East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.071010.250411a

Vol. **1**, No. 3, pp. 235-247
August 2011

Numerical Methods for Constrained Elliptic Optimal Control Problems with Rapidly Oscillating Coefficients

Yanping Chen[*,1] and Yuelong Tang[2]

[1] *School of Mathematical Sciences, South China Normal University, Guangzhou 510631, Guangdong, China.*
[2] *Hunan Key Laboratory for Computation and Simulation in Science and Engineering, School of Mathematics and Computational Science, Xiangtan University, Xiangtan 411105, Hunan, China.*

Received 7 October 2010; Accepted (in revised version) 25 April 2011

Available online 27 July 2011

Abstract. In this paper we use two numerical methods to solve constrained optimal control problems governed by elliptic equations with rapidly oscillating coefficients: one is finite element method and the other is multiscale finite element method. We derive the convergence analysis for those two methods. Analytical results show that finite element method can not work when the parameter ε is small enough, while multiscale finite element method is useful for any parameter ε.

AMS subject classifications: 65K05, 65K10, 65M60

Key words: Optimal control problems, finite element method, multiscale finite element method, homogenization, convergence analysis.

---

1. Introduction

Optimal control plays a very important role in many engineering applications. Efficient numerical methods are necessary to successful applications of optimal control. Finite element method seems to be the most widely used numerical method in computing optimal control problems, and the relevant literature is huge. It is impossible to give even a very brief review here. A systematic introduction of finite element method for PDEs and optimal control problems can be found in [1, 9, 10, 24, 26]. For elliptic and parabolic optimal control problems, a priori error estimates of finite element method were established in [18], a posteriori error estimates of residual type have been derived in [20, 21], a posteriori error estimates of recovery type have been derived in [17, 19], and some superconvergence results can be found in [2–4]. However, many fundament and practical problems in

*Corresponding author. *Email address:* `yanpingchen@scnu.edu.cn` (Y. Chen)

engineering have multiscale solutions, such as composite materials, porous media, turbulent transport in high Reynods number flows and so on. The direct numerical simulation of multiple scale problems is difficult even with modern supercomputers for the requisite of tremendous amount of computer memory and CPU time which can easily exceed the limitation of today's computer resources.

In practical applications, it is often sufficient to predict the large scale solutions to certain accuracy. Multiscale finite element method [8, 11, 13, 14, 27] provides an efficient way to capture the large scale structures of the solutions on a coarse mesh. The main idea is to construct multiscale finite element base functions which capture the local small scale information within each element. The small scale information is then brought to the large scales through the coupling of the global stiffness matrix. It is through these multiscale base functions and the finite element formulation that the effect of small scales on the large scales is correctly captured. Mixed multiscale finite element methods for multiscale problems can be found in [5, 15, 22]. Recently, Chu et al. investigated a new multiscale finite element method for high-contrast elliptic interface problems in [6] and Parvazinia considered a multiscale finite element for the solution of transport equations in [25].

The purpose of this work is to obtain the convergence analysis for finite element method and multiscale finite element method solving a constrained optimal control problems governed by elliptic equations with rapidly oscillating coefficients. Such problems often arise in composite materials and flows in porous media.

Let Ω be a bounded domain in $\mathbb{R}^n (n=2,3)$ with a Lipschitz boundary $\partial\Omega$. In this paper, we adopt the standard notation $W^{m,q}(\Omega)$ for Sobolev spaces on Ω with norm $\|\cdot\|_{W^{m,q}(\Omega)}$ and seminorm $|\cdot|_{Wm,q(\Omega)}$. We set $H_0^1(\Omega) \equiv \left\{ v \in H^1(\Omega) : v|_{\partial\Omega} = 0 \right\}$ and denote $W^{m,2}(\Omega)$ by $H^m(\Omega)$. In addition, c or C denotes a generic positive constant.

We are interested in the following optimal control problem:

$$\begin{cases} \min\limits_{u \in K} \left\{ \dfrac{1}{2}\|y^\varepsilon - y_d\|^2 + \dfrac{1}{2}\|u\|^2 \right\}, \\ -\nabla \cdot (A(x, x/\varepsilon)\nabla y^\varepsilon) = Bu, \quad \text{in } \Omega, \\ y^\varepsilon = 0, \quad \text{on } \partial\Omega, \end{cases} \tag{1.1}$$

where K is a nonempty closed convex set in $L^2(\Omega)$, $A(x, x/\varepsilon)$ is a symmetric matrix which satisfies the uniform ellipticity condition:

$$\alpha|\xi|^2 \le a_{ij}(x, x/\varepsilon)\xi_i\xi_j \le \beta|\xi|^2, \ \forall \xi \in \mathbb{R}^n,$$

with $0 < \alpha < \beta$, $y_d \in L^2(\Omega)$, B is a continuous linear operator. Further more, we assume that $a_{ij}(x, \tilde{x})$ is periodic function with respect to the unit cube I in the "fast" variable $\tilde{x} = x/\varepsilon$, and

$$K = \left\{ v \in L^2(\Omega) \ : \ a \le v \le b, \ a.e. \text{ in } \Omega \right\},$$

where a and b are constants.

The paper is organized as follows: In Section 2, we shall construct a finite element approximation scheme and a multiscale finite element approximation scheme for the model problem (1.1), respectively. In Section 3, we introduce homogenization theory and related estimates for the state equation. In Section 4, we derive the convergence analysis of the finite element approximation scheme. In Section 5, we derive the convergence analysis of the multiscale finite element approximation scheme.

