Turkish Journal of Mathematics

http://journals.tubitak.gov.tr/math/

Research Article

Turk J Math
(2018) 42: 164 – 180
© TÜBİTAK
doi:10.3906/mat-1606-111

Variational multiscale method for the optimal control problems of convection–diffusion–reaction equations

Aytekin Bayram ÇIBIK*, **Fikriye Nuray YILMAZ**
Department of Mathematics, Gazi University, Ankara, Turkey

Received: 24.06.2016 • **Accepted/Published Online:** 31.03.2017 • **Final Version:** 22.01.2018

Abstract: In this paper, we analyze a projection-based variational multiscale (VMS) method for the optimal control problems governed by the convection–diffusion–reaction equations. We derive the first-order optimality conditions by the *optimize-then-discretize* method. After expressing the discrete optimal control problem, we obtain the stability properties of state and adjoint variables. We also prove that the error in each variable is optimal. Through numerical examples, we show the efficiency of the stabilization for the solutions of the control, state, and adjoint variables.

Key words: Convection–diffusion, optimal control, finite element, VMS

1. Introduction

In this paper, we consider the optimal control problems of convection-dominated convection–diffusion–reaction equations. Let Ω be a convex polygonal domain in $\mathbb{R}^2$ with Lipschitz boundary $\Sigma = \partial\Omega$. The distributed control problem for the steady-state convection-diffusion-reaction equations with homogeneous Dirichlet boundary conditions can be stated as follows:

$$\min \quad J(y,u) = \frac{1}{2}\left\|y - y_d\right\|_\Omega^2 + \frac{\alpha}{2}\left\|u\right\|_\Omega^2 \tag{1.1}$$

$$\begin{aligned} \text{subject to} \quad -\epsilon\Delta y + \vec{b}.\nabla y + ry &= f + u \quad \text{in } \Omega, \\ y &= 0 \quad \text{on } \Sigma, \end{aligned} \tag{1.2}$$

where $f \in L^2(\Omega)$ is a fixed forcing term, $\epsilon > 0$ denotes the diffusivity, $\vec{b}$ is the fluid velocity, and r is a reaction coefficient. $\alpha > 0$ stands for the regularization parameter. Here y and u denote the state and control variables, respectively, and y_d is the desired state.

We consider the case where the diffusion coefficient ϵ is small compared to the infinite norm of the velocity field $\vec{b}$. In this case, obtaining the accurate solution is very difficult due to the dominance of convection. Thus, classical numerical methods produce nonphysical oscillations since the solution contains many scales including complex boundary and interior layers [12]. As a numerical method, the finite element method is chosen to obtain the solution of this system in this study. However, application of the standard Galerkin finite element method (GFEM) would not result in accurate numerical solutions due to the mentioned disadvantages for such

*Correspondence: abayram@gazi.edu.tr

kinds of problems. Thus, one has to consider a numerical stabilization scheme to avoid these difficulties. Some of the most used numerical stabilization techniques for flow problems are streamline upwind Galerkin (SUPG) and pressure stabilization methods, large eddy simulation (LES) methods, and variational multiscale (VMS) methods.

Analysis of some well-known stabilization techniques applied to a convection–diffusion system was given in [6] and a comparison of various stabilization techniques applied on an Oseen problem was studied in [3]. In [18], a discontinuous Galerkin finite element method (DG) with interior penalties for the optimal control problem of the convection–diffusion equation was studied and in [11] an edge-stabilized Galerkin finite element method for the same optimal control system was considered. Moreover, local error estimates for SUPG solutions of advection-dominated elliptic linear–quadratic optimal control problems were studied in [9]. Similarly, the local (DG) for the optimal control problem governed by convection–diffusion equations was analyzed in [19]. A similar study concerning the pressure stabilization technique, namely the Brezzi–Pitkaranta stabilization, was cast on a Stokes control in [5].

In this study, we solve the optimal control problem with a projection-based VMS approach similar to the idea given in [4, 13]. Within this technique, the global stabilization was added to the overall system first and then its effects were subtracted from the larger flow scales, which are defined explicitly through some projections. Thus, stabilization acts only on the smallest resolved scales for both state and adjoint equations [4].

