \boldsymbol{q}), w_h\big) = 0, \qquad \forall\ \boldsymbol{q} \in \boldsymbol{V},\ w_h \in W_h, \tag{2.44}$$
$$\|\boldsymbol{q} - \Pi_h \boldsymbol{q}\|_{0,q} \le C h^r \|\boldsymbol{q}\|_{r,q}, \qquad 1/q < r \le k+1,\ \forall\ \boldsymbol{q} \in \boldsymbol{V} \cap (W^{r,q}(\Omega))^2, \tag{2.45}$$
$$\|\mathrm{div}(\boldsymbol{q} - \Pi_h \boldsymbol{q})\| \le C h^r \|\mathrm{div}\boldsymbol{q}\|_r, \qquad 0 \le r \le k+1, \qquad \forall\ \mathrm{div}\boldsymbol{q} \in H^r(\Omega). \tag{2.46}$$

We have the commuting diagram property

$$\mathrm{div} \circ \Pi_h = P_h \circ \mathrm{div} : \boldsymbol{V} \to W_h \;\text{ and }\; \mathrm{div}(I - \Pi_h)\boldsymbol{V} \perp W_h \,, \tag{2.47}$$

where and hereafter I denotes the identity operator.

3. *A Posteriori* Error Estimates

In this section, we develop our *a posteriori* error estimates for the mixed finite element approximation to the parabolic OCP. Let $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ and $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h)$ be the solutions of (2.5)-(2.11) and (2.16)-(2.22), respectively. We decompose the errors as

$$\begin{aligned}
\boldsymbol{p} - \boldsymbol{p}_h &= \boldsymbol{p} - \boldsymbol{p}(u_h) + \boldsymbol{p}(u_h) - \boldsymbol{p}_h := r_{\boldsymbol{p}} + e_{\boldsymbol{p}} \,, \\
y - y_h &= y - y(u_h) + y(u_h) - y_h := r_y + e_y \,, \\
\boldsymbol{q} - \boldsymbol{q}_h &= \boldsymbol{q} - \boldsymbol{q}(u_h) + \boldsymbol{q}(u_h) - \boldsymbol{q}_h := r_{\boldsymbol{q}} + e_{\boldsymbol{q}} \,, \\
z - z_h &= z - z(u_h) + z(u_h) - z_h := r_z + e_z \,.
\end{aligned}$$

From Eqs. (2.5)-(2.6), (2.8)-(2.9), (2.23)-(2.24) and (2.26)-(2.27) the error equations are

$$(\alpha r_{\boldsymbol{p}}, \boldsymbol{v}) - (r_y, \mathrm{div}\boldsymbol{v}) = 0 \qquad \forall\ \boldsymbol{v} \in \boldsymbol{V} \,, \tag{3.1}$$

$$(r_{y,t}, w) + (\mathrm{div} r_{\boldsymbol{p}}, w) = (u - u_h, w) \qquad \forall\ w \in W \,, \tag{3.2}$$

$$(\alpha r_{\boldsymbol{p}}, \boldsymbol{v}) - (r_z, \mathrm{div}\boldsymbol{v}) = -(r_{\boldsymbol{p}}, \boldsymbol{v}) \qquad \forall\ \boldsymbol{v} \in \boldsymbol{V} \,, \tag{3.3}$$

$$-(r_{z,t}, w) + (\mathrm{div} r_{\boldsymbol{q}}, w) = (r_y, w) \qquad \forall\ w \in W \,. \tag{3.4}$$

Lemma 3.1. *Let* $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ *and* $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h)$ *be the solutions of* (2.5)-(2.11) *and* (2.16)-(2.22)*, respectively. Then there is a constant* $C > 0$ *independent of* h *such that*

$$\|r_{\boldsymbol{p}}\|_{L^2(J;L^2(\Omega))} + \|r_y\|_{L^\infty(J;L^2(\Omega))} \le C\|u - u_h\|_{L^2(J;L^2(\Omega))} \,, \tag{3.5}$$

$$\|r_{y,t}\|_{L^2(J;L^2(\Omega))} + \|r_{\boldsymbol{p}}\|_{L^\infty(J;L^2(\Omega))} \le C\|u - u_h\|_{L^2(J;L^2(\Omega))} \,, \tag{3.6}$$

$$\|div r_{\boldsymbol{p}}\|_{L^2(J;L^2(\Omega))} \le C\|u - u_h\|_{L^2(J;L^2(\Omega))} \,, \tag{3.7}$$

$$\|r_{\boldsymbol{q}}\|_{L^2(J;L^2(\Omega))} + \|r_z\|_{L^\infty(J;L^2(\Omega))} \le C\|u - u_h\|_{L^2(J;L^2(\Omega))} \,, \tag{3.8}$$

$$\|r_{z,t}\|_{L^2(J;L^2(\Omega))} + \|r_{\boldsymbol{q}}\|_{L^\infty(J;L^2(\Omega))} \le C\|u - u_h\|_{L^2(J;L^2(\Omega))} \,, \tag{3.9}$$

$$\|div r_{\boldsymbol{q}}\|_{L^2(J;L^2(\Omega))} \le C\|u - u_h\|_{L^2(J;L^2(\Omega))} \,. \tag{3.10}$$

Proof. Part **I**. Choosing $\boldsymbol{v} = r_{\boldsymbol{p}}$ and $w = r_y$ as the test functions and adding the two relations (3.1) and (3.2), we have

$$(\alpha r_{\boldsymbol{p}}, r_{\boldsymbol{p}}) + (r_{y,t}, r_y) = (u - u_h, r_y) \,, \tag{3.11}$$

so using the ϵ-Cauchy inequality we can find an estimate

$$(\alpha r_{\boldsymbol{p}}, r_{\boldsymbol{p}}) + (r_{y,t}, r_y) \le C\Big(\|r_y\|^2 + \|u - u_h\|^2\Big) \,. \tag{3.12}$$

Noting that

$$(r_{y,t}, r_y) = \frac{1}{2}\frac{d}{dt}\|r_y\|^2, \tag{3.13}$$

on using the assumption on A we obtain

$$\|r_{\boldsymbol{p}}\|^2 + \frac{1}{2}\frac{d}{dt}\|r_y\|^2 \leq C\Big(\|r_y\|^2 + \|u-u_h\|^2\Big). \tag{3.14}$$

Integrating (3.14) in time from 0 to t and noting that $r_y(0) = 0$, on using Gronwall's Lemma [40] we get

$$\|r_{\boldsymbol{p}}\|^2_{L^2(J;L^2(\Omega))} + \|r_y\|^2_{L^\infty(J;L^2(\Omega))} \leq C\|u-u_h\|^2_{L^2(J;L^2(\Omega))}, \tag{3.15}$$

and setting $t = 0$ and $\boldsymbol{v} = r_{\boldsymbol{p}}(0)$ in Eq. (3.1) we consequently find that $r_{\boldsymbol{p}}(0) = 0$. Then differentiating Eq. (3.1)with respect to t we also obtain

$$(\alpha r_{\boldsymbol{p},t}, \boldsymbol{v}) - (r_{y,t}, \mathrm{div}\boldsymbol{v}) = 0 \ \ \forall\ \boldsymbol{v} \in \boldsymbol{V}, \tag{3.16}$$

so setting $\boldsymbol{v} = r_{\boldsymbol{p}}$ and $w = r_{y,t}$ as the test functions and adding the two relations (3.16) and (3.2) we have

$$(\alpha r_{\boldsymbol{p},t}, r_{\boldsymbol{p}}) + (r_{y,t}, r_{y,t}) = (u-u_h, r_{y,t}). \tag{3.17}$$

On using the ϵ-Cauchy inequality, we get

$$\frac{1}{2}\frac{d}{dt}\|\alpha^{\frac{1}{2}} r_{\boldsymbol{p}}\|^2 + \|r_{y,t}\|^2 \leq C\|u-u_h\|^2, \tag{3.18}$$

and integrating (3.18) with respect to time from 0 to t we obtain

$$\|r_{y,t}\|_{L^2(J;L^2(\Omega))} + \|r_{\boldsymbol{p}}\|_{L^\infty(J;L^2(\Omega))} \leq C\|u-u_h\|_{L^2(J;L^2(\Omega))} \tag{3.19}$$

on using the assumption on A and $r_{\boldsymbol{p}}(0) = 0$. Setting $w = \mathrm{div} r_{\boldsymbol{p}}$ in Eq. (3.2), we find that

$$\|\mathrm{div} r_{\boldsymbol{p}}\|_{L^2(J;L^2(\Omega))} \leq C\Big(\|u-u_h\|_{L^2(J;L^2(\Omega))} + \|r_{y,t}\|_{L^2(J;L^2(\Omega))}\Big). \tag{3.20}$$

Combining (3.15) and (3.19) with (3.20), we obtain (3.5)-(3.7).