2. Approximation Schemes for the Model Problem

In this section, we consider a finite element approximation scheme and a multiscale finite element approximation scheme for the above model problem (1.1), respectively. For simplicity, we let $W=H_0^1(\Omega)$ with norm $\|\cdot\|_{1,\Omega}=\|\cdot\|_{H^1(\Omega)}$, $\|\cdot\|_{2,\Omega}=\|\cdot\|_{H^2(\Omega)}$ and $U=L^2(\Omega)$ with norm $\|\cdot\|=\|\cdot\|_{L^2(\Omega)}$.

First of all, we obtain the following weak formulation of (1.1): Find $(y^\varepsilon,u)\in W\times K$ such that

$$\begin{cases} \min\limits_{u\in K}\left\{\dfrac{1}{2}\|y^\varepsilon-y_d\|^2+\dfrac{1}{2}\|u\|^2\right\}, \\ (A(x,x/\varepsilon)\nabla y^\varepsilon,\nabla w)=(Bu,w), \quad \forall w\in W. \end{cases} \tag{2.1}$$

It is well known (see, e.g., [18]) that the optimal control problem (2.1) has a unique solution $(y^\varepsilon,u)\in W\times K$, and a pair $(y^\varepsilon,u)\in W\times K$ is the solution of (2.1) if and only if there is a co-state $p^\varepsilon\in W$ such that the triplet $(y^\varepsilon,p^\varepsilon,u)\in W\times W\times K$ satisfies the following optimality conditions:

$$\big(A(x,x/\varepsilon)\nabla y^\varepsilon,\nabla w\big)=(Bu,w), \qquad \forall w\in W, \tag{2.2}$$
$$(A(x,x/\varepsilon)\nabla q,\nabla p^\varepsilon)=(y^\varepsilon-y_d,q), \qquad \forall q\in W, \tag{2.3}$$
$$(u+B^*p^\varepsilon,v-u)\geq 0, \qquad \forall v\in K, \tag{2.4}$$

where B^* is the adjoint operator of B. It is well known that (2.4) is equivalent to

$$u=\max(a,\min(b,-B^*p^\varepsilon)). \tag{2.5}$$

In the second, we discuss the approximation scheme of the model problems. Let $\mathscr{T}^h$ be a regular triangulation or rectangulation of Ω and satisfy the following angle condition, namely there is a positive constant C such that for all $\tau\in\mathscr{T}^h$,

$$C^{-1}h_\tau^2\leq|\tau|\leq Ch_\tau^2,$$

where $|\tau|$ is the area of τ and h_τ is the diameter of τ. Let $h=\max_{\tau\in\mathscr{T}^h}\{h_\tau\}$.

2.1. Finite element approximation scheme for the model problem

A finite element approximation scheme is obtained by restricting the optimality conditions (2.2)-(2.4) to a finite dimensional subspace of $H_0^1(\Omega)$. Note that the regularity of the control variable is lower than the regularity of the state variable and the co-state variable, we use piecewise constant functions to approximate the control variable and use piecewise linear functions to approximate the state variable and the co-state variable, respectively. So we define the following two finite element function spaces. Associated with $\mathscr{T}^h$ are two finite dimensional subspaces:

$$W_h = \left\{ v_h \in C(\bar{\Omega}) \ : \ v_h|_\tau \in P_1, \ \forall \tau \in \mathscr{T}^h, \ v_h|_{\partial\Omega} = 0 \right\},$$
$$K_h = \left\{ v_h \in K \ : \ v_h|_\tau = \text{constant}, \ \forall \tau \in \mathscr{T}^h \right\},$$

where P_1 is the space of polynomials of total degree no more than 1.

Then the finite element approximation scheme of (2.1) is as follows: Find $(y_h, u_h) \in W_h \times K_h$ such that

$$\begin{cases} \min\limits_{u_h \in K_h} \left\{ \dfrac{1}{2}\|y_h - y_d\|^2 + \dfrac{1}{2}\|u_h\|^2 \right\}, \\ (A(x, x/\varepsilon)\nabla y_h, \nabla w_h) = (Bu_h, w_h), \quad \forall w_h \in W_h. \end{cases} \tag{2.6}$$

This optimal control problem (2.6) has a unique solution $(y_h, u_h) \in W_h \times K_h$, and a pair $(y_h, u_h) \in W_h \times K_h$ is the solution of (2.6) if and only if there is a co-state $p_h \in W_h$ such that the triplet $(y_h, p_h, u_h) \in W_h \times W_h \times K_h$ satisfies the following optimality conditions:

$$(A(x, x/\varepsilon)\nabla y_h, \nabla w_h) = (Bu_h, w_h), \qquad \forall w_h \in W_h, \tag{2.7}$$
$$(A(x, x/\varepsilon)\nabla q_h, \nabla p_h) = (y_h - y_d, q_h), \qquad \forall q_h \in W_h, \tag{2.8}$$
$$(u_h + B^* p_h, v_h - u_h) \geq 0, \qquad \forall v_h \in K_h. \tag{2.9}$$

It is easy to see that the inequality (2.9) is equivalent to

$$u_h|_\tau = \max(a, \min(b, -\overline{B^* p_h}|_\tau)), \qquad \forall \, \tau \in \mathscr{T}^h, \tag{2.10}$$

where $\overline{B^* p_h}|_\tau = \frac{1}{|\tau|}\int_\tau B^* p_h$ and $|\tau|$ is the measure of τ.

