J. Korean Math. Soc. **47** (2010), No. 1, pp. 83–100
DOI 10.4134/JKMS.2010.47.1.083

A MULTIGRID METHOD FOR AN OPTIMAL CONTROL PROBLEM OF A DIFFUSION-CONVECTION EQUATION

Hunki Baek, Sang Dong Kim, and Hyung-Chun Lee

Abstract. In this article, an optimal control problem associated with convection-diffusion equation is considered. Using Lagrange multiplier, the optimality system is obtained. The derived optimal system becomes coupled, non-symmetric partial differential equations. For discretizations and implementations, the finite element multigrid V-cycle is employed. The convergence analysis of finite element multigrid methods for the derived optimal system is shown. Some numerical simulations are performed.

1. Introduction

Optimization and control problems for systems associated with partial differential equations arise in many applications [3, 16, 20, 21] and are receiving much attention because of their importance in the industrial design process. Especially, the need for accurate and efficient solution methods for these problems has become an important issue.

Optimization or control problems have the usual three ingredients. First, one has an objective, a reason why one wants to control the state variables. Mathematically, such an objective is expressed as a cost, or performance functional. Next, one has controls or design parameters at one's disposal in order to meet the objective. Indeed, controls or design parameters are expressed in term of unknown data in the mathematical specification of the problem. Finally, one has constraints that determine what type of partial differential equations are interested in and that place direct or indirect limits on candidate optimizers. In this paper, we concern the diffusion-convection equation as the type of partial differential equations. The optimization problem is then to find optimal state and controls that minimize the objective functional subject to the requirement that the constraints are satisfied.

Received February 20, 2008; Revised October 2, 2008.
2000 *Mathematics Subject Classification.* Primary 49K20, 65N12, 65N30, 65N55.
Key words and phrases. optimal control problem, multigrid method, diffusion-convection equation.
This work was supported by KRF-2005-070-C00017.

It is well known that such constrained optimization problems can be converted to the unconstrained optimization problems by the Lagrange multiplier method [3, 16], which leads to the state equation, and adjoint equations, and an optimality condition. Because we think the diffusion-convection equation which is uniformly positive definite, one may see that the optimality system is a coupled non-symmetric and definite system. For a numerical approach we consider the finite element multigrid method to solve the discretized optimality systems. This is because multigrid methods have been extensively used to solve discretized partial differential equations successfully for a long time in many literature (for example, [4, 7, 8, 10, 11, 12, 19]). It is known that multigrid methods [7, 19] solve elliptic problems with optimal computational order. This fact has been demonstrated in the case of multigrid applied to a singular optimal control problem associated with a nonlinear elliptic equation [6]. Such techniques were applied to solve optimal control problems [5, 6, 17, 18]. Especially, in [5], Borzi and et al. proved the multigrid convergence of a finite difference method for the optimal control optimality system, which is two copies of Poisson equations (a decoupled symmetric system). However, it is difficult to adopt the methods used in [5] to the optimality system (4) because of the convection term. To avoid such difficulties, we use the perturbation operator. Thus, the main purpose of this paper is to show the convergence of a finite element multigrid method for the optimality system (4), which is a coupled nonsymmetric system. In particular, applying V multigrid algorithms to the whole coupled optimality system, we provide the multigrid V-cycle convergence analysis with same optimal convergence phenomena as the usual elliptic boundary value problems posses.

We give some examples in Section 4. We exhibit that, in the sense of approximating the state variables, the numerical approximation by V-cycle for a chosen acceleration parameter to weak solutions of the optimality system approaches the state variable. These phenomena can be verified by showing numerical errors in terms of L^2 errors. It is also shown that the numerical results verify optimal acceleration parameters for minimizing the quadratic functional.

In the following section we describe optimal control problem with some necessary introduction for multigrid methods. For a coupled optimality system, the convergence of finite element multigrid methods will be shown in Section 3. With both a model problem and a convection-diffusion problem, several numerical examples are provided for optimality systems in terms of L^2 errors in Section 4. Finally, we provide some conclusions in Section 5.

2. The optimal control problem

We consider an optimal control problem minimizing a quadratic functional

$$\mathcal{J}(u,\theta) = \frac{\alpha}{2}\int_\Omega |u-\hat{u}|^2 d\Omega + \frac{\delta}{2}\int_\Omega |\theta|^2 d\Omega \tag{1}$$

subject to $\theta \in L^2(\Omega)$ and the following uniformly positive definite elliptic problem

$$(2)\qquad \begin{cases} -\nabla\cdot(\mathbf{B}\nabla u) + \mathbf{b}\cdot\nabla u + u = \theta & \text{in} \quad \Omega, \\ u = 0 & \text{on} \quad \partial\Omega, \end{cases}$$

where Ω is a convex polygonal domain or has $C^{1,1}$ boundary condition in $\mathbb{R}^2$ and $\hat{u} \in L^2(\Omega)$ is the objective function, and $\alpha, \delta > 0$ are the weights of the cost of the control. Further we may assume that $\mathbf{B} = (B_{ij}(x))$ is symmetric, uniformly positive definite matrix with C^1-functions $B_{ij}(x)$ and each component $b_i(x)$ of $\mathbf{b}(x)$ is continuous differentiable on $\bar{\Omega}$.

The optimal control problem we consider is to seek a state u and a controller θ so that the functional (1) is minimized subject to (2). Such constrained optimization problems may be recast as unconstrained optimization problems through the Lagrange multiplier method. The existence of the optimal solution and Lagrange multiplier is well known [16, 20, 21]. Then introducing Lagrange multiplier $v \in H^2(\Omega) \cap H_0^1(\Omega)$ and using Green's formula lead the following coupled elliptic boundary value problem:

$$(3)\qquad \begin{aligned} -\nabla\cdot(\mathbf{B}\nabla u) + \mathbf{b}\cdot\nabla u + u &= \theta && \text{in } \Omega, \\ -\nabla\cdot(\mathbf{B}\nabla v) - \nabla\cdot(v\mathbf{b}) + v &= \alpha(\hat{u} - u) && \text{in } \Omega, \\ \delta\theta &= v && \text{in } \Omega, \\ \theta = 0, u = 0, v &= 0 && \text{on } \partial\Omega. \end{aligned}$$

Since the corresponding formal normal equations to (3) leads to an indefinite system of equations, we may replace θ by $\frac{1}{\delta}v$ in first equation of (3) so that we may have a positive definite system. This procedure, letting $\alpha\hat{u} = f$ for convenience, yields to the optimality system such as

$$(4)\qquad \begin{aligned} -\nabla\cdot(\mathbf{B}\nabla u) + \mathbf{b}\cdot\nabla u + u - \frac{1}{\delta}v &= 0 && \text{in } \Omega, \\ -\nabla\cdot(\mathbf{B}\nabla v) - \nabla\cdot(v\mathbf{b}) + v + \alpha u &= f && \text{in } \Omega, \\ u = 0, v &= 0 && \text{on } \partial\Omega. \end{aligned}$$

Define a bilinear form $\mathcal{A}(\cdot,\cdot;\cdot,\cdot)$ on the space $H_0^1(\Omega) \times H_0^1(\Omega)$ by

$$(5)\qquad \mathcal{A}(u, v; w, z) = \alpha\mathcal{L}(u, w) + \frac{1}{\delta}\mathcal{L}(z, v) - \frac{\alpha}{\delta}(v, w) + \frac{\alpha}{\delta}(u, z),$$

where

$$(6)\qquad \mathcal{L}(u, w) = (\mathbf{B}\nabla u, \nabla w) + (\mathbf{b}\cdot\nabla u, w) + (u, w).$$

Now, we can see that the unconstrained optimization problems becomes the variational problem for finding $(u, v) \in H_0^1(\Omega) \times H_0^1(\Omega)$ such that

$$(7)\qquad \mathcal{A}(u, v; w, z) = \frac{1}{\delta}(f, z) \text{ for all } (w, z) \in H_0^1(\Omega) \times H_0^1(\Omega),$$

where $H_0^1(\Omega)$ is the standard Sobolev space with norm $\|\cdot\|_1$. In this paper we use also the standard Sobolev space $H^s(\Omega)$ with norm $\|\cdot\|_s$ for nonnegative integer. Note that the usual $L^2(\Omega)$ space is same as $H^0(\Omega)$ with the usual $\|\cdot\|_0$ norm. It is easy to see that the bilinear form (5) is continuous. From (2), we know that $\mathbf{B}$ is uniformly positive definite. Thus, from Gårding Inequality [13], the bilinear form (5) satisfies the coercivity.

