FINITE ELEMENT METHOD AND ITS ERROR ESTIMATES FOR THE TIME OPTIMAL CONTROLS OF HEAT EQUATION

WEI GONG* AND NINGNING YAN†

Abstract. In this paper, we discuss the time optimal control problems governed by heat equation. The variational discretization concept is introduced for the approximation of the control, and the semi-discrete finite element method is applied for the controlled heat equation. We prove optimal a priori error estimate for the optimal time $T$, and quasi-optimal estimates for the optimal control $u$, the related state $y$ and adjoint state $p$.

Key words. Time optimal control problems, finite element method, error estimates.

1. Introduction

One of the most important optimal control problems is the optimal time control problem. There have been extensive researches on the theoretical parts of the time optimal control problems of ODEs (see, e.g., [4] and [5]) and time-dependent PDEs (see [14, 15, 17] and the references cited therein), but only a few works related to their numerical algorithms can be found, especially the finite element approximations and error estimates for PDEs, among them we should mention the work [8], [9] and [16].

The purpose of this work is to investigate the finite element approximations of the time optimal control problem governed by heat equation. The model problem that we shall investigate is the following time optimal control problem:

\[
\min_{u \in U_{ad}} \left\{ T : y(T; y_0, u) \in B(0,1) \right\},
\]

where $u$ is the control, the state $y$ satisfies the following controlled equation:

\[
\begin{cases}
\frac{\partial y(x,t)}{\partial t} - \Delta y(x,t) = Eu(x,t) & \text{in } \Omega \times (0, +\infty), \\
y(x,t) = 0 & \text{on } \partial \Omega \times (0, +\infty), \\
y(x,0) = y_0(x) & \text{in } \Omega.
\end{cases}
\]

The details will be specified in the next section.

Although the finite element approximations of PDE-constrained optimal control problems and related error estimates are well studied in the decades and there are huge literatures in this aspect (see, e.g., [3], [7], [10], [11], [12] and the references cited therein), the finite element method and its error estimates for time optimal control problems are addressed only in a few papers. To the best of our knowledge, the earliest work can be traced back to [8] and [9]. In both works, the finite element approximations are introduced for the time optimal control problems and the convergence analyses are provided. In [8] Knowles considered the finite dimensional control ($u = \sum_{i=1}^{m} f_i(t)g_i(x)$, where $g_i(x), i = 1, \cdots, m$, are given functions) which acts as the Robin boundary condition of the controlled equation and gave

*Received by the editors March 15, 2015.

2000 Mathematics Subject Classification. 49J20, 49K20, 65N15, 65N30.

This research was supported by the National Basic Research Program of China under grant 2012CB821204, the National Natural Science Foundation of China under grant 11201464, 91330115 and 11171337.
the error estimate for the optimal time with order $O(h^{\frac{3}{2} - \delta})$ for an arbitrary small $\delta > 0$. While in [9] Lasiecka proved the convergence (without orders) of the optimal time for the Dirichlet boundary control problem. In a recent work [16] by Wang and Zheng, the time optimal control problem (1)-(2) is discussed. The error estimate for the optimal time with order $O(h + l)$ is provided under some additional assumptions. These assumptions are not easy to verify in general cases. In this paper, we introduce the variational discretization concept (see, e.g., [6] and [7]) for the approximation of the control. Using this scheme, the error analysis becomes easier and the optimal error estimate for the optimal time is proved without those complicated assumptions. Moreover, the error estimates for the optimal control and the state are also provided in this paper which are not found elsewhere.

The plan of the paper is as follows. In section 2, we introduce the model time optimal control problem and construct its finite element approximation. The error estimates for the time optimal control problems are then analyzed in section 3, where the error estimates for the optimal time $T$, the optimal control $u$, the related state $y$ and adjoint state $p$ are provided. Finally, we give a conclusion to the results obtained in this paper and an outlook for some possible further works in the last section.

2. Time optimal control problem and its finite element approximation

In this section, we formulate the model optimal control problem and its finite element approximation.

Let $\Omega \subset \mathbb{R}^n$ ($n = 2$ or $3$) be a convex and bounded domain with sufficiently smooth boundary $\partial \Omega$. In the rest of the paper, we shall take the control space $U = L^\infty(0, +\infty; L^2(\omega))$ with $\omega \subset \Omega$. We use the standard norms $\| \cdot \|_{C([a,b]; L^2(\Omega))}$ and $\| \cdot \|_{L^2(a,b; L^2(\Omega))}$ for related Sobolev spaces. For simplicity, we denote by $\| \cdot \|$ and $(\cdot, \cdot)$ the usual norm and the inner product of $L^2(\Omega)$, respectively. In addition, $C$ denotes a general positive constant independent of the mesh size $h$.

Let

\[ U_{ad} = \{ v \in L^\infty(0, +\infty; L^2(\omega)) : \quad \| v(t) \|_{L^2(\omega)} \leq 1 \text{ for almost every } t \in [0, +\infty) \}, \]

\[ B(0, 1) = \{ w \in L^2(\Omega) : \quad \| w \| \leq 1 \}. \]

Then the model problem that we shall investigate is the following time optimal control problem (see [16]):

\[ \min_{u \in U_{ad}} \left\{ T : y(T; y_0, u) \in B(0, 1) \right\} \]

with $y(\cdot; y_0, u)$ the unique solution of the following equation

\[ \begin{aligned}
\frac{\partial y(x, t)}{\partial t} - \Delta y(x, t) &= Eu(x, t) \quad \text{in } \Omega \times (0, +\infty), \\
y(x, t) &= 0 \quad \text{on } \partial \Omega \times (0, +\infty), \\
y(x, 0) &= y_0(x) \quad \text{in } \Omega
\end{aligned} \]

(4) corresponding to the control $u$ and the initial value $y_0$, here

\[ E(f)(x) = \begin{cases} 
 f(x) & \text{if } x \in \Omega, \\
 0 & \text{if } x \in \Omega \setminus \omega.
\end{cases} \]

Throughout this paper, we will treat the solutions of (4) as functions of the time variable $t$, from $\mathbb{R}^+ := [0, +\infty)$ to the state space $L^2(\Omega)$. We call the number $\hat{T}(y_0) := \min_{u \in U_{ad}} \{ T : y(T; y_0, u) \in B(0, 1) \}$ the optimal time, while a control $\hat{u} \in U_{ad}$, and satisfying the property that $y(\hat{T}(y_0); y_0, \hat{u}) \in B(0, 1)$, is called an optimal control with corresponding optimal state $\hat{y} := y(\cdot; y_0, \hat{u})$. Clearly, $\hat{T}(\cdot)$ defines a
functional from $L^2(\Omega)$ to $\mathbb{R}^+$. Just for simplicity, in the following we denote $\bar{T}$, $\bar{u}$ and $\bar{y}$ by $T$, $u$ and $y$, respectively.