There are two different approaches for the discretization of optimal control problems: *optimize-then-discretize* (OD) and *discretize-then-optimize* (DO). We follow here the function-based approach *optimize- then-discretize*. The optimality system consists of the state and adjoint equations as coupled by an algebraic equation.

The organization of the paper is as follows. We first recall some notational issues and preliminaries in order to define the problem and its variational form. Then we discretize the problem with the finite element method with stabilization and give the stability property of the optimal control problem. We state the a priori error for each variable next. Moreover, the extension to the case with pointwise control constraints is discussed in the following section. We conclude our study with some numerical examples to verify the effectiveness of the method.

2. Notation and preliminaries

We use the standard notations used for Sobolev and Lebesgue spaces in [1] throughout the entire study. The Sobolev space $W^{k,r}(\Omega)$ on a domain $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ is given as

$$W^{k,r}(\Omega) = \{\phi \in L^r(\Omega) : \forall |s| \leq k,\ \partial^s \phi \in L^r(\Omega)\}.$$

We denote usual inner product and norm in $L^2(\Omega)$ by $(\cdot,\cdot)$ and $\|\cdot\|$, respectively. The norm and seminorm in a Sobolev space $W^{k,r}(\Omega)$ are given by $\|\cdot\|_{k,r}$ and $|\cdot|_{k,r}$. For the special case $r = 2$, the norm in the space $W^{k,2}(\Omega) = H^k(\Omega)$ is shown by $\|\cdot\|_k$. The space $H^1(\Omega)$ is of special interest and we use it frequently throughout the study. The norm in $H^1(\Omega)$ is given by $\|y\|_1 = (\|y\| + \|\nabla y\|)^{1/2}$. We would like to recall here the dual space of $H_0^1(\Omega)$, namely the space $H^{-1}(\Omega)$ equipped with the -1-norm

$$\|g\|_{-1} = \sup_{v \in Y} \frac{|< g, v >|}{\|v\|_1}. \tag{2.1}$$

Here $< .,. >$ denotes the duality pairing. The following well-known functional vector spaces are considered to define a variational formulation of (1.2).

$$\begin{aligned} Y := H_0^1(\Omega) &= \{y \in H^1(\Omega) : y = 0 \text{ on } \Sigma\}, \\ U : &= L^2(\Omega). \end{aligned}$$

We denote the state space by Y and the control space by U. Thus, one can easily obtain the variational form of the state equation (1.2) as: Find $y \in Y$ and $u \in U$ satisfying

$$\epsilon(\nabla y, \nabla v) + (\vec{b} \cdot \nabla y + ry, v) = (f + u, v), \ \forall v \in Y. \tag{2.2}$$

We also assume the classical coercivity condition for convection-diffusion equations, which states that there exists a constant β such that $r - \frac{1}{2}\nabla.\vec{b} \geq \beta > 0$. As done in [11], the left-hand side of (2.2) could be expressed as a bilinear form and by using the coercivity assumption, standard optimal control theory [16] gives that there is a unique solution of (1.1–1.2) if and only if there exists an adjoint $p \in Y$ satisfying

$$\begin{aligned} -\epsilon\Delta p - \nabla.(\vec{b}p) + rp &= y - y_d \quad \text{in } \Omega, \\ p &= 0 \quad \text{on } \Sigma, \end{aligned} \tag{2.3}$$

and the pair (u, p) satisfies

$$(\alpha u + p, w - u) \geq 0 \ \ \forall\, w \in U. \tag{2.4}$$

Thus, the variational problem corresponding to (2.3) reads as: find $p \in Y$ satisfying

$$\epsilon(\nabla p, \nabla w) - (\nabla.(\vec{b}p), w) + (rp, w) = (y - y_d, w), \ \forall w \in Y. \tag{2.5}$$

We want to note here that, throughout the entire text, C will denote a generic constant that is independent of mesh width h unless stated otherwise.