2.2. Multiscale finite element approximation scheme for the model problem

Since the basis functions of multiscale finite element method are absolutely different from the base functions of finite element method (see, e.g., [13]). We shall describe how to construct the base functions of multiscale finite element method in detail. In each element $\tau \in \mathscr{T}^h$, we define a set of nodal basis functions $\{ \phi_\tau^i, \, i = 1, \cdots, d \}$ with d being the number of nodes of τ. And ϕ_τ^i satisfies

$$-\nabla \cdot (A(x, x/\varepsilon)\nabla \phi_\tau^i) = 0, \qquad \text{in } \ \tau \in \mathscr{T}^h. \tag{2.11}$$

Let $x_j \in \bar{\tau}(j = 1, \cdots, d)$ be the nodal points of τ. As usual, we require $\phi_\tau^i(x_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta. We need to specify the Dirichlet boundary condition of ϕ_τ^i for well-posedness of (2.11). Let ϕ_τ^i are linear along $\partial\tau$. We assume that the basis functions are continuous across the boundaries of the elements, so we can define the following multiscale finite element space:

$$V_h = \text{span}\left\{ \phi_\tau^i : i = 1, \cdots, d; \ \ \forall \tau \in \mathcal{T}^h \right\} \subset H_0^1(\Omega).$$

Then the multiscale finite element approximation of (2.1) is as follows: Find $(\hat{y}_h, \hat{u}_h) \in V_h \times K_h$ such that

$$\begin{cases} \min\limits_{\hat{u}_h \in K_h} \left\{ \dfrac{1}{2}\|\hat{y}_h - y_d\|^2 + \dfrac{1}{2}\|\hat{u}_h\|^2 \right\}, \\ (A(x, x/\varepsilon)\nabla \hat{y}_h, \nabla \hat{w}_h) = (B\hat{u}_h, \hat{w}_h), \quad \forall \hat{w}_h \in V_h. \end{cases} \tag{2.12}$$

This control problem (2.12) has a unique solution $(\hat{y}_h, \hat{u}_h)$, and a pair $(\hat{y}_h, \hat{u}_h) \in V_h \times K_h$ is the solution of (2.12) if and only if there is a co-state $\hat{p}_h \in V_h$ such that the triplet $(\hat{y}_h, \hat{p}_h, \hat{u}_h) \in V_h \times V_h \times K_h$ satisfies the following optimality conditions:

$$(A(x, x/\varepsilon)\nabla \hat{y}_h, \nabla \hat{w}_h) = (B\hat{u}_h, \hat{w}_h), \qquad \forall \hat{w}_h \in V_h, \tag{2.13}$$

$$(A(x, x/\varepsilon)\nabla \hat{q}_h, \nabla \hat{p}_h) = (\hat{y}_h - y_d, \hat{q}_h), \qquad \forall \hat{q}_h \in V_h, \tag{2.14}$$

$$(\hat{u}_h + B^*\hat{p}_h, v_h - \hat{u}_h) \geq 0, \qquad \forall v_h \in K_h. \tag{2.15}$$

It is clear that the inequality (2.15) is equivalent to

$$\hat{u}_h|_\tau = \max(a, \min(b, -\overline{B^*\hat{p}_h}|_\tau)), \qquad \forall \tau \in \mathcal{T}^h, \tag{2.16}$$

where $\overline{B^*\hat{p}_h}|_\tau = \frac{1}{|\tau|}\int_\tau B^*\hat{p}_h$ and $|\tau|$ is the measure of τ.

3. Homogenization Theory and Related Estimates

In this section, we review the homogenization theory of the state equation (2.2). Let us consider the following state equation:

$$\begin{aligned} &L_\varepsilon y^\varepsilon = -\nabla \cdot (A(x, x/\varepsilon)\nabla y^\varepsilon) = Bu, \quad \text{in } \Omega, \\ &y^\varepsilon = 0, \quad \text{on } \partial\Omega. \end{aligned} \tag{3.1}$$

Following from [23,28], we may write the first equation of (3.1) as a first order system:

$$A(x, x/\varepsilon)\nabla y^\varepsilon = \boldsymbol{v}_\varepsilon, \tag{3.2}$$

$$-\nabla \cdot \boldsymbol{v}_\varepsilon = Bu, \tag{3.3}$$

and look for a formal expansion of the form

$$\begin{aligned} y^\varepsilon &= y_0(x, x/\varepsilon) + \varepsilon y_1(x, x/\varepsilon) + \cdots, \\ \boldsymbol{v}_\varepsilon &= \boldsymbol{v}_0(x, x/\varepsilon) + \varepsilon \boldsymbol{v}_1(x, x/\varepsilon) + \cdots, \end{aligned}$$

where $y_i(x, \tilde{x})$ and $\boldsymbol{v}_i(x, \tilde{x})$ are periodic in the “fast" variable $\tilde{x} = x/\varepsilon$.