It is proved in [16] that the solution of the problem (3)-(4) is exist and unique. Moreover, it can be proved that when $T$ is the optimal time, we have the unique optimal control $u$ and the adjoin state $p$ such that the following Pontryagin maximum principle holds (see [16]):

\[
\begin{aligned}
&\left\{ \begin{aligned}
\frac{\partial y}{\partial t}(x, 0) + a(y, w) = (E u, w) & \quad \forall w \in H^1_0(\Omega), \quad t \in (0, T], \\
y(x, 0) = y_0(x) & \quad x \in \Omega, \\
-(\frac{\partial p}{\partial t}, q) + a(q, p) = 0 & \quad \forall q \in H^1_0(\Omega), \quad t \in [0, T), \\
p(x, T) = g(x, T) & \quad x \in \Omega, \\
\int_0^T (E^* p, v - u)_\omega dt \geq 0 & \quad \forall v \in U_{ad},
\end{aligned} \right. \\
\end{aligned}
\]

where

\[a(y, w) = (\nabla y, \nabla w)\]

and $E^*$ is the adjoint operator of $E$ such that $E^* p(x) = p(x)$ when $x \in \omega$, and $(\cdot, \cdot)_{\omega}$ is the inner product of $L^2(\omega)$. Moreover, the last inequality in (5) can be replaced by

\[u(t) = -\frac{E^* p(t)}{\|E^* p(t)\|_\omega} \quad \text{for almost every } t \in [0, T),\]

where $\| \cdot \|_\omega$ is the norm of $L^2(\omega)$. It should be pointed out that (6) is valid because $\|E^* p(t)\|_\omega > 0$ for all $t \in [0, T - \delta)$, where $\delta$ is any positive constant, and the control $u$ has bang-bang property which reaches the boundary of $U_{ad}$; see [16] for more details.

Next, we introduce an approach to approximate the problem (3)-(4) with semi-discrete finite element method (see, e.g., [2] and [12]) for the controlled equation (4) and the variational discretization concept for the control $u$ and the optimization problem (3) (see, e.g., [6] and [7]).

Let us firstly consider the finite element approximation of the controlled equation (4). Here we consider only n-simplex elements, as they are among the most widely used ones. Also we consider only piecewise linear conforming Lagrange elements. Let $\Omega^h$ be a polygonal approximation to $\Omega$ with a boundary $\partial \Omega^h$. Let $T^h$ be a partitioning of $\Omega^h$ into disjoint regular n-simplices $\tau$, so that $\Omega^h = \bigcup_{\tau \in T^h} \tau$. Let $h_{\tau}$ be the size of the element $\tau$, and $h = \max\{h_{\tau}\}$. For simplicity, assume that $\Omega^h = \Omega$. Associated with $T^h$ is a finite dimensional subspace $S^h$ of $C(\Omega^h)$, such that $\chi_{|\tau}$ are piecewise linear functions for $\forall \chi \in S^h$ and $\tau \in T^h$. Let $W^h = S^h \cap H^1_0(\Omega)$. It is easy to see that $W^h \subset H^1_0(\Omega)$.

With above preparations we can formulate the semi-discrete finite element approximation of (4) as follows:

\[
\begin{aligned}
&\left\{ \begin{aligned}
\frac{\partial w_h}{\partial t}(x, 0) + a(y_h, w_h) = (E u_h, w_h) & \quad \forall w_h \in W^h, \quad t \in (0, +\infty), \\
y_h(x, 0) = y^h_0(x) & \quad x \in \Omega,
\end{aligned} \right. \\
\end{aligned}
\]

where $y_h \in H^1(0, T; W^h)$, the initial value $y^h_0 \in W^h$ is an approximation of $y_0$.

Then the finite element approximation with variational control discretization concept for the time optimal control problem (3)-(4) is defined by

\[\min_{u \in U_{ad}} \{ T_h : \quad y_h(T_h; y^h_0, u) \in B_h(0, 1) \}\]

such that $y_h(T_h; y^h_0, u)$ is the solution of equation (7), where $B_h(0, 1) = \{ w_h \in W^h : \|w_h\| \leq 1 \}$. In this problem, we define the number $T_h(y^h_0) := \min_{u \in U_{ad}} \{ T_h :
\[ y_h(T_h; y_0^h, u) \in B_h(0,1) \] to be the optimal time for the discrete time optimal control problem (7)-(8), while the control \( \tilde{u}_h \in U_{ad} \) which satisfies the property that \( y_h(T(y_0); y_0^h, \tilde{u}_h) \in B_h(0,1) \), is called an optimal control with associated optimal state \( \tilde{y}_h := y_h(\cdot; y_0^h, \tilde{u}_h) \). Again, we use \( T_h, u_h \) and \( y_h \) to denote \( \tilde{T}_h(y_0^h), \tilde{u}_h \) and \( \tilde{y}_h \), respectively.

It should be noticed that the approach (7)-(8) provided in this paper is different from the approach presented in [16]. Here we use the variational discretization concept to approximate the optimization problem (3). Instead of discretising the control space such that \( u_t \in \tilde{U}_{ad} \), \( t \) is the mesh size for the triangulation of \( \omega \), we require that \( u_h \in U_{ad} \). Then \( u_h \) might be an approximated function, but not necessarily in the finite element space \( W^h \), especially when the restriction of \( T^h \) on \( \omega \) does not give a triangulation of \( \omega \).