Part **II**. Selecting $\boldsymbol{v} = r_{\boldsymbol{q}}$ and $w = r_z$ as the test functions and adding the two relations (3.3) and (3.4), we obtain

$$\|\alpha^{\frac{1}{2}} r_{\boldsymbol{q}}\|^2 - \frac{1}{2}\frac{d}{dt}\|r_z\|^2 = (r_y, r_z) - (r_{\boldsymbol{p}}, r_{\boldsymbol{q}}), \tag{3.21}$$

so from the ϵ-Cauchy inequality and the assumption on A we find that

$$\|r_{\boldsymbol{q}}\|^2 - \frac{1}{2}\frac{d}{dt}\|r_z\|^2 \leq C\Big(\|r_y\|^2_{L^2(\Omega)} + \|r_{\boldsymbol{p}}\|^2_{L^2(\Omega)} + \|r_z\|^2_{L^2(\Omega)}\Big). \tag{3.22}$$

Integrating (3.22) from t to T and noting that $r_2(T) = 0$, on applying Gronwall's Lemma [40] we readily obtain the error estimate

$$\|r_{\boldsymbol{q}}\|^2_{L^2(J;L^2(\Omega))} + \|r_z\|^2_{L^\infty(J;L^2(\Omega))} \leq C\left(\|r_{\boldsymbol{p}}\|^2_{L^2(J;L^2(\Omega))} + \|r_y\|^2_{L^2(J;L^2(\Omega))}\right). \tag{3.23}$$

Differentiating Eq. (3.3) with respect to t, we get

$$(\alpha r_{\boldsymbol{q},t}, \boldsymbol{v}) - (r_{z,t}, \mathrm{div}\boldsymbol{v}) = -(r_{\boldsymbol{p},t}, \boldsymbol{v}) \ \ \forall \ \boldsymbol{v} \in \boldsymbol{V}, \tag{3.24}$$

so on setting $t = T$ in Eq. (3.3) and noting that $r_{z,t}(T) = 0$ we have

$$(\alpha r_{\boldsymbol{q}}(T), \boldsymbol{v}) = -(r_{\boldsymbol{p}}(T), \boldsymbol{v}) \ \ \forall \ \boldsymbol{v} \in \boldsymbol{V}. \tag{3.25}$$

Choosing $v = r_{\boldsymbol{q}}(T)$ in Eq. (3.25), from the ϵ-Cauchy inequality and the assumption on A

$$\|\alpha^{\frac{1}{2}} r_{\boldsymbol{q}}(T)\| \leq C\|r_{\boldsymbol{p}}(T)\|. \tag{3.26}$$

Selecting $v = -r_{\boldsymbol{q}}$ and $w = -r_{z,t}$ as the test functions and adding the two relations (3.3) and (3.4), we obtain

$$\|r_z\|^2 - \frac{1}{2}\frac{d}{dt}\|\alpha^{\frac{1}{2}} r_{\boldsymbol{q}}\|^2 = (r_{\boldsymbol{p},t}, r_{\boldsymbol{q}}) - (r_y, r_{z,t}), \tag{3.27}$$

and from Eq. (3.16) and the ϵ-Cauchy inequality we know that

$$\begin{aligned}(r_{\boldsymbol{p},t}, r_{\boldsymbol{q}}) &= (\alpha r_{\boldsymbol{p},t}, A r_{\boldsymbol{q}}) = \left(r_{y,t}, \mathrm{div}(A r_{\boldsymbol{q}})\right)\\ &\leq \|r_{y,t}\| \cdot \|\mathrm{div}(A r_{\boldsymbol{q}})\|\\ &\leq C\|r_{y,t}\| \cdot \left(\|r_{\boldsymbol{q}}\| + \|\mathrm{div} r_{\boldsymbol{q}}\|\right)\\ &\leq C(\delta)\left(\|r_{y,t}\|^2 + \|r_{\boldsymbol{q}}\|^2\right) + \delta\|\mathrm{div} r_{\boldsymbol{q}}\|^2.\end{aligned} \tag{3.28}$$

On invoking (3.28) in Eq. (3.27) and using the ϵ-Cauchy inequality we get

$$\|r_{z,t}\|^2 - \frac{1}{2}\frac{d}{dt}\|\alpha^{\frac{1}{2}} r_{\boldsymbol{q}}\|^2 \leq C\left(\|r_{y,t}\|^2 + \|r_y\|^2 + \|r_{\boldsymbol{q}}\|^2\right) + \frac{1}{8}\|\mathrm{div} r_{\boldsymbol{q}}\|^2; \tag{3.29}$$

and integrating (3.29) with respect to time from t to T, given (3.26) we arrive at

$$\begin{aligned}&\int_t^T \|r_{z,t}\|^2 ds + \|\alpha^{\frac{1}{2}} r_{\boldsymbol{q}}\|^2\\ \leq\, & C\int_t^T \left(\|r_{y,t}\|^2 + \|r_y\|^2 + \|r_{\boldsymbol{q}}\|^2\right) ds + C\|r_{\boldsymbol{p}}(T)\|^2 + \frac{1}{8}\int_t^T \|\mathrm{div} r_{\boldsymbol{q}}\|^2 ds.\end{aligned} \tag{3.30}$$

Choosing $w = \mathrm{div} r_{\boldsymbol{q}}$ as a test function in Eq. (3.4) and using the ϵ-Cauchy inequality,

$$\int_t^T \|\mathrm{div} r_{\boldsymbol{q}}\|^2 ds \leq 4\int_t^T \left(\|r_{z,t}\|^2 + \|r_y\|^2\right) ds, \tag{3.31}$$

so on invoking (3.31) in (3.30) and using Gronwall's Lemma [40] we arrive at

$$\|r_{z,t}\|_{L^2(J;L^2(\Omega))} + \|r_{\boldsymbol{q}}\|_{L^\infty(J;L^2(\Omega))} \le C\left(\|r_y\|_{H^1(J;L^2(\Omega))} + \|r_{\boldsymbol{p}}(T)\|\right). \tag{3.32}$$

Inserting (3.5) and (3.6) into (3.23), (3.31) and (3.32) respectively yields (3.8)-(3.10). □

We now proceed to derive the *a posteriori* error estimates for the control u.

Lemma 3.2. *Let $(y,\boldsymbol{p},z,\boldsymbol{q},u)$ and $(y_h,\boldsymbol{p}_h,z_h,\boldsymbol{q}_h,u_h)$ be the solutions of optimality conditions* (2.5)-(2.11) *and* (2.16)-(2.22)*, respectively. Then*

$$\|u-u_h\|^2_{L^2(J;L^2(\Omega))} \le C\eta_1^2 + C\|z_h - z(u_h)\|^2_{L^2(J;L^2(\Omega))}, \tag{3.33}$$

where

$$\eta_1^2 = \|u_h + z_h\|^2_{L^2(J;L^2(\Omega))}.$$

Proof. It follows from Eq. (2.11) that

$$\begin{aligned}
\|u-u_h\|^2_{L^2(J;L^2(\Omega))} &= \int_0^T (u-u_h, u-u_h)\,dt \\
&= \int_0^T (u+z, u-u_h)\,dt + \int_0^T (u_h+z_h, u_h-u)\,dt \\
&\quad + \int_0^T (z_h - z(u_h), u-u_h)\,dt + \int_0^T (z(u_h)-z, u-u_h)\,dt \\
&\le \int_0^T (u_h+z_h, u_h-u)\,dt + \int_0^T (z_h - z(u_h), u-u_h)\,dt \\
&\quad + \int_0^T (z(u_h)-z, u-u_h)\,dt =: I_1 + I_2 + I_3.
\end{aligned} \tag{3.34}$$

We first estimate I_1, I_2 and I_3. Thus

$$\begin{aligned}
I_1 &= \int_0^T (u_h+z_h, u_h-u)\,dt \\
&\le C(\delta)\|u_h+z_h\|^2_{L^2(J;L^2(\Omega))} + \delta\|u-u_h\|^2_{L^2(J;L^2(\Omega))} \\
&\le C(\delta)\eta_1^2 + \delta\|u-u_h\|^2_{L^2(J;L^2(\Omega))},
\end{aligned} \tag{3.35}$$

where $\delta > 0$ is an arbitrarily small number and $C(\delta)$ is dependent on δ^{-1}, and clearly

$$\begin{aligned}
I_2 &= \int_0^T (z_h - z(u_h), u-u_h)\,dt \\
&\le C(\delta)\|z_h + z(u_h)\|^2_{L^2(J;L^2(\Omega))} + \delta\|u-u_h\|^2_{L^2(J;L^2(\Omega))},
\end{aligned} \tag{3.36}$$

so let us now turn to I_3. Noting that

$$y(x,0)=y(u_h)(x,0)=y_0(x) \text{ and } z(x,T)=z(u_h)(x,T)=0\,,$$

from Eqs. (2.5)-(2.6), (2.8)-(2.9), (2.23)-(2.24) and (2.26)-(2.27), we obtain

$$\begin{aligned}
I_3 &= \int_0^T (u-u_h, z(u_h)-z)\,dt \\
&= \int_0^T \big(((y-y(u_h))_t, z(u_h)-z)+(\mathrm{div}(\boldsymbol{p}-\boldsymbol{p}(u_h)), z(u_h)-z)\big)\,dt \\
&= -\int_0^T \big((y-y(u_h)), (z(u_h)-z)_t)+(A^{-1}(\boldsymbol{q}(u_h)-\boldsymbol{q}), \boldsymbol{p}-\boldsymbol{p}(u_h))\big)\,dt \\
&\quad +\int_0^T (\boldsymbol{p}(u_h)-\boldsymbol{p}, \boldsymbol{p}-\boldsymbol{p}(u_h))\,dt \\
&= -\int_0^T \big(((z(u_h)-z)_t, y-y(u_h))+(y-y(u_h), \mathrm{div}(\boldsymbol{q}-\boldsymbol{q}(u_h)))\big)\,dt \\
&\quad +\int_0^T (\boldsymbol{p}(u_h)-\boldsymbol{p}, \boldsymbol{p}-\boldsymbol{p}(u_h))\,dt \\
&= \int_0^T \big((\boldsymbol{p}(u_h)-\boldsymbol{p}, \boldsymbol{p}-\boldsymbol{p}(u_h))+(y(u_h)-y, y-y(u_h))\big)\,dt \\
&\le 0\,.
\end{aligned} \tag{3.37}$$