Let us split the bilinear form $\mathcal{A}(\cdot,\cdot;\cdot,\cdot)$ into the symmetric bilinear form $\widehat{\mathcal{A}}(\cdot,\cdot;\cdot,\cdot)$ and nonsymmetric bilinear form $\mathcal{D}(\cdot,\cdot;\cdot,\cdot)$ such that

$$\mathcal{A}(u,v;w,z) = \widehat{\mathcal{A}}(u,v;w,z) + \mathcal{D}(u,v;w,z), \tag{8}$$

where

$$\widehat{\mathcal{A}}(u,v;w,z) = \alpha(B\nabla u,\nabla w) + \alpha(u,w) + \frac{1}{\delta}(B\nabla v,\nabla z) + \frac{1}{\delta}(v,z) \tag{9}$$

and

$$\mathcal{D}(u,v;w,z) = \alpha(\mathbf{b}\cdot\nabla u,w) + \frac{1}{\delta}(\mathbf{b}\cdot\nabla z,v) - \frac{\alpha}{\delta}(v,w) + \frac{\alpha}{\delta}(u,z). \tag{10}$$

Then we easily show that the form $\mathcal{D}(\cdot,\cdot;\cdot,\cdot)$ satisfies the inequalities, for some $C_d > 0$,

$$|\mathcal{D}(u,v;w,z)| \le C_d(\|u\|_1 + \|v\|_1)(\|w\|_0 + \|z\|_0) \tag{11}$$

and

$$|\mathcal{D}(u,v;w,z)| \le C_d(\|u\|_0 + \|v\|_0)(\|w\|_1 + \|z\|_1). \tag{12}$$

For a finite element multigrid approximation, let $\mathcal{T}_h$ be a quasi uniform triangulation of Ω which has a sequence of nested triangulations of Ω in the usual way. Let us denote $\mathcal{T}_{k+1} := \mathcal{T}_{h2^{-k}}$. We assume that a coarse triangulations $\mathcal{T}_1$ of Ω is given. Let $\mathcal{T}_k (k \ge 2)$ be obtained from $\mathcal{T}_{k-1}$ via a regular subdivision. We let the number of levels in the multigrid algorithm be determined by J. For $J \ge 1$ define $\mathbf{V}_k$ for $k = 1, 2, \ldots, J$ to be the functions which are piecewise linear with respect to $\mathcal{T}_k$ that vanish on $\partial\Omega$, so that $\mathbf{V}_1 \subset \mathbf{V}_2 \subset \cdots \subset \mathbf{V}_J \subset H_0^1(\Omega)^2$. Let h_k be the mesh size of $\mathcal{T}_k$, i.e., $h_k = \max_{T\in\mathcal{T}_k} \mathrm{diam} T$ with $h_k = \frac{1}{2}h_{k-1}$. The mesh-dependent inner product $(\cdot,\cdot)_k$ on $\mathbf{V}_k \times \mathbf{V}_k$ defined by

$$(\mathbf{w},\mathbf{z})_k = h_k^2 \sum_{i=1}^{n_k} (w_1(x_i)z_1(x_i) + w_2(x_i)z_2(x_i)), \tag{13}$$

where $\mathbf{w} = (w_1, w_2), \mathbf{z} = (z_1, z_2) \in \mathbf{V}_k$ and $\{x_i\}_{i=1}^{n_k}$ is the set of internal vertices of $\mathcal{T}_k$.

Let us denote $(u_{h_k}, v_{h_k}) \in \mathbf{V}_k$ as the finite element solution corresponding to (7). Then one may prove immediately that

$$\|u - u_{h_k}\|_1 + \|v - v_{h_k}\|_1 \le Ch_k(\|u\|_2 + \|v\|_2) \tag{14}$$

and

$$\|u - u_{h_k}\|_0 + \|v - v_{h_k}\|_0 \le Ch_k^2(\|u\|_2 + \|v\|_2), \tag{15}$$

where C is an absolute constant.

3. Multigrid convergence analysis for the optimal system

The aim of this section is to provide the convergence theory by multigrid methods for finding the solution $(u_J, v_J) \in \mathbf{V}_J$ such that

$$\mathcal{A}(u_J, v_J; w_J, z_J) = \frac{1}{\delta}(f, z_J) \quad \text{for all} \quad (w_J, z_J) \in \mathbf{V}_J. \tag{16}$$

From now, we drop the subindex J without mention.

Let $\mathbf{A}_k$ and $\widehat{\mathbf{A}}_k (k = 1, \ldots, J)$ be the matrix representations of the form $\mathcal{A}(\cdot, \cdot; \cdot, \cdot)$ and $\widehat{\mathcal{A}}(\cdot, \cdot; \cdot)$ on $\mathbf{V}_k \times \mathbf{V}_k$ with respect to the mesh-dependent inner product $(\cdot, \cdot)_k$ respectively. Let $\mathbf{D}_k = \mathbf{A}_k - \widehat{\mathbf{A}}_k$ be the matrix with respect to $\mathcal{D}(\cdot, \cdot; \cdot, \cdot)$. Then for all $\mathbf{w}, \mathbf{z} \in \mathbf{V}_k$

$$(\mathbf{A}_k \mathbf{w}, \mathbf{z})_k = \mathcal{A}(\mathbf{w}; \mathbf{z}), \quad (\widehat{\mathbf{A}}_k \mathbf{w}, \mathbf{z})_k = \widehat{\mathcal{A}}(\mathbf{w}, \mathbf{z})_k, \quad \mathbf{D}_k = \mathcal{D}(\mathbf{w}; \mathbf{z}). \tag{17}$$

We next introduce some discrete operators which play a fundamental role both in the analysis and the algorithms to be considered in this paper: Let $\mathbf{P}_k : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbf{V}_k$ and $\widehat{\mathbf{P}}_k : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbf{V}_k$ be the orthogonal projection operators with respect to $\mathcal{A}(\cdot, \cdot; \cdot, \cdot)$ and $\widehat{\mathcal{A}}(\cdot, \cdot; \cdot, \cdot)$, respectively. Indeed, for all $\mathbf{w} \in H_0^1(\Omega) \times H_0^1(\Omega)$ and $\mathbf{z} \in \mathbf{V}_k$,

$$\mathcal{A}(\mathbf{P}_k \mathbf{w}; \mathbf{z}) = \mathcal{A}(\mathbf{w}; \mathbf{z}) \quad \text{and} \quad \widehat{\mathcal{A}}(\widehat{\mathbf{P}}_k \mathbf{w}; \mathbf{z}) = \widehat{\mathcal{A}}(\mathbf{w}; \mathbf{z}).$$

The restriction operator $\mathbf{P}_{k-1}^0 : \mathbf{V}_k \to \mathbf{V}_{k-1}$ is defined by

$$(\mathbf{P}_{k-1}^0 \mathbf{w}, \mathbf{z})_{k-1} = (\mathbf{w}, \mathbf{z})_k \text{ for all } \mathbf{w}, \mathbf{z} \in \mathbf{V}_{k-1}. \tag{18}$$

Lemma 3.1. *It follows that*

$$\mathbf{A}_{k-1} \mathbf{P}_{k-1} = \mathbf{P}_{k-1}^0 \mathbf{A}_k \tag{19}$$

and

$$\widehat{\mathbf{A}}_{k-1} \widehat{\mathbf{P}}_{k-1} = \mathbf{P}_{k-1}^0 \widehat{\mathbf{A}}_k. \tag{20}$$