Again, it can be proved that the problem (7)-(8) admits a unique optimal control \( u_h \in U_{ad} \). Moreover, it can be proved that when \( T_h \) is the optimal time for the discrete problem (7)-(8), we have the unique optimal control \( u_h \) and the discrete adjoint state \( p_h \) such that

\[
\begin{aligned}
&\begin{cases}
\left( \frac{\partial y_h}{\partial t}, w_h \right) + a(y_h, w_h) = \langle E u_h, w_h \rangle \\
y_h(0, x) = y_0^h(x) \\
\int_0^{T_h} (E^* p_h, v - u_h) \omega dt \geq 0
\end{cases} & \forall w_h \in W^h, \ x \in \Omega, \\
&\begin{cases}
-(\frac{\partial y_h}{\partial t}, y_h) + a(y_h, p_h) = 0 \\
p_h(x, T_h) = y_h(x, T_h)
\end{cases} & \forall y_h \in W^h, \ t \in [0, T_h), \ x \in \Omega, \\
&\text{if } \|E^* p_h(t)\|_\omega \neq \text{if } \|E^* p_h(t)\|_\omega = 0.
\end{aligned}
\]

For the discrete optimal control \( u_h \) we actually have the following observation: for each \( t \in [0, T_h) \) there holds

\[ u_h(t) = \begin{cases} 
-\frac{E^* p_h(t)}{\|E^* p_h(t)\|_\omega} \text{ if } \|E^* p_h(t)\|_\omega \neq 0, \\
E^* v_h \text{ for } \forall v_h \in W^h \text{ satisfying } \|v_h\|_\omega \leq 1 \text{ if } \|E^* p_h(t)\|_\omega = 0.
\end{cases} \]

3. Error analysis

In this section, we will discuss the error estimate of the time optimal control problem (3)-(4) and its finite element approximation (7)-(8).

Firstly, let us introduce two lemmas, which are the standard results of finite element analysis. The proof of these lemmas can be found in many references, see, e.g., [1], [2], [12] and [13].

**Lemma 3.1.** Let \( y(\cdot; y_0, u) \) be the solution of equation (4) and \( y_h(\cdot; y_0^h, u) \) be the standard finite element approximation of \( y(\cdot; y_0, u) \) defined in (7). Assume that \( y_0 \in H^1_0(\Omega), u \in L^2(0,T; L^2(\omega)), \Omega \) is a convex domain or with smooth boundary. Then we have

\[ \|y(\cdot; y_0, u)\|_{C([0,T]; L^2(\Omega))} + \|y(\cdot; y_0, u)\|_{L^2(0,T; L^2(\Omega))} \leq C(\|y_0\|_{L^2(\Omega)} + \|E u\|_{L^2(0,T; L^2(\Omega))}) \]

and

\[ \|y(\cdot; y_0, u) - y_h(\cdot; y_0^h, u)\|_{L^2(0,T; L^2(\Omega))} + h \|y(\cdot; y_0, u) - y_h(\cdot; y_0^h, u)\|_{C([0,T]; L^2(\Omega))} \]

\[ + h \|y(\cdot; y_0, u) - y_h(\cdot; y_0^h, u)\|_{L^2(0,T; H^1(\Omega))} \leq C h^2 (\|y_0\|_{H^1(\Omega)} + \|E u\|_{L^2(0,T; L^2(\Omega))}). \]
Lemma 3.2. Let $\lambda_1$ be the first eigenvalue of the operator $-\Delta$, and $\lambda_1^h$ be the first eigenvalue of the discrete Laplace operator $-\Delta_h$, which is approximated by standard piecewise linear and continuous finite elements. Then it holds
\[
|\lambda_1 - \lambda_1^h| \leq Ch^2.
\]
Moreover, we need some regularity results for the solutions of the time optimal control problems. Since $u \in L^\infty(0, +\infty; L^2(\omega))$, it follows that $y \in L^2(0,T; H^2(\Omega) \cap H^1_0(\Omega)) \cap H^1(0,T; L^2(\Omega))$ under the assumption that $y_0 \in H^1_0(\Omega)$, which in turn implies $y \in C([0,T]; H^1_0(\Omega))$ (see [10]). From (5) we conclude that $p \in L^2(0,T; H^2(\Omega) \cap H^1_0(\Omega)) \cap H^1(0,T; L^2(\Omega))$ and $p \in C([0,T]; H^1_0(\Omega))$.

Next, we will present the property of the discrete optimal time, which has been provided in [16].

Lemma 3.3. Let $T$ and $T_h$ be the optimal times for the time optimal control problem (3)-(4) and its approximation (7)-(8), respectively. Let
\[
g(s) = \begin{cases} 
0 & s \in (0, 1], \\
\ln \frac{s}{\lambda_1^h} & s \in (1, +\infty)
\end{cases}
\]
and
\[
g_h(s) = \begin{cases} 
0 & s \in (0, 1], \\
\ln \frac{s}{\lambda_1^h} & s \in (1, +\infty),
\end{cases}
\]
where $\lambda_1$ and $\lambda_1^h$ are defined in Lemma 3.2. Then, it holds that
\[
T(y_0) \leq g(\|y_0\|)
\]
and
\[
T_h(y^h_0) \leq g(\|y^h_0\|).
\]

Now we are in the position to prove our main result: the error estimate for the approximation of the optimal time. Some of the ideas for the proof follow [16].

Theorem 3.1. Let $T$ and $T_h$ be the optimal times for the time optimal control problem (3)-(4) and its approximation (7)-(8), respectively. Assume that the conditions in Lemma 3.1 are all valid. Then
\[
|T - T_h| \leq Ch.
\]

Proof. Let us start with proving
\[
T_h - T \leq Ch.
\]
Let $u$ be the optimal control to the time optimal control problem (3)-(4) with optimal time $T$. Then it follows from (12) that
\[
\|g(T; y_0, u) - y_h(T; y^h_0, u)\| \leq Ch,
\]
where $y_h(T; y^h_0, u)$ is the solution of (7). Note that $T$ and $u$ are the optimal time and the optimal control to the time optimal control problem (3)-(4). We have that $g(T; y_0, u) \in B(0, 1)$, and therefore $\|g(T; y_0, u)\| \leq 1$. It follows that
\[
\|y_h(T; y^h_0, u)\| \leq \|g(T; y_0, u)\| + \|g(T; y_0, u) - y_h(T; y^h_0, u)\| \\
\leq 1 + Ch.
\]
If $y_h(T; y^h_0, u) \in B_h(0, 1)$, that is $y_h(T; y^h_0, u)$ takes value in $B_h(0, 1)$ at time $T$. Considering that $T_h$ is the optimal time of the discrete time optimal control problem (7)-(8), we have that $T_h \leq T$, which proves (17) in the case of $y_h(T; y^h_0, u) \in B_h(0, 1)$. 