By introducing $\nabla = \nabla_x + \varepsilon^{-1}\nabla_{\tilde{x}}$, and substituting the expansion of y^ε and $\boldsymbol{v}_\varepsilon$ into the system (3.2)-(3.3), we get

$$A(x,\tilde{x})\nabla_{\tilde{x}} y_0 = 0, \tag{3.4}$$
$$-\nabla_{\tilde{x}} \cdot \boldsymbol{v}_0 = 0, \tag{3.5}$$
$$A(x,\tilde{x})\nabla_{\tilde{x}} y_1 + A(x,\tilde{x})\nabla_x y_0 - \boldsymbol{v}_0 = 0, \tag{3.6}$$
$$-\nabla_{\tilde{x}} \cdot \boldsymbol{v}_1 - \nabla_x \cdot \boldsymbol{v}_0 = Bu. \tag{3.7}$$

From (3.1) and (3.4)-(3.7), we have $y_0 = y_0(x)$ satisfying the following equations:

$$\begin{aligned} &-\nabla \cdot (\overline{A}\nabla y_0) = Bu, \quad \text{in } \Omega, \\ &y_0 = 0, \quad \text{on } \partial\Omega, \end{aligned} \tag{3.8}$$

where $\overline{A} = (\overline{a}_{ij})_{n\times n}$ is given by

$$\overline{a}_{ij}(x) = \frac{1}{|I|}\int_I \left(a_{ij}(x,\tilde{x}) + \sum_{k=1}^n \left(a_{ik}\frac{\partial \chi^j}{\partial \tilde{x}_k} \right)(x,\tilde{x}) \right) d\tilde{x}, \tag{3.9}$$

and χ^j satisfies the following equations:

$$-\nabla_{\tilde{x}} \cdot (A(x,\tilde{x})\nabla_{\tilde{x}}\chi^j(x,\tilde{x})) = \sum_{i=1}^n \frac{\partial}{\partial \tilde{x}_i} a_{ij}(x,\tilde{x}), \ \ \tilde{x} \in I, \qquad \int_I \chi^j(x,\tilde{x}) d\tilde{x} = 0. \tag{3.10}$$

It is well known that $\overline{A}$ is a symmetric and positive definite constant matrix in Ω. Denote the homogenized operator as

$$L_0 = -\nabla \cdot (\overline{A}\nabla).$$

Then $L_0 y_0 = Bu_0$. In addition, we have

$$y_1(x,\tilde{x}) = -\chi^j \frac{\partial y_0}{\partial x_j}. \tag{3.11}$$

Note that $y_0(x,\tilde{x}) + \varepsilon y_1(x,\tilde{x}) \neq y^\varepsilon$ on $\partial\Omega$, due to the construction of y_1. We introduce a first order correction term θ_ε, satisfying

$$\begin{aligned} &L_\varepsilon \theta_\varepsilon = 0, \quad \text{in } \Omega, \\ &\theta_\varepsilon = y_1(x,\tilde{x}), \quad \text{on } \partial\Omega. \end{aligned} \tag{3.12}$$

Thus $y_0(x) + \varepsilon(y_1(x,\tilde{x}) - \theta_\varepsilon)$ satisfies the boundary condition of y^ε.

Just like the Proposition 1 of [23], we have the following lemma:

Lemma 3.1. *Let y_0 be the solution of* (3.8) *and $y_0 \in H^2(\Omega)$. Assume y_1 be given by* (3.11), *and $\theta_\varepsilon \in H^1(\Omega)$ is the solution of* (3.12). *Then there exists a constant C independent of y_0, ε and Ω such that*

$$\|y^\varepsilon - y_0 - \varepsilon(y_1 - \theta_\varepsilon)\|_{1,\Omega} \le \varepsilon C |y_0|_{2,\Omega}. \tag{3.13}$$

For the adjoint state equation (2.3), we have similar results.

4. Convergence Analysis for the Finite Element Method

In this section, we consider convergence analysis of finite element method. From [16], we have the following regularity estimates of equations (2.2) and (2.3),

$$\|y^\varepsilon\|_{2,\Omega} \le C_\varepsilon \|u\|, \tag{4.1}$$

$$\|p^\varepsilon\|_{2,\Omega} \le C_\varepsilon(\|y^\varepsilon\| + \|y_d\|), \tag{4.2}$$

where C_ε is of the order of $1/\varepsilon$.

We first give the following lemma:

Lemma 4.1. [7] *Let π_h be the standard Lagrange interpolation operator. For $m = 0$ or 1, $q > n/2$ and $v \in W^{2,q}(\Omega)$,*

$$|v - \pi_h v|_{m,q(\Omega)} \le Ch^{2-m}|v|_{W^{2,q}(\Omega)}. \tag{4.3}$$

For ease of exposition, we let

$$a(v,w) = \int_\Omega ((A(x,x/\varepsilon)\nabla v)\cdot \nabla w), \qquad \forall v,\, w \in W,$$

$$J(u) = \frac{1}{2}\|y^\varepsilon - y_d\|^2 + \frac{1}{2}\|u\|^2.$$

It can be shown that

$$(J'(u), v) = (u + B^* p^\varepsilon, v),$$

$$(J'(u_h), v) = (u_h + B^* p^\varepsilon(u_h), v),$$

where $p^\varepsilon(u_h)$ satisfies the following system:

$$(A(x,x/\varepsilon)\nabla y^\varepsilon(u_h), \nabla w) = (Bu_h, w), \qquad \forall w \in W, \tag{4.4}$$

$$(A(x,x/\varepsilon)\nabla q, \nabla p^\varepsilon(u_h)) = (y^\varepsilon(u_h) - y_d, q), \qquad \forall q \in W. \tag{4.5}$$