Proof. For the proof of (19), it is enough to see for all $\mathbf{w} \in \mathbf{V}_k$ and $\mathbf{z} \in \mathbf{V}_{k-1}$ that

$$\begin{aligned} (\mathbf{A}_{k-1} \mathbf{P}_{k-1} \mathbf{w}, \mathbf{z})_{k-1} &= \mathcal{A}(\mathbf{P}_{k-1} \mathbf{w}; \mathbf{z}) = \mathcal{A}(\mathbf{w}; \mathbf{z}) \\ &= (\mathbf{A}_k \mathbf{w}, \mathbf{z})_k = (\mathbf{P}_{k-1}^0 \mathbf{A}_k \mathbf{w}, \mathbf{z})_{k-1}. \end{aligned}$$

In a similar way, we have (20). □

Now define a scale of mesh-dependent norms $\| \cdot \|_{\hat{s},k}$ as

$$\|\mathbf{w}\|_{\hat{s},k} = \sqrt{(\widehat{\mathbf{A}}_k^s \mathbf{w}, \mathbf{w})_k} \;\text{ for all } \mathbf{w} \in \mathbf{V}_k. \tag{21}$$

We remark that $\|\cdot\|_0$ and $\|\cdot\|_{\hat{0},k}$ are equivalent (see [13]) and that $\|\cdot\|_1$ is equivalent to $\| \cdot \|_{\widehat{\mathcal{A}}}$, which allows us to assume $\|\mathbf{w}\|_1 = \|\mathbf{w}\|_{\widehat{\mathcal{A}}}$ for all $\mathbf{w} \in H_0^1(\Omega) \times H_0^1(\Omega)$.

Now let us recall the known multigrid algorithm here. For this, let $\widehat{\mathbf{R}}_k = \mathbf{R}_k$ be a symmetric relaxation operator.

Multigrid Algorithm:
Let $\mathbf{z} \in \mathbf{V}_J$. Then $\widehat{\mathbf{P}}_k\mathbf{z} \in \mathbf{V}_k$. Consider $\widehat{\mathbf{A}}_k\widehat{\mathbf{P}}_k\mathbf{z} = \mathbf{g}$ for $\mathbf{g} \in \mathbf{V}_k$.

(1) Define $\widehat{\mathbf{B}}_1 = \widehat{\mathbf{A}}_1^{-1}$.
(2) Set $\mathbf{z}_0 = 0$.
(3) For $1 \le l \le m_1$, let $\mathbf{z}_l = \mathbf{z}_{l-1} + \widehat{\mathbf{R}}_k(\mathbf{g} - \widehat{\mathbf{A}}_k\widehat{\mathbf{P}}_k\mathbf{z}_{l-1})$.
(4) Define $\mathbf{z}_{m_1+1} = \mathbf{z}_{m_1} + \mathbf{q}$, where $\mathbf{q}$ is defined by $\mathbf{q} = \widehat{\mathbf{B}}_{k-1}\mathbf{P}^0_{k-1}(\mathbf{g} - \widehat{\mathbf{A}}_k\widehat{\mathbf{P}}_k\mathbf{z}_{m_1})$.
(5) For $m_1 + 2 \le l \le m_1 + m_2 + 1$, let $\mathbf{z}_l = \mathbf{z}_{l-1} + \widehat{\mathbf{R}}_k(\mathbf{g} - \widehat{\mathbf{A}}_k\widehat{\mathbf{P}}_k\mathbf{z}_{l-1})$.
(6) Define $\widehat{\mathbf{B}}_k\mathbf{g} = \mathbf{z}_{m_1+m_2+1}$.

We may assume presmoothing process only, that is $m_1 = 1$ and $m_2 = 0$. First we consider the symmetric positive definite part of the given coupled optimality system (4). Then we will discuss the whole coupled optimality system (4).

For $k > 1$, let $\widehat{\mathbf{K}}_k := \mathbf{I} - \widehat{\mathbf{R}}_k\widehat{\mathbf{A}}_k$ (defined on $\mathbf{V}_k$) and $\widehat{\mathbf{T}}_k := \widehat{\mathbf{R}}_k\widehat{\mathbf{A}}_k\widehat{\mathbf{P}}_k$ (defined on $\mathbf{V}_J$). Set $\widehat{\mathbf{T}}_1 = \widehat{\mathbf{P}}_1$. Note that $\widehat{\mathbf{P}}_{k-1}\widehat{\mathbf{P}}_k = \widehat{\mathbf{P}}_{k-1}$. For $\mathbf{z} \in \mathbf{V}_J$, it follows that

$$\mathbf{z} - \mathbf{z}_1 = \mathbf{z} - \widehat{\mathbf{R}}_k\mathbf{g} = \mathbf{z} - \widehat{\mathbf{R}}_k\widehat{\mathbf{A}}_k\widehat{\mathbf{P}}_k\mathbf{z} = (\mathbf{I} - \widehat{\mathbf{R}}_k\widehat{\mathbf{A}}_k\widehat{\mathbf{P}}_k)\mathbf{z} = (\mathbf{I} - \widehat{\mathbf{T}}_k)\mathbf{z}. \tag{22}$$

Using (22) it follows that for $\mathbf{z} \in \mathbf{V}_J$

$$\begin{aligned} (\mathbf{I} - \widehat{\mathbf{B}}_k\widehat{\mathbf{A}}_k\widehat{\mathbf{P}}_k)(\mathbf{z} - \mathbf{z}_0) &= \mathbf{z} - \widehat{\mathbf{B}}_k\mathbf{g} = \mathbf{z} - (\mathbf{z}_1 + \mathbf{q}) \\ &= \mathbf{z} - \mathbf{z}_1 - \widehat{\mathbf{B}}_{k-1}\mathbf{P}^0_{k-1}(\mathbf{g} - \widehat{\mathbf{A}}_k\widehat{\mathbf{P}}_k\mathbf{z}_1) \\ &= \mathbf{z} - \mathbf{z}_1 - \widehat{\mathbf{B}}_{k-1}\mathbf{P}^0_{k-1}\widehat{\mathbf{A}}_k\widehat{\mathbf{P}}_k(\mathbf{z} - \mathbf{z}_1) \\ &= \mathbf{z} - \mathbf{z}_1 - \widehat{\mathbf{B}}_{k-1}\widehat{\mathbf{A}}_{k-1}\widehat{\mathbf{P}}_{k-1}\widehat{\mathbf{P}}_k(\mathbf{z} - \mathbf{z}_1) \\ &= (\mathbf{I} - \widehat{\mathbf{B}}_{k-1}\widehat{\mathbf{A}}_{k-1}\widehat{\mathbf{P}}_{k-1})(\mathbf{I} - \widehat{\mathbf{T}}_k)\mathbf{z}. \end{aligned} \tag{23}$$

The convergence results of the multigrid method will be expressed in terms of the error operators $\widehat{\mathbf{E}}_k = \mathbf{I} - \widehat{\mathbf{B}}_k\widehat{\mathbf{A}}_k\widehat{\mathbf{P}}_k$ and $\widehat{\mathbf{E}} := \widehat{\mathbf{E}}_J$. Since $\widehat{\mathbf{E}}_k = \widehat{\mathbf{E}}_{k-1}(\mathbf{I} - \widehat{\mathbf{T}}_k)$, we have

$$\widehat{\mathbf{E}}_k = (\mathbf{I} - \widehat{\mathbf{T}}_1)(\mathbf{I} - \widehat{\mathbf{T}}_2)\cdots(\mathbf{I} - \widehat{\mathbf{T}}_k). \tag{24}$$

To prove convergence of multigrid for the optimal control optimality system (4), we need to provide the convergence of multigrid algorithms for the decoupled symmetric system which is the symmetric part of (4). In this case, the decoupled symmetric system consists of two simple elliptic Dirichlet boundary value problem such as

$$\begin{aligned} -\nabla \cdot (B\nabla u) + u &= g_1 \quad \text{in } \Omega \\ -\nabla \cdot (B\nabla v) + v &= g_2 \quad \text{in } \Omega \\ u = v &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