Next, let us consider the other case that \( y_h(T; y_0^h, u) \) is outside of \( B_h(0, 1) \). Let \( y_h^* := y_h(T; y_0^h, u) \) be the initial value of the discrete time optimal control problem \((7)-(8)\), let \( T_h^* \) and \( u_h^* \) be the related optimal time and optimal control. Then the solution \( y_h(T + T_h^*; y_h^*, u_h^*) \in B_h(0, 1) \). Moreover, it follows from \((15)\) and \((18)\) that

\[
T_h^* \leq \frac{1}{\lambda_1} \ln \|y_h^*\| \leq \frac{1}{\lambda_1} \ln(1 + Ch) \leq C \frac{1}{\lambda_1} h.
\]

Using Lemma 3.2, \((19)\) implies that

\[
T_h^* \leq C \frac{1}{\lambda_1} h.
\]

We construct another control as

\[
\hat{u}(t) = \begin{cases} u(t) & t \in (0, T), \\ u_h^*(t - T) & t \in (T, +\infty). \end{cases}
\]

Then, it is easy to see that \( \hat{u} \in U_{ad} \) and \( y_h(T + T_h^*; y_0^h, \hat{u}) \in B_h(0, 1) \). Considering that \( T_h \) is the optimal time of the discrete time optimal control problem \((7)-(8)\), we have that

\[
T_h \leq T + T_h^*.
\]

Then \((17)\) follows from \((20)\) and \((21)\) immediately. In summary, \((17)\) is proved in all cases, including \( y_h(T; y_0^h, u) \) is in \( B_h(0, 1) \) and outside of \( B_h(0, 1) \).

Now, we are in the position to prove

\[
T - T_h \leq Ch.
\]

Let \( u_h \) be the optimal control to the discrete time optimal control problem \((7)-(8)\). Then it follows from \((12)\) that

\[
\|y(T_h; y_0, u_h) - y_h(T_h; y_0^h, u_h)\| \leq Ch,
\]

where \( y(T_h; y_0, u_h) \) is the solution of \((4)\) with right hand side substituted by \( Eu_h \). Considering that \( T_h \) and \( u_h \) are the optimal time and the optimal control to the discrete time optimal control problem \((7)-(8)\), we have that \( \|y_h(T_h; y_0^h, u_h)\| \leq 1 \), and therefore

\[
\|y(T_h; y_0, u_h)\| \leq \|y_h(T_h; y_0^h, u_h)\| + \|y(T_h; y_0, u_h) - y_h(T_h; y_0^h, u_h)\| \leq 1 + Ch.
\]

If \( y(T_h; y_0, u_h) \in B(0, 1) \), that is \( y(T_h; y_0, u_h) \) takes value in \( B(0, 1) \) at time \( T_h \). Considering that \( T \) is the optimal time of the time optimal control problem \((3)-(4)\), we have that \( T \leq T_h \), which proves \((22)\) in the case of \( y(T_h; y_0, u_h) \in B(0, 1) \).

Next, we consider the other case that \( y(T_h; y_0, u_h) \) is outside of \( B(0, 1) \). Let \( y^* := y(T_h; y_0, u_h) \) be the initial value of the time optimal control problem \((3)-(4)\), let \( T^* \) and \( u^* \) be the related optimal time and optimal control. Then the solution \( y(T_h + T^*; y^*, u^*) \in B(0, 1) \). Moreover, it follows from \((14)\) and \((23)\) that

\[
T^* \leq \frac{1}{\lambda_1} \ln \|y^*\| \leq \frac{1}{\lambda_1} \ln(1 + Ch) \leq C \frac{1}{\lambda_1} h.
\]

We now construct a new control

\[
\hat{u}(t) = \begin{cases} u_h(t) & t \in (0, T_h], \\ u^*(t - T_h) & t \in (T_h, +\infty). \end{cases}
\]
Then, it is easy to see that \( \hat{u} \in U_{\text{ad}} \) and \( y(T_h + T^*: y_0, \hat{u}) \in B(0, 1) \). Considering that \( T \) is the optimal time of the time optimal control problem (3)-(4), we have that

\[
T \leq T_h + T^*.
\]

Then (22) follows from (24) and (25). In summary, (22) is proved in all cases, including \( y(T_h: y_0, u_h) \) is in \( B(0, 1) \) and outside of \( B(0, 1) \).

Summing up, (16) follows from (17) and (22). This completes the proof. \( \Box \)

After deriving the error estimate between the optimal time \( T \) and the discrete optimal time \( T_h \), we are going to consider the errors between \( (y, p, u) \) and \( (y_h, p_h, u_h) \), which are the solutions of the problems (5) and (9), respectively.

**Theorem 3.2.** Let \( T \) and \( T_h \) be the optimal times for the time optimal control problem (3)-(4) and its approximation (7)-(8), let \( (y, p, u) \) and \( (y_h, p_h, u_h) \) be the solutions of problems (5) and (9), respectively. Assume that the conditions in Lemma 3.1 are all valid. Let \( \tilde{T} = \min \{T, T_h\} \). Then, for all \( t \in [0, \tilde{T}] \) and \( \varepsilon \in (0, 1) \) there hold

\[
\|y(t) - y_h(t)\| \leq C h^{1-\varepsilon}(\int_0^t \|\epsilon_t p(s)\|_\omega^2)\frac{1}{2},
\]

(26)

\[
\|p(t) - p_h(t)\| \leq \frac{C h^{1-\varepsilon}}{\|\epsilon_t p(t)\|_\omega},
\]

(27)

\[
\|u(t) - u_h(t)\|_\omega \leq C h^{1-\varepsilon}.
\]

(28)

**Proof.** To begin with, we prove an error estimate with reduced order:

\[
\|p - p_h\|_{C([0, \tilde{T}]; L^2(\Omega))} \leq C h^{\frac{1}{2}}.
\]

(29)

Let us consider two different cases.