Thus from (3.34)-(3.37) we have (3.33). □

Using Eqs. (2.37)-(2.40) in Eqs. (2.29)-(2.32), we then obtain the error equations

$$(\alpha\xi_{\boldsymbol{p}}, \boldsymbol{v})-(\xi_y, \mathrm{div}\boldsymbol{v})=0 \qquad \forall\ \boldsymbol{v}\in\boldsymbol{V}\,, \tag{3.38}$$

$$(\xi_{y,t}, w)+(\mathrm{div}\xi_{\boldsymbol{p}}, w)=(\eta_{y,t}, w) \qquad \forall\ w\in W\,, \tag{3.39}$$

$$(\alpha\xi_{\boldsymbol{q}}, \boldsymbol{v})-(\xi_z, \mathrm{div}\boldsymbol{v})=-(\xi_{\boldsymbol{p}}, \boldsymbol{v}) \qquad \forall\ \boldsymbol{v}\in\boldsymbol{V}\,, \tag{3.40}$$

$$-(\xi_{z,t}, w)+(\mathrm{div}\xi_{\boldsymbol{q}}, w)=(\xi_y, w)-(\eta_{z,t}, w) \qquad \forall\ w\in W\,. \tag{3.41}$$

Lemma 3.3. *Let mixed elliptic reconstructions $\tilde{y}$, $\tilde{\boldsymbol{p}}$, $\tilde{z}$ and $\tilde{\boldsymbol{q}}$ satisfy* (2.37)-(2.40). *Then the following properties hold:*

$$\alpha\tilde{\boldsymbol{p}}=-\nabla\tilde{y}, \qquad \alpha\tilde{\boldsymbol{q}}+\tilde{\boldsymbol{p}}-p_d=-\nabla\tilde{z}. \tag{3.42}$$

Proof. From Eqs. (3.38) and (3.40), on integrating the second term on the left-hand side by parts we have

$$(\alpha\xi_{\boldsymbol{p}}, \boldsymbol{v})=-(\nabla\xi_y, \boldsymbol{v}) \qquad \forall\ \boldsymbol{v}\in\boldsymbol{V}\,, \tag{3.43}$$

$$(\alpha\xi_{\boldsymbol{q}}, \boldsymbol{v})+(\xi_{\boldsymbol{p}}, \boldsymbol{v})=-(\nabla\xi_z, \boldsymbol{v}) \qquad \forall\ \boldsymbol{v}\in\boldsymbol{V}\,, \tag{3.44}$$

and hence

$$\alpha\xi_{\boldsymbol{p}} = -\nabla\xi_y, \quad \alpha\xi_{\boldsymbol{q}} + \xi_{\boldsymbol{p}} = -\nabla\xi_z \,. \tag{3.45}$$

Since the exact solutions $y(u_h)$ and $z(u_h)$ satisfy the zero boundary condition, we then readily obtain (3.42). □

Lemma 3.4. *Let* ξ_y *and* $\xi_{\boldsymbol{p}}$ *satisfy* (3.38)-(3.39). *Then the following estimates hold:*

$$\|\xi_{\boldsymbol{p}}\|_{L^2(J;L^2(\Omega))} + \|\xi_y\|_{L^\infty(J;L^2(\Omega))} \le C\|\xi_y(0)\| + C\|\eta_{y,t}\|_{L^2(J;H^{-1}(\Omega))}\,, \tag{3.46}$$

$$\|\xi_{y,t}\|_{L^2(J;L^2(\Omega))} + \|\xi_{\boldsymbol{p}}\|_{L^\infty(J;L^2(\Omega))} \le C\|\xi_{\boldsymbol{p}}(0)\| + C\|\eta_{y,t}\|_{L^2(J;L^2(\Omega))}\,, \tag{3.47}$$

$$\|div\xi_p\|_{L^2(J;L^2(\Omega))} \le C\|\xi_{\boldsymbol{p}}(0)\| + C\|\eta_{y,t}\|_{L^2(J;L^2(\Omega))}\,, \tag{3.48}$$

$$\|\xi_y\|_{L^2(J;L^2(\Omega))} \le C\|y_0 - y_0^h\| + C\|\eta_y\|_{L^2(J;L^2(\Omega))}\,. \tag{3.49}$$

Proof. Choosing $\boldsymbol{v} = \xi_{\boldsymbol{p}}$ and $w = \xi_y$ as the test functions and adding the two relations (3.38) and (3.39), we have

$$(\alpha\xi_{\boldsymbol{p}}, \xi_{\boldsymbol{p}}) + (\xi_{y,t}, \xi_y) = (\eta_{y,t}, \xi_y)\,. \tag{3.50}$$

We estimate the term on the right-hand side of Eq. (3.50) as

$$|(\eta_{y,t}, \xi_y)| \le \|\eta_{y,t}\|_{-1}\|\nabla\xi_y\|\,, \tag{3.51}$$

and hence from (3.42) and Young's inequality we find that

$$|(\eta_{y,t}, \xi_y)| \le C\|\eta_{y,t}\|_{-1}\|\alpha^{\frac{1}{2}}\xi_{\boldsymbol{p}}\| \le C\|\eta_{y,t}\|_{-1}^2 + \frac{1}{2}\|\alpha^{\frac{1}{2}}\xi_{\boldsymbol{p}}\|^2\,. \tag{3.52}$$

On invoking (3.52) in Eq. (3.50) and integrating with respect to time from 0 to t, we arrive at

$$\|\xi_{\boldsymbol{p}}\|_{L^2(J;L^2(\Omega))} + \|\xi_y\|_{L^\infty(J;L^2(\Omega))} \le C\|\xi_y(0)\| + C\|\eta_{y,t}\|_{L^2(J;H^{-1}(\Omega))}\,, \tag{3.53}$$

completing the estimate (3.46). To obtain (3.47), we differentiate Eq. (3.38) with respect to t, to obtain

$$(\alpha\xi_{\boldsymbol{p},t}, \boldsymbol{v}) - (\xi_{y,t}, \mathrm{div}\boldsymbol{v}) = 0 \ \ \forall\ \boldsymbol{v} \in \boldsymbol{V}\,. \tag{3.54}$$

Choosing $\boldsymbol{v} = \xi_{\boldsymbol{p}}$ and $w = \xi_{y,t}$ as the test functions and adding the two relations (3.54) and (3.39), on using the ϵ-Cauchy inequality we obtain

$$\frac{1}{2}\frac{d}{dt}\|\alpha^{\frac{1}{2}}\xi_{\boldsymbol{p}}\|^2 + \|\xi_{y,t}\|^2 \le \frac{1}{2}\|\eta_{y,t}\|^2 + \frac{1}{2}\|\xi_{y,t}\|^2\,. \tag{3.55}$$

On integrating (3.55) with respect to time from 0 to t and using the assumption on A, we find

$$\|\xi_{y,t}\|_{L^2(J;L^2(\Omega))} + \|\xi_{\boldsymbol{p}}\|_{L^\infty(J;L^2(\Omega))} \le C\|\xi_{\boldsymbol{p}}(0)\| + C\|\eta_{y,t}\|_{L^2(J;L^2(\Omega))}\,. \tag{3.56}$$

So on setting $w = \text{div} r_{\boldsymbol{p}}$ in (3.39) we obtain

$$\|\text{div}\xi_{\boldsymbol{p}}\|_{L^2(J;L^2(\Omega))} \le C\|\xi_{y,t}\|_{L^2(J;L^2(\Omega))} + C\|\eta_{y,t}\|_{L^2(J;L^2(\Omega))} . \tag{3.57}$$

On integrating Eq. (3.39) with respect to time from 0 to t and using the symbol

$$\hat{\phi}(t) := \int_0^t \phi(s)\, ds , \tag{3.58}$$

we find that

$$(\xi_y, w) + (\text{div}\hat{\xi}_{\boldsymbol{p}}, w) = (y_0^h - y_0, w) + (\eta_y, w) \quad \forall\, w \in W . \tag{3.59}$$

Setting $w = \eta_y$ in Eq. (3.59) and $\boldsymbol{v} = \hat{\xi}_{\boldsymbol{p}}$ in Eq. (3.38) and then adding the resulting equations, from the Cauchy-Schwarz and Young's inequalities we obtain

$$\|\xi_y\|^2 + \frac{d}{dt}\|\alpha^{\frac{1}{2}}\hat{\xi}_{\boldsymbol{p}}\|^2 \le 2\Big(\|y_0^h - y_0\|^2 + \|\eta_y\|^2\Big) . \tag{3.60}$$