Hence, it is enough to adopt the results of [9] and [11], in which the following assumptions (A.1) - (A.3) are used (see [8, 10, 11]). We assume that there is a constant $C_{\overline{\widehat{\mathbf{R}}}}$ such that

$$\frac{(\mathbf{w},\mathbf{w})_k}{\lambda_k} \le C_{\overline{\widehat{\mathbf{R}}}}(\overline{\widehat{\mathbf{R}}}_k\mathbf{w},\mathbf{w})_k \quad \text{for all } \mathbf{w}\in\mathbf{V}_k, \tag{A.1}$$

where $\overline{\widehat{\mathbf{R}}}_k = (\mathbf{I}-\widehat{\mathbf{K}}_k^*\widehat{\mathbf{K}}_k)\widehat{\mathbf{A}}_k^{-1}$ and λ_k is the largest eigenvalue of $\widehat{\mathbf{A}}_k$. Here and in the remainder of this paper, $*$ denotes the adjoint with respect to the inner product $\widehat{\mathcal{A}}(\cdot,\cdot)$. There is a constant $\theta < 2$ not depending on k satisfying

$$\widehat{\mathcal{A}}(\widehat{\mathbf{T}}_k\mathbf{w},\widehat{\mathbf{T}}_k\mathbf{w}) \le \theta\widehat{\mathcal{A}}(\widehat{\mathbf{T}}_k\mathbf{w},\mathbf{w}) \quad \text{for all } \mathbf{w}\in\mathbf{V}_k. \tag{A.2}$$

Then $\|\mathbf{I}-\widehat{\mathbf{T}}_k\| < 1$ can be shown under (A.2) in the following way. For any $\mathbf{w}\in\mathbf{V}_J$,

$$\begin{aligned}\widehat{\mathcal{A}}((\mathbf{I}-\widehat{\mathbf{T}}_k)\mathbf{w},(\mathbf{I}-\widehat{\mathbf{T}}_k)\mathbf{w}) &= \widehat{\mathcal{A}}(\mathbf{w},\mathbf{w}) - 2\widehat{\mathcal{A}}(\widehat{\mathbf{T}}_k\mathbf{w},\mathbf{w}) + \widehat{\mathcal{A}}(\widehat{\mathbf{T}}_k\mathbf{w},\widehat{\mathbf{T}}_k\mathbf{w})\\ &\le \widehat{\mathcal{A}}(\mathbf{w},\mathbf{w}) - (2-\theta)\widehat{\mathcal{A}}(\widehat{\mathbf{T}}_k\mathbf{w},\mathbf{w}) \le \widehat{\mathcal{A}}(\mathbf{w},\mathbf{w}).\end{aligned} \tag{25}$$

The final assumption is that for $k > 1$, there exists a constant $C_{\widehat{\mathbf{T}}}$ satisfying

$$(\widehat{\mathbf{T}}_k\mathbf{w},\widehat{\mathbf{T}}_k\mathbf{w})_k \le C_{\widehat{\mathbf{T}}}\lambda_k^{-1}\widehat{\mathcal{A}}(\widehat{\mathbf{T}}_k\mathbf{w},\mathbf{w}) \quad \text{for all } \mathbf{w}\in\mathbf{V}_k, \tag{A.3}$$

or equivalently

$$(\widehat{\mathbf{R}}_k\mathbf{w},\mathbf{w})_k \le C_R\lambda_k^{-1}\widehat{\mathcal{A}}(\mathbf{w},\mathbf{w}) \quad \text{for all } \mathbf{w}\in\mathbf{V}_k. \tag{26}$$

Then, following the same arguments in [9] and [11], we have the convergence statement.

Theorem 3.2. *Let* $\widehat{\mathbf{R}}_k$ *be any symmetric smoother satisfying* (A.1) *and* (A.2) *for* $k > 1$. *Then there exists a positive constants* $\widehat{\gamma} < 1$ *not depending on* J *such that*

$$\|\widehat{\mathbf{E}}\mathbf{w}\|_{\widehat{\mathcal{A}}} \le \widehat{\gamma}\ \|\mathbf{w}\|_{\widehat{\mathcal{A}}} \quad \textit{for all} \quad \mathbf{w}\in\mathbf{V_J}.$$

Now, let us turn to the multigrid algorithm corresponding to coupled optimality system (4) which is definitely nonsymmetric. Note that it has the same recursive form as (23) with $\mathbf{B}_k, \mathbf{E}_k$, etc., instead of $\widehat{\mathbf{B}}_k, \widehat{\mathbf{E}}_k$, etc., and thus

$$\mathbf{E}_k = (\mathbf{I}-\mathbf{T}_1)(\mathbf{I}-\mathbf{T}_2)\cdots(\mathbf{I}-\mathbf{T}_k). \tag{27}$$

To analyze the multigrid algorithm, we use the perturbation operator $\mathbf{Z}_k$ such that

$$\mathbf{T}_k = \widehat{\mathbf{T}}_k + \mathbf{Z}_k.$$

Lemma 3.3. *For any* $\mathbf{w},\mathbf{z}\in\mathbf{V}_J$ *and* $k > 1$,

$$\widehat{\mathcal{A}}(\mathbf{Z}_k\mathbf{w};\mathbf{z}) = \mathcal{D}(\mathbf{w};\widehat{\mathbf{T}}_k\mathbf{z}), \tag{28}$$

and for $k = 1$,

$$\widehat{\mathcal{A}}(\mathbf{Z}_1\mathbf{w};\mathbf{z}) = \mathcal{D}((\mathbf{I}-\mathbf{P}_1)\mathbf{w};\widehat{\mathbf{P}}_1\mathbf{z}). \tag{29}$$

Proof. Since for any $\mathbf{w}, \mathbf{z} \in \mathbf{V}_J$ and $k > 1$

$$\begin{aligned}\widehat{\mathcal{A}}(\mathbf{T}_k\mathbf{w}; \mathbf{z}) &= (\mathbf{T}_k\mathbf{w}, \widehat{\mathbf{A}}_k\widehat{\mathbf{P}}_k\mathbf{z})_k = (\mathbf{A}_k\mathbf{P}_k\mathbf{w}, \mathbf{R}_k\widehat{\mathbf{A}}_k\widehat{\mathbf{P}}_k\mathbf{z})_k \\ &= (\mathbf{A}_k\mathbf{P}_k\mathbf{w}, \widehat{\mathbf{T}}_k\mathbf{z})_k = \mathcal{A}(\mathbf{w}; \widehat{\mathbf{T}}_k\mathbf{z}) \\ &= \widehat{\mathcal{A}}(\mathbf{w}; \widehat{\mathbf{T}}_k\mathbf{z}) + \mathcal{D}(\mathbf{w}; \widehat{\mathbf{T}}_k\mathbf{z}) \text{ for all } \mathbf{w}, \mathbf{z} \in \mathbf{V}_J,\end{aligned}$$

(28) comes from the definition of $\mathbf{Z}_k$. Since

$$\begin{aligned}\widehat{\mathcal{A}}(\mathbf{P}_1\mathbf{w}; \mathbf{z}) &= \widehat{\mathcal{A}}(\mathbf{P}_1\mathbf{w}; \widehat{\mathbf{P}}_1\mathbf{z}) \\ &= \mathcal{A}(\mathbf{w}; \widehat{\mathbf{P}}_1\mathbf{z}) - \mathcal{D}(\mathbf{P}_1\mathbf{w}; \widehat{\mathbf{P}}_1\mathbf{z}) \\ &= \widehat{\mathcal{A}}(\mathbf{w}; \widehat{\mathbf{P}}_1\mathbf{z}) + \mathcal{D}(\mathbf{w}; \widehat{\mathbf{P}}_1\mathbf{z}) - \mathcal{D}(\mathbf{P}_1\mathbf{w}; \widehat{\mathbf{P}}_1\mathbf{z}) \\ &= \widehat{\mathcal{A}}(\widehat{\mathbf{P}}_1\mathbf{w}; \mathbf{z}) + \mathcal{D}((\mathbf{I} - \mathbf{P}_1)\mathbf{w}; \widehat{\mathbf{P}}_1\mathbf{z}),\end{aligned}$$