Lemma 4.2. *Let $(y^\varepsilon, p^\varepsilon, u)$ and (y_h, p_h, u_h) be the solutions of equations* (2.2)-(2.4) *and* (2.7)-(2.9)*, respectively. Assume that*

$$(J'(v) - J'(u), v - u) \ge c\|v - u\|^2, \qquad \forall u, v \in K, \tag{4.6}$$

and $u_I = \pi_h u \in K_h$ be the standard Lagrange interpolation of u such that

$$(u + B^* p^\varepsilon, u_I - u) \le Ch^2, \tag{4.7}$$

$$\|u - u_I\| \le Ch. \tag{4.8}$$

Then

$$\|u - u_h\| \le C\left(h + \|p_h - p^\varepsilon(u_h)\|\right). \tag{4.9}$$

Proof. It follows from (2.4), (2.9), (4.6)-(4.8), Hölder inequality and Young inequality that

$$
\begin{aligned}
&c\|u-u_h\|^2\\
\leq&(J'(u)-J'(u_h),u-u_h)\\
=&(u+B^*p^\varepsilon,u-u_h)-(u_h+B^*p^\varepsilon(u_h),u-u_h)\\
\leq&-(u_h+B^*p_h,u-u_h)+(B^*(p_h-p^\varepsilon(u_h)),u-u_h)\\
=&(u+B^*p^\varepsilon,u_I-u)+(u_h-u,u_I-u)+(B^*(p_h-p^\varepsilon(u_h)),u_I-u)\\
&+(B^*(p^\varepsilon(u_h)-p^\varepsilon),u_I-u)+(u_h+B^*p_h,u_h-u_I)+(B^*(p_h-p^\varepsilon(u_h)),u-u_h)\\
\leq&Ch^2+C(\delta)\|u_I-u\|^2+\delta\|u-u_h\|^2+\delta\|p^\varepsilon(u_h)-p^\varepsilon\|^2+C(\delta)\|p_h-p^\varepsilon(u_h)\|^2\\
\leq&Ch^2+\delta\|u-u_h\|^2+\delta\|p^\varepsilon(u_h)-p^\varepsilon\|^2+C(\delta)\|p_h-p^\varepsilon(u_h)\|^2. \qquad (4.10)
\end{aligned}
$$

Subtracting (2.2)-(2.3) from (4.4)-(4.5), we obtain,

$$
\begin{aligned}
&(A(x,x/\varepsilon)\nabla(y^\varepsilon(u_h)-y^\varepsilon),\nabla w)=(B(u_h-u),w), \qquad \forall w\in W, \qquad (4.11)\\
&(A(x,x/\varepsilon)\nabla q,\nabla(p^\varepsilon(u_h)-p^\varepsilon))=(y^\varepsilon(u_h)-y^\varepsilon,q), \qquad \forall q\in W. \qquad (4.12)
\end{aligned}
$$

From the regularity estimation of (4.11)-(4.12) and Friedriechs inequality, we have

$$
\begin{aligned}
&\|p^\varepsilon(u_h)-p^\varepsilon\|\leq C\|p^\varepsilon(u_h)-p^\varepsilon\|_{1,\Omega}\leq C\|y^\varepsilon(u_h)-y^\varepsilon\|, \qquad (4.13)\\
&\|y^\varepsilon(u_h)-y^\varepsilon\|\leq C\|y^\varepsilon(u_h)-y^\varepsilon\|_{1,\Omega}\leq C\|u_h-u\|. \qquad (4.14)
\end{aligned}
$$

Let δ be small enough, then (4.9) follows from (4.10) and (4.13)-(4.14). □

Theorem 4.1. *Let* $(y^\varepsilon,p^\varepsilon,u)$ *and* (y_h,p_h,u_h) *be the solutions of equations* (2.2)-(2.4) *and* (2.6)-(2.8)*, respectively. Assume that* $y^\varepsilon,p^\varepsilon\in H^2(\Omega)$ *and all the conditions in* Lemma 4.2 *are valid. Then,*

$$
\|y^\varepsilon-y_h\|_{1,\Omega}+\|p^\varepsilon-p_h\|_{1,\Omega}+\|u-u_h\|\leq C_\varepsilon h. \qquad (4.15)
$$

Proof. It follows from the assumptions on $A(x,x/\varepsilon)$ that

$$
\begin{aligned}
&c\|p^\varepsilon(u_h)-p_h\|_{1,\Omega}^2\\
\leq&a(p^\varepsilon(u_h)-p_h,p^\varepsilon(u_h)-p_h)\\
=&a(p^\varepsilon(u_h)-\pi_h(p^\varepsilon(u_h)),p^\varepsilon(u_h)-p_h)+a(\pi_h(p^\varepsilon(u_h))-p_h,p^\varepsilon(u_h)-p_h)\\
=&a(p^\varepsilon(u_h)-\pi_h(p^\varepsilon(u_h)),p^\varepsilon(u_h)-p_h)+(y^\varepsilon(u_h)-y_h,\pi_h(p^\varepsilon(u_h))-p_h)\\
\leq&C\|p^\varepsilon(u_h)-p_h\|_{1,\Omega}\|p^\varepsilon(u_h)-\pi_h(p^\varepsilon(u_h))\|_{1,\Omega}+C\|y^\varepsilon(u_h)-y_h\|\|\pi_h(p^\varepsilon(u_h))-p_h\|\\
\leq&C(\delta)\|p^\varepsilon(u_h)-\pi_h(p^\varepsilon(u_h))\|_{1,\Omega}^2+C(\delta)\|y^\varepsilon(u_h)-y_h\|^2+C\delta\|\pi_h(p^\varepsilon(u_h))-p_h\|_{1,\Omega}^2\\
&+C\delta\|(p^\varepsilon(u_h))-p_h\|_{1,\Omega}^2\\
\leq&C(\delta)\|p^\varepsilon(u_h)-\pi_h(p^\varepsilon(u_h))\|_{1,\Omega}^2+C(\delta)\|y^\varepsilon(u_h)-y_h\|^2+C\delta\|p^\varepsilon(u_h)-p_h\|_{1,\Omega}^2. \qquad (4.16)
\end{aligned}
$$