Integrating with respect to time from 0 to t, we then arrive at

$$\|\xi_y\|_{L^2(J;L^2(\Omega))} \le 2T\|y_0 - y_0^h\| + 2\|\eta_y\|_{L^2(J;L^2(\Omega))} , \tag{3.61}$$

and combining (3.53), (3.56) and (3.57) with (3.61) completes the proof. □

Lemma 3.5. *Let* ξ_z *and* $\xi_{\boldsymbol{q}}$ *satisfy* (3.40)-(3.41). *Then the following estimates hold:*

$$\begin{aligned}
&\|\xi_{\boldsymbol{q}}\|_{L^2(J;L^2(\Omega))} + \|\xi_z\|_{L^\infty(J;L^2(\Omega))} \\
&\le C\|\xi_y(0)\| + C\|\eta_{y,t}\|_{L^2(J;H^{-1}(\Omega))} + C\|\xi_z(T)\| + C\|\eta_{z,t}\|_{L^2(J;H^{-1}(\Omega))} , && (3.62)\\
&\|\xi_{z,t}\|_{L^2(J;L^2(\Omega))} + \|\xi_{\boldsymbol{q}}\|_{L^\infty(J;L^2(\Omega))} \\
&\le C\|\xi_y(0)\| + C\|\xi_{\boldsymbol{p}}(0)\| + C\|\eta_{y,t}\|_{L^2(J;L^2(\Omega))} + C\|\xi_{\boldsymbol{q}}(T)\| + C\|\eta_{z,t}\|_{L^2(J;L^2(\Omega))} , && (3.63)\\
&\|div\xi_{\boldsymbol{q}}\|_{L^2(J;L^2(\Omega))} \\
&\le C\|\xi_y(0)\| + C\|\xi_{\boldsymbol{p}}(0)\| + C\|\eta_y\|_{H^1(J;L^2(\Omega))} + C\|\xi_{\boldsymbol{q}}(T)\| + C\|\eta_{z,t}\|_{L^2(J;L^2(\Omega))} \\
&\quad + C\|y_0 - y_0^h\| . && (3.64)
\end{aligned}$$

Proof. Selecting $\boldsymbol{v} = \xi_{\boldsymbol{q}}$ and $w = \xi_z$ as the test functions and adding the two relations (3.40) and (3.41), we obtain

$$\|\alpha\xi_{\boldsymbol{q}}\|^2 - \frac{1}{2}\frac{d}{dt}\|\xi_z\|^2 = (\xi_y, \xi_z) - (\eta_{z,t}, \xi_z) - (\xi_{\boldsymbol{p}}, \xi_{\boldsymbol{q}}) . \tag{3.65}$$

We estimate the term on the right-hand side of Eq. (3.65) as

$$|(\eta_{z,t}, \xi_z)| \le \|\eta_{z,t}\|_{-1}\|\nabla\xi_z\| , \tag{3.66}$$

so using (3.42) with Young's inequality we find that

$$|(\eta_{z,t},\xi_z)\le C\|\eta_{z,t}\|_{-1}\|\alpha\xi_{\boldsymbol{q}}+\xi_{\boldsymbol{p}}\|\le C\|\eta_{z,t}\|_{-1}^2+C\|\xi_{\boldsymbol{p}}\|^2+\frac{1}{2}\|\alpha^{\frac{1}{2}}\xi_{\boldsymbol{q}}\|^2\,. \tag{3.67}$$

Then using (3.65)-(3.67) and the ϵ-Cauchy inequality, we obtain

$$\|\alpha^{\frac{1}{2}}\xi_{\boldsymbol{q}}\|^2-\frac{1}{2}\frac{d}{dt}\|\xi_z\|^2\le C\left(\|\eta_{z,t}\|_{-1}^2+\|\xi_y\|_{L^2(\Omega)}^2+\|\xi_{\boldsymbol{p}}\|_{L^2(\Omega)}^2+\|\xi_z\|_{L^2(\Omega)}^2\right)\,. \tag{3.68}$$

Integrating (3.68) from t to T and applying the assumption on A and Gronwall's Lemma [40], we readily obtain the error estimate

$$\begin{aligned}&\|\xi_{\boldsymbol{q}}\|_{L^2(J;L^2(\Omega))}+\|\xi_z\|_{L^\infty(J;L^2(\Omega))}\\ \le\,& C\|\xi_{\boldsymbol{p}}\|_{L^2(J;L^2(\Omega))}+C\|\eta_{y,t}\|_{L^2(J;H^{-1}(\Omega))}+C\|\eta_{z,t}\|_{L^2(J;H^{-1}(\Omega))}+C\|\xi_z(T)\|\,.\end{aligned} \tag{3.69}$$

From Eqs. (3.38) and (3.40),

$$\begin{aligned}&(\alpha\xi_{\boldsymbol{q}},\boldsymbol{v})-(\xi_z,\mathrm{div}\boldsymbol{v})\\ =&-(\xi_{\boldsymbol{p}},\boldsymbol{v})=-(\alpha\xi_{\boldsymbol{p}},A\boldsymbol{v})=-\big(\xi_y,\mathrm{div}(A\boldsymbol{v})\big)\,.\end{aligned} \tag{3.70}$$

On differentiating Eq. (3.70) with respect to t, we get

$$(\alpha\xi_{\boldsymbol{q},t},\boldsymbol{v})-(\xi_{z,t},\mathrm{div}\boldsymbol{v})=-\big(\xi_{y,t},\mathrm{div}(A\boldsymbol{v})\big)\ \ \forall\ \boldsymbol{v}\in\boldsymbol{V}\,. \tag{3.71}$$

Selecting $\boldsymbol{v}=-\xi_{\boldsymbol{q}}$ and $w=-\xi_{z,t}$ as the test functions and adding the two relations (3.40) and (3.41), we obtain

$$\|\xi_z\|^2-\frac{1}{2}\frac{d}{dt}\|\alpha^{\frac{1}{2}}\xi_{\boldsymbol{q}}\|^2=\big(\xi_{y,t},\mathrm{div}(A\xi_{\boldsymbol{q}})\big)+(\eta_{z,t},\xi_{z,t})-(\xi_y,\xi_{z,t})\,. \tag{3.72}$$

Now from the ϵ-Cauchy inequality we know that

$$\big(\xi_{y,t},\mathrm{div}(A\xi_{\boldsymbol{q}})\big)\le C\big(\|\xi_{y,t}\|^2+\|\xi_{\boldsymbol{q}}\|^2\big)+\frac{1}{8}\|\mathrm{div}\xi_{\boldsymbol{q}}\|^2\,; \tag{3.73}$$

and on invoking (3.73) in Eq. (3.72) and using the ϵ-Cauchy inequality, we get

$$\begin{aligned}&\|\xi_{z,t}\|^2-\frac{1}{2}\frac{d}{dt}\|\alpha^{\frac{1}{2}}\xi_{\boldsymbol{q}}\|^2\\ \le\,& C\big(\|\xi_{y,t}\|^2+\|\eta_{z,t}\|^2+\|\xi_y\|^2+\|\xi_{\boldsymbol{q}}\|^2\big)+\frac{1}{8}\|\xi_{z,t}\|^2+\frac{1}{8}\|\mathrm{div}\xi_{\boldsymbol{q}}\|^2\,.\end{aligned} \tag{3.74}$$

Choosing $w=\mathrm{div}\xi_{\boldsymbol{q}}$ as a test function in Eq. (3.41) and using the ϵ-Cauchy inequality, we find that

$$\int_t^T\|\mathrm{div}\xi_{\boldsymbol{q}}\|^2ds\le 4\int_t^T\big(\|\xi_{z,t}\|^2+\|\xi_y\|^2+\|\eta_{z,t}\|^2\big)\,ds\,. \tag{3.75}$$

Integrating (3.74) with respect to time from t to T and invoking (3.75), we arrive at

$$\begin{aligned}&\int_t^T \|\xi_{z,t}\|^2 ds + \|\alpha^{\frac{1}{2}}\xi_{\boldsymbol{q}}\|^2 \\ &\le C\int_t^T \left(\|\xi_{y,t}\|^2 + \|\eta_{y,t}\|^2 + \|\xi_y\|^2 + \|\xi_{\boldsymbol{q}}\|^2\right) ds + C\|\xi_{\boldsymbol{q}}(T)\|^2 . \end{aligned} \tag{3.76}$$

From the assumption on A and Gronwall's Lemma [40] we then obtain

$$\begin{aligned}&\|\xi_{z,t}\|_{L^2(J;L^2(\Omega))} + \|\xi_{\boldsymbol{q}}\|_{L^\infty(J;L^2(\Omega))} \\ &\le C\left(\|\xi_y\|_{H^1(J;L^2(\Omega))} + \|\eta_{y,t}\|_{L^2(J;L^2(\Omega))} + \|\xi_{\boldsymbol{q}}(T)\|^2\right), \end{aligned} \tag{3.77}$$

so (3.69), (3.75) and (3.77) together with Lemma 3.4 completes the proof. □

We now briefly consider the use of mixed elliptic reconstructions combined with parabolic duality arguments to derive the estimates for $\|\xi_y\|_{L^2(J;L^2(\Omega))}$, $\|\xi_y\|_{L^\infty(J;L^2(\Omega))}$, $\|\xi_{\boldsymbol{p}}\|_{L^2(J;L^2(\Omega))}$, $\|\xi_z\|_{L^2(J;L^2(\Omega))}$, $\|\xi_z\|_{L^\infty(J;L^2(\Omega))}$ and $\|\xi_{\boldsymbol{q}}\|_{L^2(J;L^2(\Omega))}$.