(i) At first we consider the case \( T_h \leq T \). Let \( (y_h(u), p_h(u)) \) be the solutions of the following auxiliary equations

\[
\left\{
\begin{array}{ll}
\left( \frac{\partial y_h(u)}{\partial t}, w_h \right) + a(y_h(u), w_h) = (E u, w_h) & \forall w_h \in W^h, t \in (0, T_h], \\
y_h(u)(x, 0) = y^h_0(x) & x \in \Omega,
\end{array}
\right.
\]

\[
\left\{
\begin{array}{ll}
-(\frac{\partial p_h(u)}{\partial t}, q_h) + a(q_h, p_h(u)) = 0 & \forall q_h \in W^h, t \in [0, T_h),
\end{array}
\right.
\]

\[
\left\{
\begin{array}{ll}
p_h(u)(x, T_h) = y_h(u)(x, T_h) & x \in \Omega.
\end{array}
\right.
\]

(30)

Note that \( u \) is defined in \([0, T]\), and then (30) is well defined when \( T_h \leq T \).

Here and later we extend \( u_h \) from \([0, T_h]\) to \([0, T]\) by setting \( u_h(t) = u(t) \) when \( t \in (T_h, T] \). It is clear that \( u \) and \( u_h \) are all in the control set \( U_{\text{ad}} \). It follows from (5) and (9) that

\[
\int_0^T (E^* p, u_h - u) \omega dt \geq 0
\]

and

\[
\int_0^{T_h} (E^* p_h, u - u_h) \omega dt \geq 0.
\]
Adding the above two inequalities together leads to

\[
\int_0^{T_h} (E^*p - E^*p_h(u), u_h - u)_\omega dt + \int_0^{T_h} (E^*p, u_h - u)_\omega dt
\]

\[
= \int_0^{T_h} (p_h(u) - p_h, Eu_h - Eu)_dt
\]

\[
= \int_0^{T_h} \left( \frac{\partial(y_h - y_h(u))}{\partial t}, p_h(u) - p_h \right) + a(y_h - y_h(u), p_h(u) - p_h) \]

\[
+ (y_h(T_h) - y_h(u)(T_h), y_h(u)(T_h) - y_h(T_h)) \]

\[
= \int_0^{T_h} (E^*(p - p_h(y)), u_h - u)_\omega dt
\]

\[
\leq \int_0^{T_h} (E^*(p - p_h(y)), u_h - u)_\omega dt
\]

\[
\leq \int_0^{T_h} (E^*(p - p_h(y)), u_h - u)_\omega dt + \int_0^{T_h} (E^*(p_h(y) - p_h(u)), u_h - u)_\omega dt
\]

(34) and (35) imply that

\[
\|y_h(T_h) - y_h(u)(T_h)\|^2 \leq \int_0^{T_h} (E^*(p - p_h(u)), u_h - u)_\omega dt
\]

\[
\leq \int_0^{T_h} (E^*(p - p_h(y)), u_h - u)_\omega dt
\]

\[
\leq \int_0^{T_h} (E^*(p - p_h(y)), u_h - u)_\omega dt
\]

where \(p_h(y)\) is the solution of the following equation

\[
\begin{cases}
-\frac{\partial p_h(y)}{\partial t} + a(q_h, p_h(y)) = 0 & \forall q_h \in W^h, \ t \in [0, T_h), \\
p_h(y)(x, T_h) = P_h y(x, T_h) & x \in \Omega
\end{cases}
\]

with \(P_h y \in W^h\) an appropriate approximation of \(y\). It is easy to see that \(p_h(y)\) is the standard finite element approximation of \(\tilde{p}\), where \(\tilde{p}\) is the solution of the equation

\[
\begin{cases}
-(\frac{\partial \tilde{p}}{\partial t}, q) + a(q, \tilde{p}) = 0 & \forall q \in H_0^1(\Omega), \ t \in [0, T_h), \\
\tilde{p}(x, T_h) = y(x, T_h) & x \in \Omega
\end{cases}
\]

and it follows from (12) that

\[
\|\tilde{p} - p_h(y)\|_{C([0, T_h]; L^2(\Omega))} \leq C h.
\]
Similar to the proof of (35) we have
\[
\int_0^{T_h} (E^*(p_h(y) - p_h(u)), u_h - u) dt
= \int_0^{T_h} (p_h(y) - p_h(u), E(u_h - u)) dt
= (y_h(T_h) - y_h(u)(T_h), P_h y(T_h) - y_h(u)(T_h)).
\]
(40)

Then (36), (40) and \(\varepsilon\)-Young inequality imply that
\[
\|y_h(T_h) - y_h(u)(T_h)\| \leq C\|P_h y(T_h) - y_h(u)(T_h)\| + C\int_0^{T_h} (E^*(p - p_h(y)), u_h - u) dt.
\]
(41)

It is obvious that \(y(T_h) \in H^1_0(\Omega)\). If we choose \(P_h y(T_h)\) as the \(L^2\) projection of \(y(T_h)\) onto \(W^h\), we can conclude from the standard projection error estimate ([2]) and (12) that
\[
\|P_h y(T_h) - y_h(u)(T_h)\| \leq \|P_h y(T_h) - y(T_h)\| + \|y(T_h) - y_h(u)(T_h)\| \leq C h.
\]
(42)

Note that \(u, u_h \in U_{ad}\) and \(p, p_h(u) \in L^2(0, T_h; L^2(\Omega))\). Then (39) implies that
\[
\int_0^{T_h} (E^*(p - p_h(y)), u_h - u) dt
\leq \|p - p_h(y)\|_{L^2(0,T_h;L^2(\omega))}\|u - u_h\|_{L^2(0,T_h,L^2(\omega))}
\leq C\|p - p_h(y)\|_{L^2(0,T_h,L^2(\omega))}
\leq C\|\tilde{p} - p\|_{L^2(0,T_h,L^2(\omega))} + C\|\tilde{p} - p_h(y)\|_{L^2(0,T_h,L^2(\omega))}
\leq C\|\tilde{p} - p\|_{L^2(0,T_h,L^2(\omega))} + Ch^2.
\]
(43)