Let δ to be small enough, using Lemma 4.1, we have

$$\begin{aligned}\|p^\varepsilon(u_h)-p_h\|_{1,\Omega}\le & C\|p^\varepsilon(u_h)-\pi_h(p^\varepsilon(u_h))\|_{1,\Omega}+C\|y^\varepsilon(u_h)-y_h\|\\ \le & Ch\|p^\varepsilon(u_h)\|_{2,\Omega}+C\|y^\varepsilon(u_h)-y_h\|.\end{aligned}$$

Note that Ω is convex. From regularity estimates, we get

$$\begin{aligned}\|p^\varepsilon(u_h)\|_{2,\Omega}\le & \|p^\varepsilon\|_{2,\Omega}+\|p^\varepsilon(u_h)-p^\varepsilon\|_{2,\Omega}\\ \le & \|p^\varepsilon\|_{2,\Omega}+C\|y^\varepsilon(u_h)-y^\varepsilon\|\\ \le & \|p^\varepsilon\|_{2,\Omega}+C\|u_h-u\|\le C_\varepsilon.\end{aligned}$$

Therefore,

$$\|p^\varepsilon(u_h)-p_h\|_{1,\Omega}\le C_\varepsilon h+C\|y^\varepsilon(u_h)-y_h\|. \tag{4.17}$$

Similarly, we can prove that

$$\|y^\varepsilon(u_h)-y_h\|_{1,\Omega}\le Ch\|y^\varepsilon(u_h)\|_{2,\Omega}\le C_\varepsilon h. \tag{4.18}$$

It follows from (4.17) and (4.18) that

$$\|p^\varepsilon(u_h)-p_h\|\le\|p^\varepsilon(u_h)-p_h\|_{1,\Omega}\le C_\varepsilon h. \tag{4.19}$$

From the Lemma 4.2 and (4.19), we have

$$\|u-u_h\|\le C_\varepsilon h. \tag{4.20}$$

Note that

$$\|y^\varepsilon-y_h\|_{1,\Omega}\le\|y^\varepsilon-y^\varepsilon(u_h)\|_{1,\Omega}+\|y^\varepsilon(u_h)-y_h\|_{1,\Omega}, \tag{4.21}$$

$$\|p^\varepsilon-p_h\|_{1,\Omega}\le\|p^\varepsilon-p^\varepsilon(u_h)\|_{1,\Omega}+\|p^\varepsilon(u_h)-p_h\|_{1,\Omega}, \tag{4.22}$$

and

$$\|p^\varepsilon-p^\varepsilon(u_h)\|_{1,\Omega}\le C\|y^\varepsilon-y^\varepsilon(u_h)\|_{1,\Omega}\le C\|u-u_h\|. \tag{4.23}$$

Then, (4.15) follows from (4.17)-(4.23). □

Remark 4.1. Since C_ε is usually of the order of $1/\varepsilon$, when ε is small enough, we have to use very fine mesh size which should be smaller than ε to approximate y^ε and p^ε. It is often beyond our computational powers even for two-dimensional cases. Thus in this case the finite element method can not work.

5. Convergence Analysis for the Multiscale Finite Element Method

In this section, we study convergence analysis of multiscale finite element method for the model problem (1.1). Just like the Lemma 4.2, it is easy to prove the following lemma:

Lemma 5.1. *Let* $(y^\varepsilon, p^\varepsilon, u)$ *and* $(\hat{y}_h, \hat{p}_h, \hat{u}_h)$ *be the solutions of equations* (2.2)-(2.4) *and* (2.13)-(2.15)*, respectively. Assume that*

$$(J'(v) - J'(u), v - u) \geq c\|v - u\|^2, \qquad \forall u, v \in K, \tag{5.1}$$

and $u_I = \pi_h u \in K_h$ *be the standard Lagrange interpolation of* u *such that*

$$(u + B^* p^\varepsilon, u_I - u) \leq Ch^2, \tag{5.2}$$

$$\|u - u_I\| \leq Ch. \tag{5.3}$$

Then

$$\|u - \hat{u}_h\| \leq C(h + \|\hat{p}_h - p^\varepsilon(\hat{u}_h)\|), \tag{5.4}$$

where $p^\varepsilon(\hat{u}_h)$ *satisfies the following system:*

$$(A(x, x/\varepsilon)\nabla y^\varepsilon(\hat{u}_h), \nabla w) = (B\hat{u}_h, w), \qquad \forall w \in W, \tag{5.5}$$

$$(A(x, x/\varepsilon)\nabla q, \nabla p^\varepsilon(\hat{u}_h)) = (y^\varepsilon(\hat{u}_h) - y_d, q), \qquad \forall q \in W, \tag{5.6}$$

with $\hat{u}_h \in K_h$.