Here we only prove the estimates for $\|\xi_y\|_{L^\infty(J;L^2(\Omega))}$ and $\|\xi_z\|_{L^\infty(J;L^2(\Omega))}$, for which we need the following well known stability results for the following dual equations (cf. Ref. [14] for details):

$$\begin{cases} \phi_t - \mathrm{div}(A\nabla\phi) = 0\,, & x\in\Omega, t\in[t^*,T]\,, \\ \phi|_{\partial\Omega} = 0\,, & t\in[t^*,T]\,, \\ \phi(x,t^*) = \phi_0(x)\,, & x\in\Omega\,, \end{cases} \tag{3.78}$$

and

$$\begin{cases} -\psi_t - \mathrm{div}(A\nabla\psi) = 0\,, & x\in\Omega, t\in[0,t^*]\,, \\ \psi|_{\partial\Omega} = 0\,, & t\in[0,t^*]\,, \\ \psi(x,t^*) = \psi_0(x)\,, & x\in\Omega\,. \end{cases} \tag{3.79}$$

Lemma 3.6 ([14]). *Let ϕ and ψ be the solutions of* (3.78) *and* (3.79)*, respectively. Let Ω be a convex domain. Then*

$$\begin{aligned}&\int_\Omega |\phi(x,t)|^2 dx \le C\|\phi_0\|^2_{L^2(\Omega)}, \quad \forall\ t\in[t^*,T]\,, \\ &\int_{t^*}^T\int_\Omega |\nabla\phi|^2 dx\,dt \le \|\phi_0\|^2_{L^2(\Omega)}\,, \\ &\int_{t^*}^T\int_\Omega |t-t^*||D^2\phi|^2 dx\,dt \le \|\phi_0\|^2_{L^2(\Omega)}\,, \\ &\int_{t^*}^T\int_\Omega |t-t^*||\phi_t|^2 dx\,dt \le \|\phi_0\|^2_{L^2(\Omega)}\,,\end{aligned}$$

and

$$
\begin{aligned}
&\int_\Omega |\psi(x,t)|^2 dx \le C\|\psi_0\|^2_{L^2(\Omega)}, \quad \forall\ t\in[0,t^*]\,,\\
&\int_0^{t^*}\int_\Omega |\nabla\psi|^2 dx\,dt \le \|\psi_0\|^2_{L^2(\Omega)}\,,\\
&\int_0^{t^*}\int_\Omega |t-t^*||D^2\psi|^2 dx\,dt \le \|\psi_0\|^2_{L^2(\Omega)}\,,\\
&\int_0^{t^*}\int_\Omega |t-t^*||\psi_t|^2 dx\,dt \le \|\psi_0\|^2_{L^2(\Omega)}\,,
\end{aligned}
$$

where $|D^2\phi| = \max\{|\partial^2\phi/\partial x_i\partial x_j|, 1\le i,j\le 2\}$, *and* $|D^2\psi|$ *is defined similarly.*

Lemma 3.7. *Let* ξ_y, $\xi_{\boldsymbol{p}}$, ξ_z *and* $\xi_{\boldsymbol{q}}$ *satisfy the error equations* (3.38)-(3.41). *Then the estimates* (3.46) *and* (3.62) *hold. Moreover, for* $\tau\in(0,t^*)$ *and* $\rho\in(t^*,T)$ *we have*

$$
\begin{aligned}
\|\xi_y(t^*)\| \le& C\left(1+\left(\ln\left(\frac{t^*}{\tau}\right)\right)^{\frac{1}{2}}\right)\max_{0\le t\le t^*}\|\eta_y(t)\| + C\|e_y(0)\|\\
&+C\|\eta_{y,t}\|_{L^2(t^*-\tau,t^*;H^{-1}(\Omega))}\,, \qquad (3.80)\\
\|\xi_z(t^*)\| \le& C\left(1+\left(\ln\left(\frac{T-t^*}{\rho-t^*}\right)\right)^{\frac{1}{2}}\right)\max_{t^*\le t\le T}\|\eta_z(t)\| + C\|\xi_{\boldsymbol{p}}\|_{L^2(t^*,T;L^2(\Omega))}\\
&+C\|\xi_y\|_{L^2(t^*,T;L^2(\Omega))}+C\|\eta_{z,t}\|_{L^2(t^*,\rho;H^{-1}(\Omega))}\,. \qquad (3.81)
\end{aligned}
$$

Proof. Let ψ is the solution of (3.78) with $\psi_0(x)=\xi_y(x,t^*)$. Then it follows from Eqs. (3.38) and (3.39) that

$$
\begin{aligned}
&\|\xi_y(t^*)\|^2_{L^2(\Omega)} = \big(\xi_y(t^*),\psi(t^*)\big)\\
=&\int_0^{t^*}\big((\xi_{y,t},\psi)-(\xi_y,\mathrm{div}(A\nabla\psi))\big)\,dt + \big(\xi_y(0),\psi(0)\big)\\
=&\int_0^{t^*}\big((\eta_{y,t},\psi)-(\mathrm{div}\xi_{\boldsymbol{p}},\psi)\big)\,dt + \big(\xi_y(0),\psi(0)\big) - \int_0^{t^*}\big((\xi_y,\mathrm{div}(A\nabla\psi))\big)\,dt\\
=&\int_0^{t^*}(\eta_{y,t},\psi)\,dt + \big(\xi_y(0),\psi(0)\big)\,. \qquad (3.82)
\end{aligned}
$$

From Eq. (3.82) and the stability result in Lemma 3.6, we directly obtain the estimate

$$
\|\xi_y(t^*)\|^2_{L^2(\Omega)} \le C\|\eta_{y,t}\|^2_{L^2(0,t^*;H^{-1}(\Omega))} + C\|\xi_y(0)\|^2\,, \qquad (3.83)
$$

which conforms to our estimate (3.46). For $\tau\in(0,t^*)$, let us rewrite the first term on the right-hand side of Eq. (3.82) as

$$\begin{aligned}&\int_0^{t^*}(\eta_{y,t},\psi)\,dt\\ =&\big(\eta_y(t^*-\tau),\psi(t^*-\tau)\big)-\big(\eta_y(0),\psi(0)\big)-\int_0^{t^*-\tau}(\eta_y,\psi_t)\,dt+\int_{t^*-\tau}^{t^*}(\eta_{y,t},\psi)\,dt\,. \end{aligned}\tag{3.84}$$

For the third term on the right-hand side of Eq. (3.84), we note that

$$\int_0^{t^*-\tau}(\eta_y,\psi_t)\,dt\le\left(\ln\left(\frac{t^*}{\tau}\right)\right)^{\frac12}\max_{0\le t\le t^*}\|\eta_y(t)\|\left(\int_0^{t^*}(t^*-t)\|\psi_t(t)\|^2\right)^{\frac12}, \tag{3.85}$$

and the estimate of the last term on the right-hand side of Eq. (3.84) is

$$\int_{t^*-\tau}^{t^*}(\eta_{y,t},\psi)\,dt\le C\|\eta_{y,t}\|_{L^2(t^*-\tau,t^*;H^{-1}(\Omega))}\|\nabla\psi\|_{L^2(0,t^*;L^2(\Omega))}\,, \tag{3.86}$$

so together with the stability estimates in Lemma 3.6 we obtain (3.80).

Similarly, let ϕ be the solution of (3.79) with $\phi_0(x)=\xi_z(x,t^*)$. Then it follows from Eqs. (3.40) and (3.41) that

$$\begin{aligned}&\|\xi_z(t^*)\|^2_{L^2}=\big(\xi_z(t^*),\phi(t^*)\big)\\ =&\int_{t^*}^T\big(-(\xi_{z,t},\phi)-(\xi_z,\mathrm{div}(A\nabla\phi))\big)\,dt+\big(\xi_z(T),\phi(T)\big)\\ =&\int_{t^*}^T\big((\xi_y,\phi)-(\eta_{z,t},\phi)-(\mathrm{div}\xi_{\boldsymbol q},\phi)\big)\,dt+\big(\xi_z(T),\phi(T)\big)-\int_{t^*}^T(\alpha\xi_{\boldsymbol q}+\xi_{\boldsymbol p},A\nabla\phi)\,dt\\ =&\int_{t^*}^T\big((\xi_y,\phi)-(\eta_{z,t},\phi)\big)\,dt+(\xi_z(T),\phi(T))-\int_{t^*}^T(\xi_{\boldsymbol p},A\nabla\phi)\,dt\,, \end{aligned}\tag{3.87}$$

so from Lemma 3.6 we get

$$\|\xi_z(t^*)\|^2_{L^2(\Omega)}\le C\int_{t^*}^T\big(\|\xi_y\|^2+\|\xi_{\boldsymbol p}\|^2+\|\eta_{z,t}\|^2_{-1}\big)\,dt+C\|\xi_z(T)\|^2\,. \tag{3.88}$$

Similar to the above consideration of Eq. (3.84), for $\rho\in(t^*,T)$ we rewrite the second term on the right-hand side of Eq. (3.87) — i.e.

$$\begin{aligned}&-\int_{t^*}^T(\eta_{z,t},\phi)\,dt\\ =&\big(\eta_z(\rho),\phi(\rho)\big)-\big(\eta_z(T),\phi(T)\big)-\int_{t^*}^{\rho}(\eta_{z,t},\phi)\,dt+\int_\rho^T(\eta_z,\phi_t)\,dt\,. \end{aligned}\tag{3.89}$$

For the last term on the right-hand side of Eq. (3.89), we note that

$$\int_{\rho}^{T}(\eta_z,\phi_t)\,dt \le \left(\ln\left(\frac{T-t^*}{\rho-t^*}\right)\right)^{\frac{1}{2}} \max_{t^*\le t\le T}\|\eta_z(t)\| \left(\int_{t^*}^{T}(t-t^*)\|\phi_t(t)\|^2\right)^{\frac{1}{2}}, \qquad (3.90)$$

so from (3.87), (3.89) and (3.90) together with Lemma 3.6 we obtain (3.81). □

From Eqs. (2.37)-(2.40), we obtain the error equations

$$
\begin{aligned}
&(\alpha\eta_{\boldsymbol{p}},\boldsymbol{v}_h)-(\eta_y,\mathrm{div}\boldsymbol{v}_h)=0 && \forall\ \boldsymbol{v}_h\in\boldsymbol{V}_h\,, && (3.91)\\
&(\mathrm{div}\eta_{\boldsymbol{p}},w_h)=0 && \forall\ w_h\in W_h\,, && (3.92)\\
&(\alpha\eta_{\boldsymbol{q}},\boldsymbol{v}_h)-(\eta_z,\mathrm{div}\boldsymbol{v}_h)=-(\eta_{\boldsymbol{p}},\boldsymbol{v}_h) && \forall\ \boldsymbol{v}_h\in\boldsymbol{V}_h\,, && (3.93)\\
&(\mathrm{div}\eta_{\boldsymbol{q}},w_h)=(\eta_y,w_h) && \forall\ w_h\in W_h\,. && (3.94)
\end{aligned}
$$

To prove the main theorem, we need to establish relevant *a posteriori* estimates of η_y, $\eta_{y,t}$, $\eta_{\boldsymbol{p}}$, $\mathrm{div}\eta_{\boldsymbol{p}}$, η_z, $\eta_{z,t}$, $\eta_{\boldsymbol{q}}$ and $\mathrm{div}\eta_{\boldsymbol{q}}$ for the mixed elliptic reconstructions (2.37)-(2.40) as follows.