Moreover, it can be shown from (11) that
\[
\|\tilde{p} - p\|_{L^2(0,T_h,L^2(\omega))} + \|\tilde{p} - p\|_{C([0,T_h];L^2(\Omega))} \leq C\|\tilde{p}(T_h) - p(T_h)\|.
\]
(44)

When \(T_h \leq T\) and \(h\) is small enough, the initial conditions in (5), (38) and the error estimate (16) imply that
\[
\|\tilde{p}(T_h) - p(T_h)\| = \|y(T_h) - p(T_h)\|
\leq \|y(T_h) - e^{\Delta(T - T_h)} y(T)\|
\leq \|y(T_h) - e^{\Delta(T - T_h)}(e^{\Delta(T - T_h)} y(T_h) + \int_0^T e^{\Delta(T-s)} E u(s) ds)\|
\leq \|(1 - e^{2\Delta(T - T_h)}) y(T_h)\| + \int_0^T e^{\Delta(T-s)} E u(s) ds\|
\leq \|1 - e^{2\Delta(T - T_h)}\| \|y(T_h)\| + \int_0^T e^{-\lambda_1(T-s)} ds
\leq C\lambda_1(T - T_h)\|y(T_h)\| + \frac{C}{\lambda_1}(T - T_h)
\leq C|T - T_h| \leq C h.
\]
(45)

It follows from (44) and (45) that
\[
\|\tilde{p} - p\|_{L^2(0,T_h,L^2(\omega))} + \|\tilde{p} - p\|_{C([0,T_h];L^2(\Omega))} \leq C h.
\]
(46)
Therefore, we can conclude from (41), (43), (42) and (46) that
\[
\|y_h(T_h) - y_h(u)(T_h)\| \leq C h^{\frac{1}{2}}.
\]
It is easy to see from (12) that
\[
\|y(T_h) - y_h(u)(T_h)\| \leq C h.
\]
Then (47) and (48) imply that
\[
\|y(T_h) - y_h(T_h)\| \leq C h^{\frac{1}{2}}.
\]
We can then conclude from (9), (11), (37) and (49) that
\[
\|p_h - p_h(y)\|_{C([0,T_h];L^2(\Omega))} \leq C\|p_h(T_h) - p_h(y)(T_h)\| \leq C\|y_h(T_h) - P_h(y(T_h))\| \leq C h^{\frac{1}{2}}.
\]
Thus, it follows from (39), (46) and (50) that
\[
\|p - p_h\|_{C([0,T_h];L^2(\Omega))} \leq C h^{\frac{1}{2}},
\]
which proves the error estimate (29) with the case $T_h \leq T$.

(ii) Next we consider the case $T < T_h$. Let $(y(u_h), p(u_h))$ be the solutions of the following auxiliary equations
\[
\begin{aligned}
\left\{ \begin{array}{l}
\frac{\partial y(u_h)}{\partial t}, w) + a(y(u_h), w) = (Eu_h, w) & \forall w \in H^1_0(\Omega), t \in (0, T],
y(u_h)(x, 0) = y_0(x) & x \in \Omega,
-
\frac{\partial p(u_h)}{\partial t}, q) + a(q, p(u_h)) = 0 & \forall q \in H^1_0(\Omega), t \in [0, T),
p(u_h)(x, T) = y(u_h)(x, T) & x \in \Omega.
\end{array} \right.
\end{aligned}
\]
For the case of $T < T_h$, we set $u(t) = u_h(t)$ for $t \in (T, T_h]$. Then, it follows from (31) and (32) that
\[
\int_0^T (E^* - E^* p(u_h), u_h - u) dt + \int_T^{T_h} (E^* p_h, u - u_h) dt
\]
\[
+ \int_0^T (E^* p(u_h) - E^* p_h, u_h - u) dt \geq 0.
\]
Considering the definition of $u$ in $(T, T_h]$, we have that $u = u_h$ in $(T, T_h]$. Thus (53) can be replaced as
\[
\int_0^T (E^* p - E^* p(u_h), u_h - u) dt + \int_0^T (E^* p(u_h) - E^* p_h, u_h - u) dt \geq 0.
\]
It follows from (5) and (52) that
\[
\int_0^T (E^* p - E^* p(u_h), u_h - u) dt = \int_0^T (p - p(u_h), Eu_h - Eu) dt
\]
\[
= \int_0^T \left( - \frac{\partial (y(u_h) - y)}{\partial t}, p - p(u_h) \right) + a(y(u_h) - y, p - p(u_h)) dt
\]
\[
+ (y(u_h)(T) - y(T), y(T) - y(u_h)(T)) - (y_0 - y_0, p(0) - p(u_h)(0))
\]
\[
= -\|y(T) - y(u_h)(T)\|^2.
\]
Therefore, (54) and (55) imply that

\[
\|y(T) - y(u_h)(T)\|^2 \leq \int_0^T (E^*(p(u_h) - p_h), u_h - u) dt \\
\leq \int_0^T (E^*(p(u_h) - p(y_h)), u_h - u) dt \\
+ \int_0^T (E^*(p(y_h) - p_h), u_h - u) dt,
\]

(56)

where \( p(y_h) \) is the solution of the following equation

\[
\begin{align*}
\begin{cases}
\left(-\frac{\partial p(y_h)}{\partial t}, q\right) + a(q, p(y_h)) = 0 &\forall q \in H_0^1(\Omega), \ t \in [0,T), \\
\frac{\partial p}{\partial n}(y_h)(x, T) = y_h(x, T) & x \in \Omega.
\end{cases}
\end{align*}
\]

Let \( \tilde{p}_h \in W^h \) be the solution of the equation

\[
\begin{align*}
\begin{cases}
\left(-\frac{\partial \tilde{p}_h}{\partial t}, q_h\right) + a(q_h, \tilde{p}_h) = 0 &\forall q_h \in W^h, \ t \in [0,T), \\
\tilde{p}_h(x, T) = y_h(x, T) & x \in \Omega,
\end{cases}
\end{align*}
\]

it is easy to see that \( \tilde{p}_h \) is the standard finite element approximation of \( p(y_h) \), and it follows from (12) that

\[
\|\tilde{p}_h - p(y_h)\|_{C([0,T];L^2(\Omega))} \leq Ch.
\]