(29) comes immediately by noting $\mathbf{P}_1 = \mathbf{T}_1$ and $\widehat{\mathbf{P}}_1 = \widehat{\mathbf{T}}_1$. □

Lemma 3.4. *Assume that* (A.1)-(A.3) *hold. For* $k > 0$*, we get*

$$\|\mathbf{Z}_k\|_{\widehat{\mathcal{A}}} \leq C_Z h_k,$$

where C_Z *is a constant not dependent on* k*.*

Proof. Let $\mathbf{w} \in \mathbf{V}_J$. Using the $\mathcal{A}$- and $\widehat{\mathcal{A}}$-orthogonal property of $\mathbf{P}_k$ and $\widehat{\mathbf{P}}_k$ respectively, we have

(30)
$$\|(\mathbf{I} - \mathbf{P}_k)\mathbf{w}\|_0 \leq Ch_k\|(\mathbf{I} - \mathbf{P}_k)\mathbf{w}\|_1 \leq Ch_k\|\mathbf{w}\|_1 \quad \text{and} \quad \|\widehat{\mathbf{P}}_k\|_1 \leq C \text{ for all } k > 0.$$

Applying (12) and (30) to (29) give

(31)
$$|\widehat{\mathcal{A}}(\mathbf{Z}_1\mathbf{w}; \mathbf{z})| = |\mathcal{D}((\mathbf{I} - \mathbf{P}_1)\mathbf{w}; \widehat{\mathbf{P}}_1\mathbf{z})| \leq C\|(\mathbf{I} - \mathbf{P}_1)\mathbf{w}\|_0\|\widehat{\mathbf{P}}_k\mathbf{z}\|_1 \leq Ch_1\|\mathbf{w}\|_1\|\mathbf{z}\|_1.$$

Now, consider $k > 1$. Applying (11) and (A.3) to (28) give

$$\begin{aligned}|\widehat{\mathcal{A}}(\mathbf{Z}_k\mathbf{w}; \mathbf{z})| &= |\mathcal{D}(\mathbf{w}; \widehat{\mathbf{T}}_k\mathbf{z})| \leq C_d\|\mathbf{w}\|_1\|\widehat{\mathbf{T}}_k\mathbf{z}\|_0 \leq C\|\mathbf{w}\|_1(\widehat{\mathbf{T}}_k\mathbf{z}; \widehat{\mathbf{T}}_k\mathbf{z})_k^{\frac{1}{2}} \\ &\leq Ch_k\|\mathbf{w}\|_1\widehat{\mathcal{A}}(\widehat{\mathbf{T}}_k\mathbf{z}; \mathbf{z})_k^{\frac{1}{2}} \leq Ch_k\|\mathbf{w}\|_1\|\mathbf{z}\|_1.\end{aligned}$$

Therefore, we have the conclusion. □

Theorem 3.5. *Assume that* (A.1)-(A.3) *hold. There exists a* h_0 *such that for all* $h_1 < h_0$*,*

$$\|\mathbf{E}\mathbf{w}\|_{\widehat{\mathcal{A}}} \leq \gamma^*\|\mathbf{w}\|_{\widehat{\mathcal{A}}} \quad \textit{for all} \;\; \mathbf{w} \in \mathbf{V}_J,$$

where $\gamma^* = \widehat{\gamma} + Ch_1 < 1$ *and* $\widehat{\gamma}$ *is in the Theorem 3.2.*

Proof. It is from (25) and Lemma 3.4 that the $\widehat{\mathcal{A}}$-norm of $(\mathbf{I} - \mathbf{T}_k) = (\mathbf{I} - \widehat{\mathbf{T}}_k - \mathbf{Z}_k)$ is less than or equal to $1 + C_Z h_k$. Hence, it follows that

$$\|\mathbf{E}_k\|_{\widehat{\mathcal{A}}} \leq \prod_{i=1}^{k}(1 + C_Z h_i) < C_1 \text{ for some } C_1 \text{ not dependent on } k.$$

Consider the difference of the error operators:

$$
\begin{aligned}
\mathbf{E}_k - \widehat{\mathbf{E}}_k &= \mathbf{E}_{k-1}(\mathbf{I} - \mathbf{T}_k) - \widehat{\mathbf{E}}_{k-1}(\mathbf{I} - \widehat{\mathbf{T}}_k) (\text{by (24) and (27)}) \\
&= \mathbf{E}_{k-1}(\mathbf{I} - \widehat{\mathbf{T}}_k) - \widehat{\mathbf{E}}_{k-1}(\mathbf{I} - \widehat{\mathbf{T}}_k) - \mathbf{E}_{k-1}\mathbf{T}_k + \mathbf{E}_{k-1}\widehat{\mathbf{T}}_k \\
&= (\mathbf{E}_{k-1} - \widehat{\mathbf{E}}_{k-1})(\mathbf{I} - \widehat{\mathbf{T}}_k) - \mathbf{E}_{k-1}\mathbf{Z}_k.
\end{aligned} \tag{32}
$$

By (25) and Lemma 3.4, for $k > 1$,

$$
\|\mathbf{E}_k - \widehat{\mathbf{E}}_k\|_{\widehat{\mathcal{A}}} \le \|\mathbf{E}_{k-1} - \widehat{\mathbf{E}}_{k-1}\|_{\widehat{\mathcal{A}}} + C_1 C_Z h_k. \tag{33}
$$

Repetitively applying (33) and using

$$
\|\mathbf{E}_1 - \widehat{\mathbf{E}}_1\|_{\widehat{\mathcal{A}}} = \|\mathbf{Z}_1\|_{\widehat{\mathcal{A}}} \le C_Z h_1
$$

gives that, for some C not dependent on k,

$$
\|\mathbf{E} - \widehat{\mathbf{E}}\|_{\widehat{\mathcal{A}}} \le C_Z h_1 + C_1 C_Z \sum_{k=2}^{J} h_k \le C \sum_{k=1}^{\infty} h_k = C \sum_{k=1}^{\infty} 2^{1-k} h_1 \le C h_1.
$$

The results follow from the triangle inequality and Theorem 3.2. □

4. Numerical experiments

TABLE 1. The values are derived from (35) and (36) when $\alpha = 1$.

δ	$\|u - \hat{u}\|_0$	$\|\theta\|_0$	$\mathcal{J}(u,\theta)$
1	4.9884e−001	2.4053e−002	1.2471e−001
10^{-1}	4.8863e−001	2.3561e−001	1.2215e−001
10^{-2}	4.0568e−001	1.9561e+000	1.0147e−001
10^{-3}	1.5037e−001	7.2508e+000	3.7594e−002
10^{-4}	2.0618e−002	9.9419e+000	5.1547e−003
10^{-5}	2.1413e−003	1.0325e+001	5.3534e−004
10^{-6}	2.1496e−004	1.0365e+001	5.3741e−005
10^{-7}	2.1504e−005	1.0369e+001	5.3762e−006
10^{-8}	2.1505e−006	1.0369e+001	5.3764e−007
10^{-9}	2.1505e−007	1.0369e+001	5.3764e−008

In this section, we present the usage of multigrid methods applied to a discretized optimal control problem for two dimensional diffusion-convection problem. We focus on the role of weight parameters α and δ in the quadratic functional $\mathcal{J}(u_h, \theta_h)$ in the sense of how it affects the convergence rate. We report both the L^2 error of the target velocity $\widehat{u}$ and the approximate velocity u_h from multigrid methods and the value of the quadratic functional $\mathcal{J}(u_h, \theta_h)$ where θ_h is understood as the approximate value corresponding to (2). The computational domain is triangularized uniformly with the grid interval h ranging from 2^{-2} to 2^{-5} for each direction. We use the single approximation space of continuous piecewise linear polynomials for the approximations of all unknowns. We use the preconditioned Richardson method for smoothing iteration and fix $m_1 = m_2 = 1$ in the multigrid algorithm. We easily see that this relaxation scheme satisfies the assumptions of Theorem 3.5. (See [8]). We set the tolerance of the errors to be 10^{-6} and the maximum number of iterations to be 500. In order to see that the multigrid methods works well for

TABLE 2. The numerical results with $\alpha = 1$ when $h = 1/32$.