Since $a_{ij}(x, \tilde{x})$ is periodic in the "fast" variable $\tilde{x} = x/\varepsilon$, from Lemma 5.3 in [14], it is easy to prove the following lemma:

Lemma 5.2. *Let* $p^\varepsilon(\hat{u}_h) \in H^1(\Omega)$ *satisfy equation* (5.6) *and* $p^\varepsilon(\hat{u}_h)_I \in V_h$ *be the interpolation of its homogenized solution* $p_0(\hat{u}_h)$*, using the multiscale base functions* ϕ^i*. Then there exists constant* C *independent of* ε *and* h*, such that*

$$\|p^\varepsilon(\hat{u}_h)_I - p^\varepsilon(\hat{u}_h)\|_{1,\Omega} \leq C\left(h(\|y^\varepsilon(\hat{u}_h)\| + \|y_d\|) + \sqrt{\varepsilon/h}\right). \tag{5.7}$$

Theorem 5.1. *Let* $(y^\varepsilon, p^\varepsilon, u)$ *and* $(\hat{y}_h, \hat{p}_h, \hat{u}_h)$ *be the solutions of equations* (2.2)-(2.4) *and* (2.13)-(2.15)*, respectively. Assume that all the conditions in* Lemmas 5.1-5.2 *are valid. Moveover, assume that* $y^\varepsilon, p^\varepsilon \in H^2(\Omega)$*. Then there exists a constant* C *independent of* ε *and* h *such that*

$$\|y^\varepsilon - \hat{y}_h\|_{1,\Omega} + \|p^\varepsilon - \hat{p}_h\|_{1,\Omega} + \|u - \hat{u}_h\| \leq C\left(h + \sqrt{\varepsilon/h}\right). \tag{5.8}$$

Proof. It follows from the assumptions on $A(x,x/\varepsilon)$, Lemma 5.2, ε-Cauchy inequality and Friedriechs inequality, we have

$$\begin{aligned}
&c\|p^\varepsilon(\hat{u}_h)-\hat{p}_h\|_{1,\Omega}^2\\
\le& a(p^\varepsilon(\hat{u}_h)-\hat{p}_h,p^\varepsilon(\hat{u}_h)-\hat{p}_h)\\
=& a(p^\varepsilon(\hat{u}_h)-p^\varepsilon(\hat{u}_h)_I,p^\varepsilon(\hat{u}_h)-\hat{p}_h)+a(p^\varepsilon(\hat{u}_h)_I-\hat{p}_h,p^\varepsilon(\hat{u}_h)-\hat{p}_h)\\
=& a(p^\varepsilon(\hat{u}_h)-p^\varepsilon(\hat{u}_h)_I,p^\varepsilon(\hat{u}_h)-\hat{p}_h)+(y^\varepsilon(\hat{u}_h)-\hat{y}_h,p^\varepsilon(\hat{u}_h)_I-\hat{p}_h)\\
\le& C\|p^\varepsilon(\hat{u}_h)-\hat{p}_h\|_{1,\Omega}\|p^\varepsilon(\hat{u}_h)-p^\varepsilon(\hat{u}_h)_I\|_{1,\Omega}+C\|y^\varepsilon(\hat{u}_h)-\hat{y}_h\|\|p^\varepsilon(\hat{u}_h)_I-\hat{p}_h\|\\
\le& C(\delta)\|p^\varepsilon(\hat{u}_h)-p^\varepsilon(\hat{u}_h)_I\|_{1,\Omega}^2+C(\delta)\|y^\varepsilon(\hat{u}_h)-\hat{y}_h\|^2+C\delta\|p^\varepsilon(\hat{u}_h)_I-\hat{p}_h\|_{1,\Omega}^2\\
&+C\delta\|p^\varepsilon(\hat{u}_h)-\hat{p}_h\|_{1,\Omega}^2\\
\le& C(\delta)\|p^\varepsilon(\hat{u}_h)-p^\varepsilon(\hat{u}_h)_I\|_{1,\Omega}^2+C(\delta)\|y^\varepsilon(\hat{u}_h)-\hat{y}_h\|^2+C\delta\|p^\varepsilon(\hat{u}_h)-\hat{p}_h\|_{1,\Omega}^2\\
\le& C\left(h(\|y^\varepsilon(\hat{u}_h)\|+\|y_d\|)+\sqrt{\varepsilon/h}\right)^2+C(\delta)\|y^\varepsilon(\hat{u}_h)-\hat{y}_h\|^2+C\delta\|p^\varepsilon(\hat{u}_h)-\hat{p}_h\|_{1,\Omega}^2. \quad (5.9)
\end{aligned}$$

Let δ be small enough, we get

$$\|p^\varepsilon(\hat{u}_h)-\hat{p}_h\|_{1,\Omega}\le C\left(h(\|y^\varepsilon(\hat{u}_h)\|+\|y_d\|)+\sqrt{\varepsilon/h}\right)+C(\delta)\|y^\varepsilon(\hat{u}_h)-\hat{y}_h\|. \tag{5.10}$$

Similarly, it is easy to prove

$$\|y^\varepsilon(\hat{u}_h)-\hat{y}_h\|_{1,\Omega}\le C\left(h\|\hat{u}_h\|+\sqrt{\varepsilon/h}\right). \tag{5.11}$$

Thus,

$$\|y^\varepsilon(\hat{u}_h)-\hat{y}_h\|_{1,\Omega}+\|p^\varepsilon(\hat{u}_h)-\hat{p}_h\|_{1,\Omega}\le C\left(h+\sqrt{\varepsilon/h}\right). \tag{5.12}$$