3. Discretization with VMS finite element method

In this section, we will discretize our continuous problems with a projection-based stabilized finite element method. We let $Y^h \subset Y$ and $U^h \subset U$ be the finite element spaces with a conforming triangulation τ^H of Ω. For $H \geq h$, let τ^h be a refinement of τ^H. Let L^H be a vector-valued finite element subspace of $L^2(\Omega)$. We assume that finite element spaces have the following properties. We consider Y^h to be the space of continuous piecewise polynomials of degree r. We also make the standard assumptions that the space Y^h satisfies the following approximation properties for a given integer $1 \leq s \leq r$:

$$\inf_{y^h \in Y^h} \left\{ \|(y - y^h)\| + h\|\nabla(y - y^h)\| \right\} \leq Ch^{s+1}(\|y\|_{s+1}) \tag{3.1}$$

for $(y \in (Y \cap H^{s+1}(\Omega))$. We also assume that the control variable u satisfies

$$\|u - \tilde{u}\| \leq Ch^{s+1}\|u\|_{s+1} \quad for\ u \in U \cap H^{s+1}(\Omega), \tag{3.2}$$

where $\tilde{u}$ is the L^2 projection from U to U^h. We also use the fact that L^2 orthogonal projections of L^H satisfy

$$\|G - P_H G\| \leq CH^s|G|_s, \quad 1 \leq s \leq r \tag{3.3}$$

for $G \in (L^2(\Omega) \cap H^s(\Omega))$. We also use the well-known property of these operators that

$$\|I - P_H\| \leq 1. \tag{3.4}$$

Through our scheme, we add the global stabilization first and then we subtract its effect onto large scales only. In this way, stabilization acts only on the smallest resolved scales. To do this, additional diffusion acts on all discrete scales and its effects are subtracted from scales resolvable on τ^H. We follow similar steps as done in [10] and [15] to obtain the discrete form of PDE's and we obtain the optimal control problem (1.1)–(1.2) as follows: Find $y^h \in Y^h$,$g^H \in L^H$ and $u^h \in U^h$ such that

$$\min\ J(y^h, u^h) = \frac{1}{2}\left\|y^h - y_d\right\|^2 + \frac{\alpha}{2}\left\|u^h\right\|^2 \tag{3.5}$$

subject to

$$\begin{aligned} &\epsilon(\nabla y^h, \nabla v^h) + (\sigma(\nabla y^h - g^H), \nabla v^h) + (\vec{b}.\nabla y^h, v^h) + (ry^h, v^h) \\ &= (f, v^h) + (u^h, v^h),\ \forall v^h \in Y^h, &(3.6)\\ &(g^H - \nabla y^h, l^H) = 0,\ \forall l^h \in L^H. &(3.7) \end{aligned}$$

Here σ stands for a nonnegative user selected stabilization parameter depending on the mesh width h. These parameters can be thought of as an additional viscosity in the coarse space.

Remark 3.1 *One should select the large-scale space L^H explicitly in multiscale formulations as introduced in this study. The discretization we study adds additional diffusion acting on all discrete scales and then antidiffuses on the scales resolvable on large ones. Since each flow structure contains large and small scales together, distinguishing these scales is of importance. Here our regular step size or mesh parameter h stands for the small scales and H denotes the large scales used for stabilization and projection issues used within. If one chooses zero subspace for L^H, then the standard Galerkin formulation is obtained. Here we take $L^H = \nabla Y^H$ choice of [15] to obtain the results in this paper. See [12] for other possible choices.*

Here we note that equation (3.7) implies that g^H is the L^2 orthogonal projection of ∇y^h. Denoting this projection with P_H, the properties of the projection operator lead us to a new variational formulation of the problem: Find $y^h \in Y^h$ and $u^h \in U^h$ such that

$$\min\ J(y^h, u^h) = \frac{1}{2}\left\|y^h - y_d\right\|^2 + \frac{\alpha}{2}\left\|u^h\right\|^2 \tag{3.8}$$

subject to

$$\begin{aligned} &\epsilon(\nabla y^h, \nabla v^h) + \sigma((I - P_H)\nabla y^h, (I - P_H)\nabla v^h) \\ &+(\vec{b}.\nabla y^h, v^h) + (ry^h, v^h) = (f, v^h) + (u^h, v^h),\ \forall v^h \in Y^h, &(3.9) \end{aligned}$$

where I stands for the identity operator.