δ	$\|u_h - \hat{u}\|_0$	$\|\theta_h\|_0$	$\mathcal{J}(u_h, \theta_h)$
1	4.9885e−001	2.3984e−002	1.2471e−001
10^{-1}	4.8870e−001	2.3508e−001	1.2217e−001
10^{-2}	4.0597e−001	1.9545e+000	1.0151e−001
10^{-3}	1.5074e−001	7.2607e+000	3.7721e−002
10^{-4}	2.0795e−002	9.9598e+000	5.1761e−003
10^{-5}	2.2087e−003	1.0347e+001	5.3782e−004
10^{-6}	4.8266e−004	1.0389e+001	5.4082e−005
10^{-7}	4.2144e−004	1.0394e+001	5.4914e−006
10^{-8}	4.1711e−004	1.0400e+001	6.2780e−007
10^{-9}	4.1660e−004	1.0412e+001	1.4099e−007

whole system (4), we take the unit square $\Omega = (0,1) \times (0,1)$ as domains and $\hat{u}(x,y) = \sin \pi x \sin \pi y$ as a desired state (target velocity).

Example 1. For the first numerical experiments for our optimal control problem, we choose **B** and **b** as the identity matrix and zero vector in (4) respectively. Then, the optimality system becomes

$$
\begin{cases}
-\Delta u + u - \dfrac{v}{\delta} = & 0 \quad \text{in} \quad \Omega, \\
-\Delta v + v + \alpha u = & \alpha \sin \pi x \sin \pi y \quad \text{in} \quad \Omega, \\
u = & 0 \quad \text{on} \quad \partial\Omega, \\
v = & 0 \quad \text{on} \quad \partial\Omega.
\end{cases}
\tag{34}
$$

The exact optimality solutions can easily obtained as

$$
u = \frac{1}{1 + \frac{\delta}{\alpha}(1+2\pi^2)^2} \sin \pi x \sin \pi y, \quad v = \frac{\delta(1+2\pi^2)}{1 + \frac{\delta}{\alpha}(1+2\pi^2)^2} \sin \pi x \sin \pi y.
$$

Now, we consider the quadratic functional in (1). Note that all terms of (1)

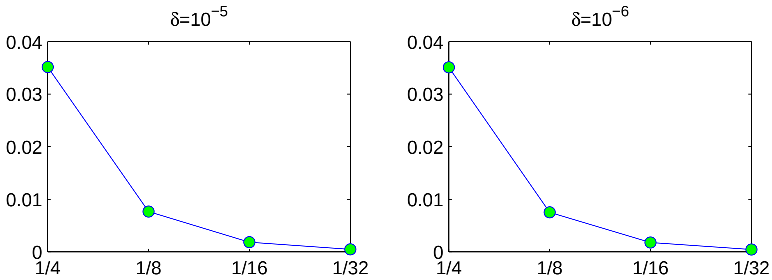

FIGURE 1. The norms $\|u_h - \hat{u}\|_0$ for $\delta = 10^{-5}$ and $\delta = 10^{-6}$ when $h = 1/32$.

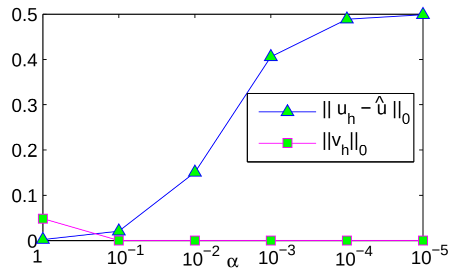

FIGURE 2. The norms $\|u_h - \hat{u}\|_0$ and $\|v_h\|_0$ with $\delta = 10^{-5}$ when $h = 1/32$.

can be easily calculated exactly. Indeed, we have

$$\frac{\alpha}{2}\int_\Omega |u-\hat{u}|^2 d\Omega = \frac{\alpha}{8}\left(1 - \frac{1}{1+\frac{\delta}{\alpha}(1+2\pi^2)^2}\right)^2 \tag{35}$$

and

$$\frac{\delta}{2}\int_\Omega |\theta|^2 d\Omega = \frac{\delta(1+2\pi^2)^2}{8(1+\frac{\delta}{\alpha}(1+2\pi^2)^2)^2}. \tag{36}$$

From (35) and (36) (or seeing the construction of target velocity and exact solution), we know immediately that if we choose $\alpha = 1$ and $\delta = 0$, then the target velocity $\hat{u}$ and the exact solution u matches completely, but we can not make $\delta = 0$ because of the relationship $\theta = \frac{v}{\delta}$. Hence we will make δ approaches 0. From now on, we consider the weights of the costs satisfying the relationship, without loss of generality,

$$0 < \delta \le \alpha = 1.$$

First, we display the values of each term in (1) with the quadratic functional in Table 1 with δ ranging from 1 to 10^{-8} for fixed $\alpha = 1$.

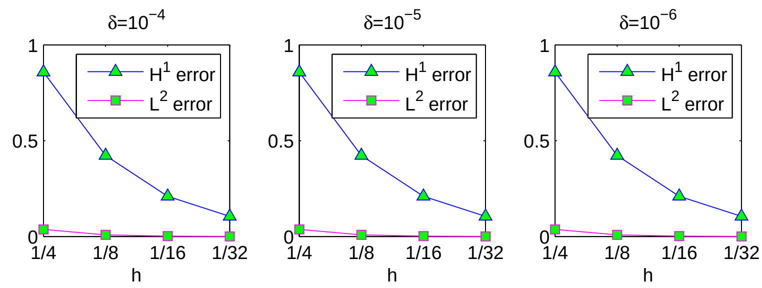

FIGURE 3. The graph of L^2 and H^1 errors when $\alpha = 1$.

TABLE 3. The convergence rates for $L^2(\Omega)$ and $H^1(\Omega)$-norms errors when $\alpha = 1$.

δ	h	L^2	H^1	δ	h	L^2	H^1
1	$\frac{1}{8}$	1.905	0.964	10^{-5}	$\frac{1}{8}$	2.196	1.049
	$\frac{1}{16}$	1.975	0.991		$\frac{1}{16}$	2.063	1.017
	$\frac{1}{32}$	1.993	0.997		$\frac{1}{32}$	2.016	1.004
10^{-1}	$\frac{1}{8}$	1.876	0.977	10^{-6}	$\frac{1}{8}$	2.226	1.055
	$\frac{1}{16}$	1.967	0.995		$\frac{1}{16}$	2.093	1.022
	$\frac{1}{32}$	1.991	0.999		$\frac{1}{32}$	2.026	1.006
10^{-2}	$\frac{1}{8}$	1.878	0.984	10^{-7}	$\frac{1}{8}$	2.229	0.105
	$\frac{1}{16}$	1.969	0.997		$\frac{1}{16}$	2.110	0.102
	$\frac{1}{32}$	1.992	0.999		$\frac{1}{32}$	2.041	0.100
10^{-3}	$\frac{1}{8}$	2.002	0.971	10^{-8}	$\frac{1}{8}$	2.229	0.105
	$\frac{1}{16}$	2.002	0.993		$\frac{1}{16}$	2.112	0.102
	$\frac{1}{32}$	2.000	0.998		$\frac{1}{32}$	2.052	0.100
10^{-4}	$\frac{1}{8}$	2.129	1.021	10^{-9}	$\frac{1}{8}$	2.229	1.054
	$\frac{1}{16}$	2.040	1.008		$\frac{1}{16}$	2.112	1.022
	$\frac{1}{32}$	2.010	1.002		$\frac{1}{32}$	2.054	1.008