Similar to the proof of (55) we have

\[
\begin{align*}
\int_0^T (E^*(p(u_h) - p(y_h)), u_h - u) dt \\
= \int_0^T (p(u_h) - p(y_h), E(u_h - u)) dt \\
= \|y(u_h)(T) - y(T), y(u_h)(T) - y_h(T)\|.
\end{align*}
\]

Then (56), (60) and \( \varepsilon \)-Young inequality imply that

\[
\|y(T) - y(u_h)(T)\|^2 \leq C\|y(u_h)(T) - y_h(T)\|^2 \\
+ C \int_0^T (E^*(p(y_h) - p_h), u_h - u) dt.
\]

(61)

From (12), we also have

\[
\|y(u_h)(T) - y_h(T)\| \leq Ch.
\]

Note that \( u, u_h \in U_{ad} \) and \( p(y_h), p_h \in L^2(0,T; L^2(\Omega)) \). Then (59) implies that

\[
\begin{align*}
\int_0^T (E^* p(y_h) - E^* p_h, u_h - u) dt \\
\leq \|p(y_h) - p_h\|_{L^2(0,T;L^2(\omega))} \|u - u_h\|_{L^2(0,T;L^2(\omega))} \\
\leq C\|p(y_h) - p_h\|_{L^2(0,T;L^2(\omega))} \\
\leq C\|\tilde{p}_h - p_h\|_{L^2(0,T;L^2(\Omega))} + C\|p(y_h)\|_{L^2(0,T;L^2(\Omega))} \\
\leq C\|\tilde{p}_h - p_h\|_{L^2(0,T;L^2(\Omega))} + Ch^2.
\end{align*}
\]

(63)

Moreover, it can be shown from (11) that

\[
\|\tilde{p}_h - p_h\|_{L^2(0,T;L^2(\Omega))} + \|\tilde{p}_h - p_h\|_{C([0,T];L^2(\Omega))} \leq C\|\tilde{p}_h(T) - p_h(T)\|.
\]

(64)
Using the initial conditions in (9), (58) and the error estimate (16) again, we have that
\[
\|\tilde{p}_h(T) - p_h(T)\| = \|y(T) - y_h(T)\| = \|y(T) - e^{\Delta_h(T_h-T)}y_h(T)\| \\
= \|y(T) - e^{\Delta_h(T_h-T)}(e^{\Delta_h(T_h-T)}y_h(T) + \int_T^{T_h} e^{\Delta_h(T_h-s)}E u_h(s)ds)\| \\
\leq \|(1 - e^{2\Delta_h(T_h-T)})y_h(T)\| + \|\int_T^{T_h} e^{\Delta_h(T_h-s)}E u_h(s)ds\| \\
\leq \|1 - e^{2\Delta_h(T_h-T)}\| \|y_h(T)\| + \int_T^{T_h} e^{-\lambda_1(T_h-s)}ds \\
\leq C\lambda_1^h(T_h - T)\|y_h(T)\| + \frac{C}{\lambda_1^h}(T_h - T)
\]
(65) \quad \leq C|T - T_h| \leq Ch.

This together with (64) gives
\[
\|\tilde{p}_h - p_h\|_{L^2(0,T;L^2(\Omega))} + \|\tilde{p}_h - p_h\|_{C([0,T];L^2(\Omega))} \leq Ch.
\]
Therefore, it follows from (61), (62), (63) and (66) that
\[
\|y(T) - y(u_h)(T)\| \leq C^\frac{1}{2}.
\]
Combing (62) and (67) we are led to
\[
\|y(T) - y_h(T)\| \leq C^\frac{1}{2}.
\]
We can conclude from (9), (11), (57) and (68) that
\[
\|p - p_h\|_{C([0,T];L^2(\Omega))} \leq C\|y(T) - y_h(T)\| \leq C^\frac{1}{2}.
\]
Thus, it follows from (59), (66) and (69) that
\[
\|p - p_h\|_{C([0,T];L^2(\Omega))} \leq C^\frac{1}{2}.
\]
Again we proved the error estimate (29) with the case \(T < T_h\).

We can observe from the proof of the error estimate (29) that the main reason for the reduction of convergence order for \(\|p - p_h\|_{C([0,T];L^2(\Omega))}\) lies in the estimates of (43) and (63), where \(\|u - u_h\|_{L^2(0,T;L^2(\Omega))}\) is bounded from above by a constant \(C\). So we need the estimate of \(\|u - u_h\|_{C([0,T];L^2(\Omega))}\) to improve the estimate of \(p\).

Using (6), (10), (51) and (70), we have that for all \(t \in [0,\tilde{T})\), if \(\|E^*p_h(t)\|_\omega \neq 0\) then
\[
\|u(t) - u_h(t)\|_\omega = \left\| \frac{E^*p(t)}{\|E^*p(t)\|_\omega} - \frac{E^*p_h(t)}{\|E^*p_h(t)\|_\omega} \right\|_\omega \\
= \frac{1}{\|E^*p_h(t)\|_\omega \|E^*p(t)\|_\omega} \left\| E^*p_h(t) \|E^*p(t)\|_\omega - \|E^*p(t)\| \|E^*p_h(t)\|_\omega \right\|_\omega \\
\leq \frac{1}{\|E^*p_h(t)\|_\omega \|E^*p(t)\|_\omega} \left( \|E^*p(t)\|_\omega \|E^*p_h(t)\|_\omega + \|E^*p(t)\| \|E^*p_h(t)\|_\omega \right) \\
\leq \frac{C}{\|E^*p(t)\|_\omega} \|E^*p_h(t) - E^*p(t)\|_\omega \leq \frac{Ch^\frac{1}{2}}{\|E^*p(t)\|_\omega}.
\]
Then (71) and (72) imply that for all $t$

$$\|u(t) - u_h(t)\|_\omega = \left\| \frac{E^*p(t)}{\|E^*p(t)\|_\omega} - \frac{E^*p(t)}{\sqrt{|\omega| \|E^*p(t)\|_\omega}} \right\|_\omega$$

$$= \frac{1}{\sqrt{|\omega| \|E^*p(t)\|_\omega}} \left\| \sqrt{|\omega| E^*p(t) - \|E^*p(t)\|_\omega} \right\|_\omega$$

$$= \frac{1}{\sqrt{|\omega| \|E^*p(t)\|_\omega}} \left\| \sqrt{|\omega| (E^*p(t) - p_h(t))} - \|E^*p(t)\|_\omega + \|E^*p_h(t)\|_\omega \right\|_\omega$$