As in the continuous case, the control problem (3.8)–(3.9) admits a unique solution (y^h, u^h) if and only if there is a unique adjoint state p^h satisfying the optimality conditions:

$$\epsilon(\nabla y^h,\nabla v^h)+\sigma((I-P_H)\nabla y^h,(I-P_H)\nabla v^h)+(\vec{b}.\nabla y^h,v^h)+(ry^h,v^h) \tag{3.10}$$
$$=(f,v^h)+(u^h,v^h),\ \forall v^h\in Y^h,$$
$$\epsilon(\nabla p^h,\nabla w^h)+\sigma((I-P_H)\nabla p^h,(I-P_H)\nabla w^h)-(\nabla.(\vec{b}p^h),w^h)+(rp^h,w^h) \tag{3.11}$$
$$=(y^h-y_d,w^h),\ \forall w^h\in Y^h,$$
$$(\alpha u^h+p^h,w^h-u^h)\geq 0,\ \forall w^h\in Y^h. \tag{3.12}$$

The new variational form of the adjoint equation seen in (3.11) is obtained exactly the same as done for (3.10) previously. In the next lemma, we state stability results for the state and adjoint state equations.

Lemma 3.1 *Under the coercivity assumption* $r-\frac{1}{2}\nabla.\vec{b}\geq\beta>0$*, the discrete optimal control problem (3.10)–(3.11) is stable in the sense that:*

$$\epsilon\left\|\nabla y^h\right\|^2+\sigma\left\|(I-P_H)\nabla y^h\right\|^2+\frac{\beta}{2}\left\|y^h\right\|^2\leq\frac{2}{\beta}(\|f\|^2+\left\|u^h\right\|^2), \tag{3.13}$$

and

$$\epsilon\left\|\nabla p^h\right\|^2+\sigma\left\|(I-P_H)\nabla p^h\right\|^2+\frac{\beta}{2}\left\|p^h\right\|^2\leq\frac{1}{\beta}\left\|y^h-y_d\right\|^2. \tag{3.14}$$

Proof To prove the state part, we put $v^h=y^h$ in (3.10) to get

$$\begin{aligned}\epsilon(\nabla y^h,\nabla y^h)+\sigma((I-P_H)\nabla y^h,(I-P_H)\nabla y^h) &+ (\vec{b}\cdot\nabla y^h+ry^h,y^h)\\ &= (f+u^h,y^h).\end{aligned}$$

Making use of the integration by parts in the last terms on the left-hand side, we obtain

$$\begin{aligned}\epsilon(\nabla y^h,\nabla y^h)+\sigma((I-P_H)\nabla y^h,(I-P_H)\nabla y^h) &+ ((r-\frac{1}{2}\nabla.\vec{b})y^h,y^h)\\ &\leq (f+u^h,y^h).\end{aligned}$$

The Cauchy–Schwartz and Young's inequalities give the desired result along with the coercivity assumption for the state part. For the adjoint equation, we set $w^h=p^h$ in (3.11) to obtain

$$\begin{aligned}\epsilon\left\|\nabla p^h\right\|^2+\sigma\left\|(I-P_H)\nabla p^h\right\|^2 &- (\nabla.(\vec{b}p^h),p^h)+(rp^h,p^h)\\ &= (y^h-y_d,p^h).\end{aligned} \tag{3.15}$$

Now, making use of integration by parts for the term $(\nabla.(\vec{b}p^h),p^h)$ in the last equation, one has

$$-(\nabla.(\vec{b}p^h),p^h)=(\vec{b}p^h,\nabla p^h)=\frac{1}{2}\int\vec{b}\cdot\nabla(p^h)^2dx_\Omega=-\frac{1}{2}((\nabla.\vec{b})p^h,p^h).$$

Thus, we have

$$-(\nabla.(\vec{b}p^h), p^h) + (rp^h, p^h) = ((r - \frac{1}{2}\nabla.\vec{b})p^h, p^h). \tag{3.16}$$

Finally, putting the above bound into the left-hand side of (3.15), and making use of Cauchy–Schwartz and Young's inequalities on the right-hand side would yield the desired result as in the state case. □

4. Error estimates for the optimal control problem

In this section, we will derive the error estimate for the state, adjoint state, and control variables.