In Table 2, we present numerical results for the functional values using the V cycle multigrid method when $h = 1/32$ in the sense of L^2. According to Tables 1 and 2, we figure out that the approximate solution $u_h(\delta)$ converges slowly to the target velocity $\hat{u}$ after $\delta \leq 10^{-6}$ though the approximate control norm $||\theta_h||_0$ converges to the norm $||\theta||_0$ as $\delta \to 0$. Thus the values of the error $\|u_h - \hat{u}\|_0$ are not agreeable to those of Table 1 after 10^{-6}. Let us explain why this phenomenon can be occurred. If v is replaced with $-\delta\Delta u + \delta u$ in the second equation of (34), we get the 4^{th} order elliptic boundary value problem

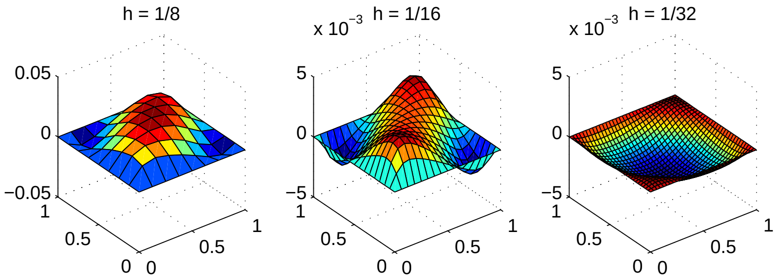

FIGURE 4. The pointwise error graph of $u_h - \hat{u}$ with h when $\alpha = 1$, $\delta = 10^{-5}$.

with zero boundary condition

$$-\Delta\Delta u + 2\Delta u - \left(1 + \frac{\alpha}{\delta}\right) u = -f. \tag{37}$$

Using the standard finite element analysis for (37), we can induce that for $u \in H^2(\Omega)$ and $\alpha = 1$

$$\|u - u_h\|_0 \leq C\left(1 + \frac{1}{\delta}\right)\|u\|_2, \tag{38}$$

where C is independent of h. Now we can explain why the computed solution u_h can not approach to $\hat{u}$ as δ approaches 0, by considering $\|u_h - \hat{u}\|_0$ which is bounded by $\|u - u_h\|_0 + \|u - \hat{u}\|_0$. As shown in (35) (or see Table 1), the values of $\|u - \hat{u}\|_0$ approach to 0 as δ goes to zero. But from the error analysis (38) we cannot affirm that the value $\|u - u_h\|_0$ are affected by δ when δ is small. As we see in Figure 1, the values of $\|u - u_h\|_0$ are not influenced by the parameter δ. It is proper that we select the acceleration parameter $\delta = 10^{-5}$ for fixed

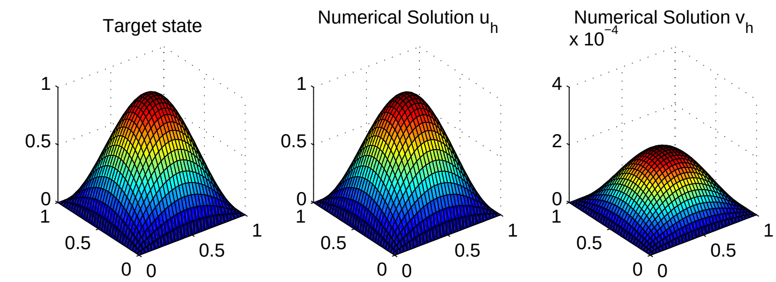

FIGURE 5. The graphes of the target function and numerical solution u_h and v_h for $\alpha = 1$, $\delta = 10^{-5}$ when $h = 1/32$.

$\alpha = 1$ according to Tables 1 and 2. Now, we fix $\delta = 10^{-5}$ for a moment to investigate the effect of parameter α. Figure 2 shows the norm $\|u_h - \hat{u}\|_0$ of the error between the controlled optimal solution u_h and the desired state $\hat{u}$ and the norm $\|v_h\|_0$ related to the control function plotted for different values of α with a fixed $\delta = 10^{-5}$. As expected, as α becomes smaller, the norm v_h is forced to become smaller, and as is shown in Figure 2, the error $\|u_h - \hat{u}\|_0$ becomes less controlled. In this point of view, the choice of $\alpha = 1$ and $\delta = 10^{-5}$ seems to be the best choice among the presented results. With these choices, we exhibit L^2-error $\|u - u_h\|_0 + \|v - v_h\|_0$ and H^1- error $\|u - u_h\|_1 + \|v - v_h\|_1$ in Figure 3. Then convergence rates reported in Table 3 for the approximated solution u_h and v_h by the multigrid V-cycle are measured by

$$\log_2 \frac{\|u - u_h\|_0 + \|v - v_h\|_0}{\|u - u_{h/2}\|_0 + \|v - v_{h/2}\|_0} \text{ and } \log_2 \frac{\|u - u_h\|_1 + \|v - v_h\|_1}{\|u - u_{h/2}\|_1 + \|v - v_{h/2}\|_1}.$$

It is from Table 3 that the convergence rates are like $O(h^2)$ in L^2 and $O(h)$ in H^1. These facts imply that the resulting convergence rates asymptotically approach to those of the theoretically predicted. Figure 4 illustrates that the numerical optimal solution u_h converging to the objective function $\hat{u}$, which can be evidently by pointwise error figure between u_h and $\hat{u}$. Specially, we display the target function and numerical solutions u_h and v_h for $\alpha = 1$, $\delta = 10^{-5}$ when $h = \frac{1}{32}$ in Figure 5.

TABLE 4. The numerical results for the quadratic functional with $\alpha = 1$ when $h = 1/32$.

δ	$\|u_h - \hat{u}\|_0$	$\|\theta_h\|_0$	$\mathcal{J}(u_h, \theta_h)$
1	4.9887e−001	2.3727e−002	1.2471e−001
10^{-1}	4.8894e−001	2.3253e−001	1.2223e−001
10^{-2}	4.0781e−001	1.9387e+000	1.0194e−001
10^{-3}	1.5403e−001	7.2858e+000	3.8404e−002
10^{-4}	2.2013e−002	1.0099e+001	5.3428e−003
10^{-5}	2.6502e−003	1.0535e+001	5.5848e−004
10^{-6}	5.8359e−004	1.0597e+001	5.6322e−005
10^{-7}	4.2769e−004	1.0613e+001	5.7240e−006
10^{-8}	4.1734e−004	1.0623e+001	6.5143e−007
10^{-9}	4.1661e−004	1.0630e+001	1.4328e−007

Example 2. For the second numerical experiments for the model optimal problem, we choose $\mathbf{B}$ as the identity matrix and the convection coefficient $\mathbf{b} = [1, 1]^t$. First note that with the above $\mathbf{B}$ and $\mathbf{b}$ the constraint equation (2) becomes a uniformly positive definite system. The corresponding optimality

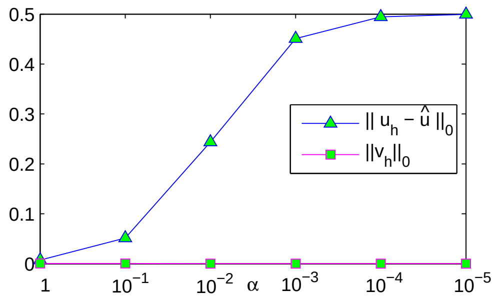

FIGURE 6. The norms $\|u_h - \hat{u}\|_0$ and $\|v_h\|_0$ with $\delta = 10^{-7}$ when $h = 1/32$.

system is now

$$(39)\qquad \begin{aligned} -\Delta u + \mathbf{b}\cdot\nabla u + u - \frac{1}{\delta}v &= 0 && \text{in} \quad \Omega, \\ -\Delta v - \nabla\cdot(v\mathbf{b}) + v + \alpha u &= \alpha \sin\pi x \sin\pi y && \text{in} \quad \Omega, \\ u = 0, v &= 0 && \text{on} \quad \partial\Omega. \end{aligned}$$

It is hard to solve the optimality equation (39) analytically. Moreover, it is well known that for convection dominated problems standard finite element discretizations applied to (39) lead to strongly oscillatory solutions unless the mesh size h is sufficiently small with respect to the ratio between the rate of convection of a flow and the rate of diffusion. For this reason, we adopt standard finite element discretization with stabilization used in [2, 15] and use the notations in theirs. Since the Peclet number $\text{Pe} = h||\mathbf{b}|| = \sqrt{2}h < 1$ for $h < \frac{1}{2}$, the positive stabilization parameter is equal to zero. Thus, it follows that the bilinear form (5) is the stabilization bilinear form of the optimality system (39).