Lemma 3.8. *For Raviart-Thomas elements, there exists a positive constant C which depends only on the coefficient matrix A, the domain Ω, the shape regularity of the elements and polynomial degree k such that for* $l=0,1$

$$
\begin{aligned}
\|\eta_y\|_{-l}^2 \le\ & C\left(\|h^{1+l}\mathbf{r}_2\|^2+\min_{w_h\in W_h}\left\|h^{1+l}(\alpha\boldsymbol{p}_h-\nabla_h w_h)\right\|^2\right)\\
\le\ & C\left(\left\|h^{1+l}(y_{h,t}+\nabla\cdot\boldsymbol{p}_h-f-u_h)\right\|^2+\min_{w_h\in W_h}\left\|h^{1+l}(\alpha\boldsymbol{p}_h-\nabla_h w_h)\right\|^2\right), \qquad (3.95)\\
\|\eta_{y,t}\|_{-l}^2 \le\ & C\left(\|h^{1+l}\mathbf{r}_2\|^2+\min_{w_h\in W_h}\left\|h^{1+l}(\alpha\boldsymbol{p}_{h,t}-\nabla_h w_h)\right\|^2\right)\\
\le\ & C\left(\left\|h^{1+l}(y_{h,t}+\nabla\cdot\boldsymbol{p}_h-f-u_h)_t\right\|^2+\min_{w_h\in W_h}\left\|h^{1+l}(\alpha\boldsymbol{p}_{h,t}-\nabla_h w_h)\right\|^2\right), \qquad (3.96)\\
\|\eta_{\boldsymbol{p}}\|_{-l}^2 \le\ & C\left(\|h\mathbf{r}_2\|^2+\left\|h^{\frac{1}{2}}J(\alpha\boldsymbol{p}_h\cdot\mathbf{t})\right\|_{0,\Gamma_h}^2+\left\|hcurl_h(\alpha\boldsymbol{p}_h)\right\|^2\right)\\
\le\ & C\Big(\left\|h^{\frac{1}{2}}J(\alpha\boldsymbol{p}_h\cdot\mathbf{t})\right\|_{0,\Gamma_h}^2+\left\|hcurl_h(\alpha\boldsymbol{p}_h)\right\|^2\\
&+\left\|h(y_{h,t}+\nabla\cdot\boldsymbol{p}_h-f-u_h)\right\|^2\Big), \qquad (3.97)\\
\|\eta_{\boldsymbol{p},t}\|_{-l}^2 \le\ & C\left(\|h\mathbf{r}_{2,t}\|^2+\left\|h^{\frac{1}{2}}J(\alpha\boldsymbol{p}_{h,t}\cdot\mathbf{t})\right\|_{0,\Gamma_h}^2+\left\|hcurl_h(\alpha\boldsymbol{p}_{h,t})\right\|^2\right)\\
\le\ & C\Big(\left\|h^{\frac{1}{2}}J(\alpha\boldsymbol{p}_{h,t}\cdot\mathbf{t})\right\|_{0,\Gamma_h}^2+\left\|hcurl_h(\alpha\boldsymbol{p}_{h,t})\right\|^2\\
&+\left\|h(y_{h,t}+\nabla\cdot\boldsymbol{p}_h-f-u_h)_t\right\|^2\Big), \qquad (3.98)\\
\|div\eta_{\boldsymbol{p}}\|^2 \le\ & C\|y_{h,t}+\nabla\cdot\boldsymbol{p}_h-f-u_h\|^2, \qquad (3.99)
\end{aligned}
$$

$$
\begin{aligned}
\|\eta_z\|_{-l}^2 \le\ & C\Big(\min_{w_h\in W_h}\left\|h^{1+l}(\alpha\boldsymbol{q}_h+\boldsymbol{p}_h-\boldsymbol{p}_d-\nabla_h w_h)\right\|^2+\|h^{1+l}\mathbf{r}_4\|^2+\|\eta_y\|_{-l}^2+\|\eta_{\boldsymbol{p}}\|^2\Big)\\
\le\ & C\Big(\left\|h^{1+l}(z_{h,t}-\nabla\cdot\boldsymbol{q}_h+y_h-y_d)\right\|^2+\|\eta_y\|_{-l}^2+\|\eta_{\boldsymbol{p}}\|^2\\
&+\min_{w_h\in W_h}\left\|h^{1+l}(\alpha\boldsymbol{q}_h+\boldsymbol{p}_h-\boldsymbol{p}_d-\nabla_h w_h)\right\|^2\Big), \qquad (3.100)\\
\|\eta_{z,t}\|_{-l}^2 \le\ & C\Big(\min_{w_h\in W_h}\left\|h^{1+l}(\alpha\boldsymbol{q}_{h,t}+\boldsymbol{p}_{h,t}-\boldsymbol{p}_{d,t}-\nabla_h w_h)\right\|^2+\|h^{1+l}\mathbf{r}_{4,t}\|^2+\|\eta_{y,t}\|_{-l}^2\\
&+\|\eta_{\boldsymbol{p},t}\|^2\Big)\\
\le\ & C\Big(\left\|h^{1+l}(z_{h,t}-\nabla\cdot\boldsymbol{q}_h+y_h-y_d)_t\right\|^2+\|\eta_{y,t}\|_{-l}^2+\|\eta_{\boldsymbol{p},t}\|^2\\
&+\min_{w_h\in W_h}\left\|h^{1+l}(\alpha\boldsymbol{q}_{h,t}+\boldsymbol{p}_{h,t}-\boldsymbol{p}_{d,t}-\nabla_h w_h)\right\|^2\Big), \qquad (3.101)\\
\|\eta_{\boldsymbol{q}}\|^2 \le\ & C\Big(\|h\mathbf{r}_2\|^2+\left\|h\,curl_h(\alpha\boldsymbol{q}_h+\boldsymbol{p}_h-\boldsymbol{p}_d)\right\|^2+\|\eta_y\|_{-l}^2\\
&+\left\|h^{\frac12}J((\alpha\boldsymbol{q}_h+\boldsymbol{p}_h-\boldsymbol{p}_d)\cdot\mathbf{t})\right\|_{0,\Gamma_h}^2+\|\eta_{\boldsymbol{p}}\|^2\Big)\\
\le\ & C\Big(\left\|h(z_{h,t}-\nabla\cdot\boldsymbol{q}_h+y_h-y_d)\right\|^2+\|\eta_y\|_{-l}^2+\left\|h^{\frac12}J((\alpha\boldsymbol{q}_h+\boldsymbol{p}_h-\boldsymbol{p}_d)\cdot\mathbf{t})\right\|_{0,\Gamma_h}^2\\
&+\|\eta_{\boldsymbol{p}}\|^2+\left\|h\,curl_h(\alpha\boldsymbol{q}_h+\boldsymbol{p}_h-\boldsymbol{p}_d)\right\|^2\Big), \qquad (3.102)\\
\|div\eta_{\boldsymbol{q}}\|^2 \le\ & C\|z_{h,t}-\nabla\cdot\boldsymbol{q}_h+y_h-y_d\|^2+C\|\eta_y\|^2, \qquad (3.103)
\end{aligned}
$$

where $J(\boldsymbol{v}\cdot\mathbf{t})$ denotes the jump of $\boldsymbol{v}\cdot\mathbf{t}$ across an element edge E for all $\boldsymbol{v}\in\boldsymbol{V}$ with $\mathbf{t}$ denoting the tangential unit vector along the edge $E\in\Gamma_h$.