Note that

$$\|y^\varepsilon-\hat{y}_h\|_{1,\Omega}\le\|y^\varepsilon-y^\varepsilon(\hat{u}_h)\|_{1,\Omega}+\|y^\varepsilon(\hat{u}_h)-\hat{y}_h\|_{1,\Omega}, \tag{5.13}$$

$$\|p^\varepsilon-\hat{p}_h\|_{1,\Omega}\le\|p^\varepsilon-p^\varepsilon(\hat{u}_h)\|_{1,\Omega}+\|p^\varepsilon(\hat{u}_h)-\hat{p}_h\|_{1,\Omega}, \tag{5.14}$$

and

$$\|p^\varepsilon-p^\varepsilon(\hat{u}_h)\|_{1,\Omega}\le C\|y^\varepsilon-y^\varepsilon(\hat{u}_h)\|_{1,\Omega}\le C\|u-u_h\|. \tag{5.15}$$

Then, (5.8) follows from (5.11) and (5.12)-(5.15). □

Remark 5.1. Since the multiscale finite element method is of the order $h+\sqrt{\varepsilon/h}$ convergence. It is useful for any parameter ε.

Acknowledgments

The first author is supported by the Foundation for Talent Introduction of Guangdong Provincial University, Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme (2008), National Science Foundation of China (10971074). The second author is supported by the Hunan Provinical Innovation Foundation for Postgraduate (lx2009B120).

References

[1] N. Arada, E. Casas and F. Tröltzsch, *Error estimates for a semilinear elliptic optimal control problems*, Copmut. Optim. Appl., 23 (2002), pp. 201-229.

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

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

[4] Y. Chen, Y. Huang, W. B. Liu and N. N. Yan, *Error estimates and superconvergence of mixed finite element methods for convex optimal control problems* , J. Sci. Comput., 42(4) (2010), pp. 382-403.

[5] Z. Chen and T. Y. Hou, *A mixed multiscale finite element method for elliptic problems with oscillating coefficients*, Math. Comp., 72 (2003), pp. 541-576.

[6] C. C. Chu, G. Graham and T. Y. Hou, *A new multiscale finite element method for high-contrast elliptic interface problems*, Math. Comp., 79 (2010), pp. 1915-1955.

[7] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[8] Y. R. Efendiev, T. Y. Hou and X. H. Wu, *The convergence of non-conforming multiscale finite element methods*, SIAM J. Numer. Anal., 37 (2000), pp. 888-910.

[9] R. Falk, *Approximation of a class of optimal control problems with order of convergence estimates*, J. Math. Anal. Appl., 44 (1973), pp. 28-47.

[10] T. Geveci, *On the appoximation of the solution of an elliptic control problem governed by an elliptic equation*, RAIRO Anal. Numer., 13 (1979), pp. 13-328.

[11] W. M. He and J. Z. Cui, *A finite element method for elliptic problems with rapidly oscillating coefficients*, BIT Numer. Math., 47 (2007), pp. 77-102.

[12] W. M. He and J. Z. Cui, *A new approximation method for second order elliptic problems with rapidly oscillating coefficients based on the method of multiscale asymptotic expansion*, J. Math. Anal. Appl., 335 (2007), pp. 657-668.

[13] T. Y. Hou and X. H. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys., 134 (1997), pp. 169-189.

[14] T. Y. Hou, X. H. Wu and Z. Cai, *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients*, Math. Comp., 68 (1999), pp. 913-943.

[15] L. Jiang and Y. Efendiev, *Mixed multisclae finite element analysis for wave equations using global information*, Doma. Deco. Meth. Sci. Eng. XIX: Lecture Notes Comput. Sci. Eng., 78 (2011), pp. 181-188.

[16] O. A. Ladyzhenskaya and N. N. Ural'tseva, *Linear and Quasilinear Elliptic Equations*, Academic Press, New York, 1968.

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

[18] J. L. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin, 1971.

[19] H. LIU AND N. N. YAN, *Recovery type superconvergence and a posteriori error estimates for control problems governed by Stokes equations*, J. Comput. Appl. Math., 209 (2007), pp. 187-207.

[20] W. B. LIU AND N. N. YAN, *A posteriori error estimates for distributed convex optimal control problems*, Adv. Comput. Math., 15 (2001), pp. 285-309.

[21] W. B. LIU AND N. N. YAN, *A posteriori error estimates for optimal control problems governed by parabolic equations*, Numer. Math., 93 (2003), pp. 497-521.

[22] X. Ma and N. Zabaras, *A stochastic mixed finite element heterogenous multiscale method for flow in porous media*, J. Comput. Phys., 230 (2011), pp. 4696-4772.

[23] S. MOSKOW AND M. VOGELIUS, *First order corrections to the homogenized eigenvalues of a periodic composite medium*, Proc. Roy. Soc. Edinburgh Sect. A., 127 (1997), pp.1263-1299.

[24] P. NEITTAANMAKI AND D. TIBA, *Optimal Control Of Nonlinear Parabolic Systems: Thoery, Algorithms And Applications*, M. Dekker, New York, 1994.

[25] M. PARVAZINIA, *A multiscale finite element for the solution of transport equations*, Finite Elem. Anal. Des., 47 (2011), pp. 211-219.

[26] L. R. SCOTT AND S. ZHANG, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483-493.

[27] E. WEINAN, P. MING AND P. ZHANG, *Analysis of the heterogeneous multi-scale method for elliptic homogenization problems*, J. Amer. Math. Soc., 18(1) (2005), pp. 121-156.

[28] V. V. ZHIOV, S. M. KOZLOV AND O. A. OLEINIK, *Homogenization of Differential Operators and Integral Functionals*, Springer-Verlag, Heidelberg, 1994.