As the example 1, we fix $\alpha = 1$. Then we display the errors between the numerical solution u_h and the target velocity $\hat{u}$ for $\alpha = 1$ in Table 4.

As seen in Table 4, we can choose $\delta = 10^{-7}$ as an optimal weight of the cost of the control for fixed $\alpha = 1$ if we deliberate on the stability of the norm $\|\theta_h\|_0$ of control function. In addition, we can figure out from Figure 6 that, for fixed $\delta = 10^{-7}$, $\alpha = 1$ seems to be the best choice among the presented results. Finally, we plot the pointwise error figure in Figure 7 and illustrate the target state, the numerical solution u_h and v_h when $\alpha = 1$ and $\delta = 10^{-7}$ for $h = 1/32$ in Figure 8.

5. Conclusion

We have shown that the coupled optimal system can be solved well by multigrid methods. As pointed in (38), one may not allow δ to be arbitrary small for

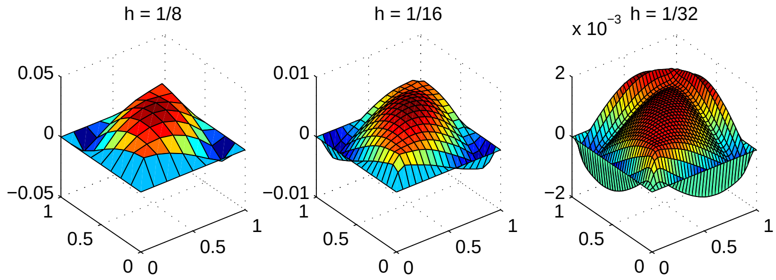

FIGURE 7. The pointwise error graph of $u_h - \hat{u}$ with h when $\alpha = 1$, $\delta = 10^{-7}$.

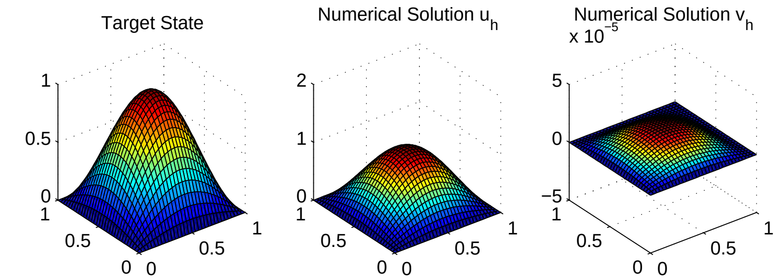

FIGURE 8. The graphes of the target function and numerical solution u_h and v_h for $\alpha = 1$, $\delta = 10^{-7}$ when $h = 1/32$.

a give mesh size. But, for a reasonable approximation to a given objective function, it is explained that it is enough to choose a relative small δ. For a coupled elliptic boundary value problem, the V-cycle multigrid convergence analysis is provided in case that those equations are coupled with reaction terms. One may try to provide the similar convergence analysis for those equations coupled by diffusion terms. One may also decompose the coupled optimal system (4) into an uncoupled optimal system for numerical implementations. In this case, one may have some restrictions on the weights α and δ of the cost of the controls even for a simple diffusion-reaction control problem. This topic will be dealt with in a coming paper.

Acknowledgement. We would thank the referee for carefully reading the manuscript and for suggesting improvements.

References

[1] Z.-Z. Bai and M. K. Ng, *On inexact preconditioners for nonsymmetric matrices*, SIAM J. Sci. Comput. **26** (2005), no. 5, 1710–1724.

[2] R. Becker and B. Vexler, *Optimal control of the convection-diffusion equation using stabilized finite element methods*, Numer. Math. **106** (2007), no. 3, 349–367.

[3] P. B. Bochev and M. D. Gunzburger, *Least-squares finite element methods for optimality systems arising in optimization and control problems*, SIAM J. Numer. Anal. **43** (2006), no. 6, 2517–2543

[4] A. Borzi, *Multigrid methods for parabolic distributed optimal control problems*, J. Comput. Appl. Math. **157** (2003), no. 2, 365–382.

[5] A. Borzi and K. Kunisch, *The numerical solution of the steady state solid fuel ignition model and its optimal control*, SIAM J. Sci. Comput. **22** (2000), no. 1, 263–284.

[6] A. Borzi, K. Kunisch, and D. Y. Kwak, *Accuracy and convergence properties of the finite difference multigrid solution of an optimal control optimality system*, SIAM J. Control Optim. **41** (2002), no. 5, 1477–1497.

[7] J. H. Bramble, *Multigrid Methods*, Pitman Res. Notes Math, Ser. 294, John Wiley, New York, 1993.

[8] J. H. Bramble, D. Y. Kwak, and J. E. Pasciak, *Uniform convergence of multigrid V-cycle iterations for indefinite and nonsymmetric problems*, SIAM J. Numer. Anal. **31** (1994), no. 6, 1746–1763.

[9] J. H. Bramble and J. E. Pasciak, *New estimates for multilevel algorithms including the V-cycle*, Math. Comp. **60** (1993), no. 202, 447–471.

[10] J. H. Bramble, J. E. Pasciak, and J. Xu, *The analysis of multigrid algorithms for nonsymmetric and indefinite elliptic problems*, Math. Comp. **51** (1988), no. 184, 389–414.

[11] ———, *The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms*, Math. Comp. **56** (1991), no. 193, 1–34.

[12] S. C. Brenner, *Convergence of the multigrid V-cycle algorithm for second-order boundary value problems without full elliptic regularity*, Math. Comp. **71** (2002), no. 238, 507–525.

[13] S. C. Brenner and L. R. Scott, *The mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 2002.

[14] L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, 19. American Mathematical Society, Providence, RI, 1998.

[15] J. L. Guermond, *Stabilization of Galerkin approximations of transport equations by subgrid modeling*, M2AN Math. Model. Numer. Anal. **33** (1999), no. 6, 1293–1316.

[16] M. D. Gunzburger, *Perspectives in Flow Control and Optimization*, Advances in Design and Control, 5. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2003.

[17] M. D. Gunzburger and H.-C. Lee, *Analysis, approximation, and computation of a coupled solid/fluid temperature control problem*, Comput. Methods Appl. Mech. Engrg. **118** (1994), no. 1-2, 133–152.

[18] ———, *Analysis and approximation of optimal control problems for first-order elliptic systems in three dimensions*, Appl. Math. Comput. **100** (1999), no. 1, 49–70.

[19] W. Hackbusch, *Multigrid Methods and Applications*, Springer Series in Computational Mathematics, 4. Springer-Verlag, Berlin, 1985.

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

[21] P. Neittaanmaki and D. Tiba, *Optimal Control of Nonlinear Parabolic Systems*, Marcel Dekker, New York, 1994.

[22] M. H. van Raalte and P. W. Hemker, *Two-level multigrid analysis for the convection-diffusion equation discretized by a discontinuous Galerkin method*, Numer. Linear Algebra Appl. **12** (2005), no. 5-6, 563–584.

[23] A. H. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Math. Comp. **28** (1974), 959–962.

[24] J. Wang, *Convergence analysis of multigrid algorithms for nonselfadjoint and indefinite elliptic problems*, SIAM J. Numer. Anal. **30** (1993), no. 1, 275–285.

Hunki Baek
Department of Mathematics
Kyungpook National University
Daegu 702-701, Korea
E-mail address: hkbaek@knu.ac.kr

Sang Dong Kim
Department of Mathematics
Kyungpook National University
Daegu 702-701, Korea
E-mail address: skim@knu.ac.kr

Hyung-Chun Lee
Department of Mathematics
Ajou University
Suwon 443-749, Korea
E-mail address: hclee@ajou.ac.kr