Proof. Based on the tools developed in Refs. [10, 11], it is straightforward to derive *a posteriori* error estimates for η_y, $\eta_{\boldsymbol{p}}$, div$\eta_{\boldsymbol{p}}$, η_z, $\eta_{\boldsymbol{q}}$ and div$\eta_{\boldsymbol{q}}$. Consequently, we only need discuss results for the negative norm estimates η_y and η_z, and we appeal to Aubin-Nitsche duality arguments. Thus we consider $\Phi\in H_0^1(\Omega)\cap H^{2+l}(\Omega)$ as the solution of the elliptic problem:

$$-\nabla\cdot(A\nabla\Phi)=\Psi,\quad \text{in }\Omega \qquad (3.104)$$

that satisfies the elliptic regularity result

$$\|\Phi\|_{2+l}\le C\|\Psi\|_l\,. \qquad (3.105)$$

From Eq. (3.93) and the definition of Π_h, on integrating by parts and invoking the property

(3.42) we obtain

$$\begin{aligned}(\eta_z,\Psi) &= \big(\eta_z,-\nabla\cdot(A\nabla\Phi)\big)\\ &= \big(\tilde{z},-\nabla\cdot(A\nabla\Phi)\big)+\big(z_h,\nabla\cdot(A\nabla\Phi)\big)\\ &= \big(z_h,\nabla\cdot\Pi_h(A\nabla\Phi)\big)+(A\nabla\tilde{z},\nabla\Phi)\\ &= -(\tilde{\boldsymbol{q}}+A\tilde{\boldsymbol{p}}-A\boldsymbol{p}_d,\nabla\Phi)+\big(z_h,\nabla\cdot\Pi_h(A\nabla\Phi)\big)\\ &= -(\eta_{\boldsymbol{q}},\nabla\Phi)-(A\eta_{\boldsymbol{p}},\nabla\Phi)+\big(z_h,\nabla\cdot\Pi_h(A\nabla\Phi)\big)-(\alpha\boldsymbol{q}_h+\boldsymbol{p}_h-\boldsymbol{p}_d,A\nabla\Phi)\,.\end{aligned}$$

From Eq. (3.94), on integrating by parts and invoking

$$\big(\nabla_h w_h,(I-\Pi_h)(A\nabla\Phi)\big)=0\,, \tag{3.106}$$

we arrive at

$$\begin{aligned}(\eta_z,\Psi) &= (\nabla\eta_{\boldsymbol{q}},\Phi-P_h\Phi)+(\eta_y,P_h\Phi)-(A\eta_{\boldsymbol{p}},\nabla\Phi)\\ &\quad-\big(\alpha\boldsymbol{q}_h+\boldsymbol{p}_h-\boldsymbol{p}_d-\nabla_h w_h,(I-\Pi_h(A\nabla\Phi))\big)\\ &= (z_{h,t}-\nabla\cdot\boldsymbol{q}_h+y_h-y_d,\Phi-P_h\Phi)+(\eta_y,\Phi)-(A\eta_{\boldsymbol{p}},\nabla\Phi)\\ &\quad-\big(\alpha\boldsymbol{q}_h+\boldsymbol{p}_h-\boldsymbol{p}_d-\nabla_h w_h,(I-\Pi_h(A\nabla\Phi))\big)\\ &\le C\Big(\big\|h^{1+l}(z_{h,t}-\nabla\cdot\boldsymbol{q}_h+y_h-y_d)\big\|\,\|\Phi\|_{1+l}+\|\eta_y\|_{-1}\|\Phi\|_1\\ &\quad+\|A\eta_{\boldsymbol{p}}\|\|\nabla\Phi\|+\big\|h^{1+l}(\alpha\boldsymbol{q}_h+\boldsymbol{p}_h-\boldsymbol{p}_d-\nabla_h w_h)\big\|\,\|A\nabla\Phi\|_{1+l}\Big)\\ &\le C\Big(\big\|h^{1+l}(z_{h,t}-\nabla\cdot\boldsymbol{q}_h+y_h-y_d)\big\|+\|\eta_y\|_{-1}+\|\eta_{\boldsymbol{p}}\|\\ &\quad+\big\|h^{1+l}(\alpha\boldsymbol{q}_h+\boldsymbol{p}_h-\boldsymbol{p}_d-\nabla_h w_h)\big\|\Big)\|\Phi\|_{2+l}\,.\end{aligned} \tag{3.107}$$

Using the elliptic regularity (3.105) in (3.107), we obtain

$$\begin{aligned}\frac{(\eta_z,\Psi)}{\|\Psi\|_l} \le&\Big(\big\|h^{1+l}(z_{h,t}-\nabla\cdot\boldsymbol{q}_h+y_h-y_d)\big\|+\|\eta_y\|_{-1}+\|\eta_{\boldsymbol{p}}\|\\ &+\min_{w_h\in W_h}\big\|h^{1+l}(\alpha\boldsymbol{q}_h+\boldsymbol{p}_h-\boldsymbol{p}_d-\nabla_h w_h)\big\|\Big)\,,\end{aligned} \tag{3.108}$$

and hence on taking the supremum over Ψ we obtain the estimate (3.100). Following the analysis given above, the estimates for $\|\eta_{z,t}\|_{-l}$, $\|\eta_y\|_{-l}$ and $\|\eta_{y,t}\|_{-l}$ are also readily derived. Finally, from Lemma 5.1 of [11] we arrive at $\eta_{\boldsymbol{p}}$, $\eta_{\boldsymbol{p},t}$, div$\eta_{\boldsymbol{p}}$, $\eta_{\boldsymbol{q}}$ and div$\eta_{\boldsymbol{q}}$ to complete the proof. □

Remark 3.1. For a negative norm estimate (i.e. when l=1), we need to use mixed finite element spaces of index $k\ge 1$, but otherwise we can bound $\|\eta_y\|_{-1}\le C\|\eta_y\|$.

Collecting Lemmas 3.1-3.5, let us now summarise our main results as follows.

Theorem 3.1. *Let* $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ *and* $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h)$ *be the solution of* (2.5)-(2.11) *and* (2.16)-(2.22)*, respectively. Then the following a posteriori estimates hold for* $l = 0, 1$ *and* $t \in [0, T]$*:*

$$
\begin{aligned}
\|u-u_h\|^2_{L^2(J;L^2(\Omega))} \le C\Big(&\eta_1^2 + \|\eta_y(0)\|^2 + \|\eta_z(T)\|^2 + \|\eta_{y,t}\|^2_{L^2(J;H^{-1}(\Omega))}\\
&+ \|y_0 - y_0^h\|^2 + \|\eta_{z,t}\|^2_{L^2(J;H^{-1}(\Omega))} + \|\eta_z\|^2_{L^\infty(J;L^2(\Omega))}\Big),\\
\|y-y_h\|^2_{L^\infty(J;L^2(\Omega))} + \|\boldsymbol{p}-\boldsymbol{p}_h\|^2_{L^2(J;L^2(\Omega))} \le C\Big(&\|u-u_h\|^2_{L^2(J;L^2(\Omega))} + \|\eta_y(0)\|^2\\
&+ \|\eta_{\boldsymbol{p}}\|^2_{L^2(J;L^2(\Omega))} + \|\eta_y\|^2_{L^\infty(J;L^2(\Omega))} + \|y_0 - y_0^h\|^2\\
&+ \|\eta_y\|^2_{L^\infty(J;L^2(\Omega))} + \|\eta_{y,t}\|^2_{L^2(J;H^{-1}(\Omega))}\Big),\\
\|(y-y_h)_t\|^2_{L^2(J;L^2(\Omega))} + \|\boldsymbol{p}-\boldsymbol{p}_h\|^2_{L^\infty(J;L^2(\Omega))} \le C\Big(&\|\eta_{y,t}\|^2_{L^2(J;L^2(\Omega))} + \|\eta_{\boldsymbol{p}}\|^2_{L^\infty(J;L^2(\Omega))}\\
&+ \|u-u_h\|^2_{L^2(J;L^2(\Omega))} + \|A\nabla y_0 + \boldsymbol{p}_h(0)\|^2\Big),\\
\|div(\boldsymbol{p}-\boldsymbol{p}_h)\|^2_{L^2(J;L^2(\Omega))} \le C\Big(&\|u-u_h\|^2_{L^2(J;L^2(\Omega))} + \|div\eta_{\boldsymbol{p}}\|^2_{L^2(J;L^2(\Omega))}\\
&+ \|\eta_{y,t}\|^2_{L^2(J;L^2(\Omega))} + \|\eta_{\boldsymbol{p}}(0)\|^2 + \|A\nabla y_0 + \boldsymbol{p}_h(0)\|^2\Big),\\
\|z-z_h\|^2_{L^\infty(J;L^2(\Omega))} + \|\boldsymbol{q}-\boldsymbol{q}_h\|^2_{L^2(J;L^2(\Omega))} \le C\Big(&\|u-u_h\|^2_{L^2(J;L^2(\Omega))} + \|\eta_z\|^2_{L^\infty(J;L^2(\Omega))}\\
&+ \|\eta_{z,t}\|^2_{L^2(J;H^{-1}(\Omega))} + \|\eta_y\|^2_{L^\infty(J;L^2(\Omega))} + \|\eta_{\boldsymbol{q}}\|^2_{L^2(J;L^2(\Omega))} + \|y_0 - y_0^h\|^2\\
&+ \|\eta_{y,t}\|^2_{L^2(J;H^{-1}(\Omega))} + \|\eta_y(0)\|^2 + \|\eta_z(T)\|^2\Big),\\
\|(z-z_h)_t\|^2_{L^\infty(J;L^2(\Omega))} + \|\boldsymbol{q}-\boldsymbol{q}_h\|^2_{L^\infty(J;L^2(\Omega))} \le C\Big(&\|u-u_h\|^2_{L^2(J;L^2(\Omega))} + \|y_0 - y_0^h\|^2\\
&+ \|\eta_{z,t}\|^2_{L^2(J;L^2(\Omega))} + \|\eta_{\boldsymbol{p}}\|^2_{L^\infty(J;L^2(\Omega))} + \|\eta_{y,t}\|^2_{L^2(J;L^2(\Omega))} + \|\eta_y(0)\|^2\\
&+ \|\eta_{\boldsymbol{q}}\|^2_{L^\infty(J;L^2(\Omega))} + \|A\boldsymbol{p}_h(T) - A\boldsymbol{p}_d(T) + \boldsymbol{q}_h(T)\|^2\\
&+ \|\eta_{\boldsymbol{q}}(T)\|^2 + \|\eta_{\boldsymbol{p}}(0)\|^2 + \|A\nabla y_0 + \boldsymbol{p}_h(0)\|^2\Big),\\
\|div(\boldsymbol{q}-\boldsymbol{q}_h)\|^2_{L^2(J;L^2(\Omega))} \le C\Big(&\|u-u_h\|^2_{L^2(J;L^2(\Omega))} + \|\eta_y\|^2_{H^1(J;L^2(\Omega))} + \|y_0 - y_0^h\|^2\\
&+ \|\eta_{z,t}\|^2_{L^2(J;H^{-1}(\Omega))} + \|\eta_{\boldsymbol{p}}\|^2_{L^\infty(J;L^2(\Omega))} + \|div\eta_{\boldsymbol{q}}\|^2_{L^2(J;L^2(\Omega))}\\
&+ \|A\boldsymbol{p}_h(T) - A\boldsymbol{p}_d(T) + \boldsymbol{q}_h(T)\|^2 + \|\eta_{\boldsymbol{q}}(T)\|^2\\
&+ \|\eta_y(0)\|^2 + \|\eta_{\boldsymbol{p}}(0)\|^2 + \|A\nabla y_0 + \boldsymbol{p}_h(0)\|^2\Big),
\end{aligned}
$$