(72) $$\leq \frac{C \sqrt{|\omega| \|E^*p(t)\|_\omega}}{\|E^*p(t)\|_\omega} \|E^*p(t) - E^*p_h(t)\|_\omega \leq \frac{Ch^\frac{1}{2}}{\|E^*p(t)\|_\omega}.$$  

Then (71) and (72) imply that for all $t \in (0, \tilde{T})$ there holds

$$\|u(t) - u_h(t)\|_\omega \leq \frac{Ch^\frac{1}{2}}{\|E^*p(t)\|_\omega}.$$  

Inserting (73) into both the estimates of (43) and (63) and proceeding as above we are led to

$$\|p(t) - p_h(t)\| \leq Ch^\frac{1}{2} \left( \frac{1}{\|E^*p(t)\|_\omega} \right)^\frac{1}{2},$$  

which in turn implies

$$\|u(t) - u_h(t)\|_\omega \leq Ch^\frac{1}{2} \left( \frac{1}{\|E^*p(t)\|_\omega} \right)^\frac{1}{2}.$$  

Repeat the above process, a bootstrapping like technique enables us to estimate that for all positive integer $n > 1$ and all $t \in (0, \tilde{T}),$

$$\|p(t) - p_h(t)\| \leq \frac{Ch^{\frac{1}{2} + \frac{1}{2n} + \cdots + \frac{1}{2^n}}}{\|E^*p(t)\|_\omega^{\frac{1}{2} + \frac{1}{2n} + \cdots + \frac{1}{2^n}}}$$

$$= \frac{Ch^{\frac{1}{2} + \frac{1}{2n} + \cdots + \frac{1}{2^n}}}{\|E^*p(t)\|_\omega^{\frac{1}{2} + \frac{1}{2n} + \cdots + \frac{1}{2^n}}} \leq \frac{C \|E^*p(t)\|_\omega^{\frac{1}{2} + \frac{1}{2n} + \cdots + \frac{1}{2^n}}}{\|E^*p(t)\|_\omega} \leq \frac{1}{\|E^*p(t)\|_\omega}.$$  

because that $\|E^*p(t)\|_\omega \leq C$. Then (27) is proved. Moreover, we can also obtain

$$\|u(t) - u_h(t)\|_\omega \leq \frac{Ch^{1-\varepsilon}}{\|E^*p(t)\|_\omega}$$

for all $t \in (0, \tilde{T})$ and $\varepsilon \in (0, 1)$, this gives (28).

Last, we consider the error for the state $y$. It is easy to see that $y_h(u)$ is the standard finite element approximation of $y$ on the time interval $[0, T_h]$ when $T_h \leq T$ and $y_h$ is the standard finite element approximation of $y(u_h)$ on the time interval $[0, T]$ when $T \leq T_h$. Then it follows from (12) that

(77) $$\|y - y_h(u)\|_{C([0, T_h], L^2(\Omega))} \leq Ch \quad \text{when} \quad T_h \leq T$$

and

(78) $$\|y_h - y(u_h)\|_{C([0, T], L^2(\Omega))} \leq Ch \quad \text{when} \quad T \leq T_h.$$
Moreover, it can be proven from (11) and (28) that
\[
\|y(t) - y(u_h)(t)\| \leq C\|u - u_h\|_{L^2(0,T;L^2(\omega))} \leq Ch^{1-\varepsilon}(t \int_0^t ds \|E^*p(s)\|_{L^2(\omega)})^{1/2},
\]
and
\[
\|y_h(u)(t) - y_h(t)\| \leq C\|u - u_h\|_{L^2(0,T;L^2(\omega))} \leq Ch^{1-\varepsilon}(t \int_0^t ds \|E^*p(s)\|_{L^2(\omega)})^{1/2}.
\]
Collecting the above results we obtain
\[
\|y(t) - y_h(t)\| \leq Ch^{1-\varepsilon}(t \int_0^t ds \|E^*p(s)\|_{L^2(\omega)})^{1/2},
\]
which proves (26). Then the proof of the theorem is completed.

**Remark 3.1.** It should be noted that when \(\|E^*p(t)\|_{L^2(\omega)} \geq C > 0\) for all \(t \in [0,\tilde{T}]\), Theorem 3.2 implies that for all \(t \in [0,\tilde{T}]\) and \(\varepsilon \in (0,1)\),
\[
\|y - y_h\|_{C([0,\tilde{T}];L^2(\Omega))} + \|p - p_h\|_{C([0,\tilde{T}];L^2(\Omega))} + \|u - u_h\|_{C([0,\tilde{T}];L^2(\omega))} \leq Ch^{1-\varepsilon}.
\]
This is the quasi-optimal error estimate.

4. Conclusion and outlook

In this paper, we discussed the finite element approximation to the time optimal control problem. The variational discretization concept was introduced for the approximation of the control, and the semi-discrete finite element method for the controlled equation was applied. The error estimates for the optimal time \(T\), the optimal control \(u\), the related state \(y\) and adjoint state \(p\) were provided. The quasi-optimal error estimates were obtained.

There are still many important issues to be addressed in this area. The fully discrete finite element method for the controlled equation and the numerical algorithm for the time optimal control problem should be investigated. It is also very interesting to study a posteriori error estimates and superconvergence for the time optimal control problems. Moreover, the research for efficient numerical algorithms of time optimal control problem and their application remains a challenge, which calls for more new techniques.

Acknowledgments

The authors would like to thank Professor Gengsheng Wang from Wuhan University for his helpful suggestion and discussion on the topic of this paper.

References


*LSEC, Institute of Computational Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. Email: wgong@lsec.cc.ac.cn.
†NCMIS, LSEC, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. Email: ynn@amss.ac.cn.