where η_1 *is defined in Lemma* 3.2 *and the estimates for* η_y*,* $\eta_{y,t}$*,* $\eta_{\boldsymbol{p}}$*,* $div\eta_{\boldsymbol{p}}$*,* η_z*,* $\eta_{z,t}$*,* $\eta_{\boldsymbol{q}}$ *and* $div\eta_{\boldsymbol{q}}$ *are given in Lemma* 3.8.

4. Conclusion and Future Work

We have derived *a posteriori* error estimates for semidiscrete mixed finite element solutions of quadratic optimal control problems (OCP) governed by parabolic equations. Our *a posteriori* error estimates for the linear parabolic OCP that we have obtained via mixed finite element methods seem to be new. In future, we intend to explore *a posteriori* analysis for a completely discrete mixed approximation based on the backward Euler method, and design relevant adaptive mixed finite element algorithms. Furthermore, we will also consider *a posteriori* error estimates of mixed finite element methods for hyperbolic OCP.

Acknowledgments

This work is supported by grants (91430104, 11271145) from the National Science Foundation of China, and in addition a Specialized Research Fund for the Doctoral Program of Higher Education (20114407110009).

References

[1] N. Arada, E. Casas and F.A. Tröltzsch, *Error estimates for the numerical approximation of a semilinear elliptic control problem*, Comp. Optim. Appl. **23**, 201-229 (2002).
[2] I. Babuska and T. Strouboulis, *The Finite Element Method and its Reliability*, Oxford University Press (2001).
[3] I. Babuska and S. Ohnimus, *A posteriori error estimation for the semidiscrete finite element method of parabolic partial differential equations*, Comp. Methods Appl. Mech. Eng. **190**, 65-187 (2001).
[4] I. Babuska, M. Feistauer and P. Solin, *On one approach to a posteriori error estimates for evolution problems solved by the method of lines*, Num. Math. **89**, 225-256 (2001).
[5] H. Brunner and N. Yan, *Finite element methods for optimal control problems governed by integral equations and integro-differential equations*, Num. Math. **101**, 1-27 (2005).
[6] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer, New York (1991).
[7] Y. Chen, *Superconvergence of quadratic optimal control problems by triangular mixed finite elements*, Int. J. Num. Meths. Eng. **75**, 881-898 (2008).
[8] Y. Chen, Y. Huang, W.B. Liu and N. Yan, Error estimates and superconvergence of mixed finite element methods for convex optimal control problems, J. Sc. Comp. **42**, 382-403 (2009).
[9] Y. Chen, N. Yi and W.B. Liu, *A Legendre Galerkin spectral method for optimal control problems governed by elliptic equations*, SIAM J. Num. Anal. **46**, 2254-2275 (2008).
[10] Y. Chen and W.B. Liu, *A posteriori error estimates for mixed finite element solutions of convex optimal control problems*, J. Comp. Appl. Math. **211**, 76-89 (2008).
[11] C. Carstensen, *A posteriori error estimate for the mixed finite element method*, Math. Comp. **66**, 465-476 (1997).
[12] J. Douglas and J.E. Roberts, *Global estimates for mixed finite element methods for second order elliptic equations*, Math. Comp. **44**, 39-52 (1985).
[13] A. Demlow, O. Lakkis and C. Makridakes, *A posteriori eror estimates in maximum norm for parabolic problems*, SIAM J. Num. Anal. **47**, 2157-2176 (2009).
[14] K. Eriksson and C. Johnson, *Adaptive finite elements methods for parabolic problems I: A linear model problem*, SIAM J. Num. Anal. **28**, 43-77 (1991).
[15] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems IV: Nonlinear problems*, SIAM J. Num. Anal. **32**, 1729-1749 (1995).
[16] E. Georgoulis, O. Lakkis and J.M. Virtanen, *A posteriori error control for discontinuous Galerkin methods for parabolic problems*, SIAM J. Num. Anal. **49**, 427-458 (2011).
[17] W. Gong and N. Yan, *A posteriori error estimate for boundary control problems governed by the parabolic partial differential equations*, J. Comp. Math. **27**, 68-88 (2009).

[18] L. Hou and J.C. Turner, *Analysis and finite element approximation of an optimal control problem in electrochemistry with current density controls*, Num. Math. **71**, 289-315 (1995).

[19] J. Haslinger and P. Neittaanmaki, *Finite Element Approximation for Optimal Shape Design*, John Wiley and Sons, Chichester (1989).

[20] R.H.W. Hoppe, Y. Iliash, C. Iyyunni and N.H. Sweilam, *A posteriori error estimates for adaptive finite element discretizations of boundary control problems*, J. Num. Math. **14**, 57-82 (2006).

[21] C. Johnson, Y. Nie and V. Thomée, *An a posteriori error estimate and adaptive time step control for a backward Euler discretisation of a parabolic problem*, SIAM J. Num. Anal. **27**, 277-291 (1990).

[22] G. Knowles, *Finite element approximation of parabolic time optimal control problems*, SIAM J. Control Optim. **20**, 414-427 (1982).

[23] J. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin (1971).

[24] J. Lions and E. Magenes, *Non-homogeneous Boundary Value Problems and Applications*, Grandlehre B. 181, Springer (1972).

[25] W.B. Liu, H. Ma, T. Tang and N. Yan, *A posteriori error estimates for discontinuous Galerkin time-stepping method for optimal control problems governed by parabolic equations*, SIAM J. Num. Anal. **42**, 1032-1061 (2004).

[26] W.B. Liu and N. Yan, *A posteriori error estimates for convex boundary control problems*, SIAM J. Num. Anal. **39**, 73-99 (2001).

[27] W.B. Liu and N. Yan, *A posteriori error analysis for convex distributed optimal control problems*, Adv. Comp. Math. **15**, 285-309 (2001).

[28] W.B. Liu and N. Yan, *A posteriori error estimates for optimal control problems governed by Stokes equations*, SIAM J. Num. Anal. **40**, 1850-1869 (2003).

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

[30] R. Li, W.B. Liu, H. Ma and T. Tang, *Adaptive finite element approximation of elliptic control problems*, SIAM J. Control Optim. **41**, 1321-1349 (2002).

[31] O. Lakkis and C. Makridakis, *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems*, Math. Comp. **75**, 1627-1658 (2006).

[32] R. Mcknight and W. Bosarge Jr., *The Ritz-Galerkin procedure for parabolic control problems,* SIAM J. Control Optim. **11**, 510-524 (1973).

[33] C. Makridakis and R.H. Nochetto, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Num. Anal. **41**, 1585-1594 (2003).

[34] R.H. Nochetto, G. Savaré and C. V erdi, *A posteriori error estimates for variable time step discretizations of nonlinear evolution equations*, Comm. Pure Appl. Math. **53**, 525-589 (2000).

[35] P. Neittaanmaki and D. Tiba, *Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms and Applications*, Dekker, New York (1994).

[36] D. Tiba, Lectures on the Optimal Control of Elliptic Problems, University of Jyvaskyla Press, Finland (1995).

[37] F. Tröltzsch, *Semidiscrete Ritz-Galerkin approximation of nonlinear parabolic boundary control problems-strong convergence of optimal control*, Appl. Math. Optim. **29**, 309-329 (1994).

[38] R. Verfürth, *A posteriori error estimates for nonlinear problems: $L^r(0,T;W^{1,\rho}(\Omega))$-error estimates for finite element discretisation of parabolic equations*, Num. Meth. PDE. **14**, 487-518 (1998).

[39] R. Verfürth, *A posteriori error estimates for nonlinear problems: $L^r(0,T;L^\rho(\Omega))$-error estimates for finite element discretisation of parabolic equations*, Math. Comp. **67**, 1335-1360 (1998).

[40] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer (1997).

[41] M.F. Wheeler, *A priori L_2 error estimates for Galerkin approximations to parabolic partial differential equations*, SIAM J. Num. Anal. **10**, 723-759 (1973).