We consider the solution operator $S : U \mapsto H_0^1(\Omega) \cap H^2(\Omega)$. Then we define the reduced cost function:

$$J(y, u) = J(S(u), u) := j(u), \tag{4.1}$$

where $S(u)$ solves

$$\begin{aligned} -\epsilon\Delta y(u) + \vec{b} \cdot \nabla y(u) + ry &= u \text{ in } \Omega, \\ y(u) &= 0 \text{ on } \Sigma. \end{aligned} \tag{4.2}$$

The optimization techniques by using the Lagrange approach give the reduced gradient as:

$$j'(u)(\tilde{u} - u) = (p(u) + \alpha u, \tilde{u} - u), \quad \forall \tilde{u} \in U, \tag{4.3}$$

with $p(u)$ solving the following system:

$$\begin{aligned} -\epsilon\Delta p(u) - \nabla.(\vec{b}p(u)) + rp &= y(u) - y_d \quad \text{in } \Omega, \\ p(u) &= 0 \qquad\qquad \text{on } \Sigma. \end{aligned} \tag{4.4}$$

Since our cost function J and the operator $e(y, u)$ are twice-differentiable, then we can use the second-order sufficient optimality condition to get the positive definiteness of the reduced hessian [2]

$$j''(u)(\delta u, \delta u) \geq \alpha \left\|\delta u\right\|^2 \quad \forall \delta u \in U. \tag{4.5}$$

Similar to the continuous case, we can define the discrete solution operator S^h such that $S^h(u) = y^h(u)$. Then there holds

$$j_h'(u)(\tilde{u} - u) = (p^h(u) + \alpha u, \tilde{u} - u), \quad \forall \tilde{u} \in U, \tag{4.6}$$

and

$$j_h''(u)(\delta u, \delta u) \geq \alpha \left\|\delta u\right\|_{L^2(\Omega)}^2 \quad \forall \delta u \in U, \tag{4.7}$$

where $(y^h(u), p^h(u))$ solves

$$\begin{aligned} &\epsilon(\nabla y^h(u), \nabla v^h) + \sigma_1((I - P_H)\nabla y^h(u), (I - P_H)\nabla v^h) + (\vec{b} \cdot \nabla y^h(u) + ry^h(u), v^h) \\ &\qquad = (f + u, v^h), \ \forall v^h \in Y^h, \end{aligned} \tag{4.8}$$

$$\begin{aligned} &\epsilon(\nabla p^h(u), \nabla w^h) + \sigma_2((I - P_H)\nabla p^h(u), (I - P_H)\nabla w^h) + (-\nabla.(\vec{b}p^h(u)) + rp^h(u), w^h) \\ &\qquad = (y^h(u) - y_d, w^h), \ \forall w^h \in U^h. \end{aligned} \tag{4.9}$$

Lemma 4.1 *Let y and $y^h(u)$ be solutions of (2.2) and (4.8), respectively. Then we have*

$$\epsilon\|\nabla(y(u)-y^h(u))\|^2+\sigma\|((I-P_H)\nabla(y(u)-y^h(u))\|^2+\beta\left\|(y(u)-y^h(u))\right\|^2\leq$$
$$C\{(\epsilon+\beta^{-1})\|\nabla(y(u)-\tilde{y})\|^2+\sigma\|(I-P_H)\nabla(y(u)-\tilde{y})\|^2+\sigma\|(I-P_H)\nabla y(u)\|^2\},$$

where $\tilde{y}$ is the best approximation of y in Y^h.

Proof We begin the analysis by constructing an error equation through subtracting (4.8) from (2.2) via the same test function v^h. Thus we get

$$\begin{aligned}&\epsilon(\nabla(y(u)-y^h(u)),\nabla v^h)+\sigma((I-P_H)\nabla(y(u)-y^h(u)),(I-P_H)\nabla v^h)\\&\quad+(\vec{b}\cdot\nabla(y(u)-y^h(u)),v^h)+(r(y(u)-y^h(u)),v^h)\\&\quad=\sigma((I-P_H)\nabla y(u),(I-P_H)\nabla v^h),\ \forall v^h\in Y^h.\end{aligned}\tag{4.10}$$

Now we split the error term $y(u)-y^h(u)$ as $y(u)-y^h(u)=y(u)-\tilde{y}-(y^h(u)-\tilde{y})=\eta-\phi^h$, where $\tilde{y}$ is the best approximation of $y(u)$ in Y^h. Hence we modify our error equation as:

$$\begin{aligned}&\epsilon(\nabla\phi^h,\nabla v^h)+\sigma((I-P_H)\nabla\phi^h,(I-P_H)\nabla v^h)+(\vec{b}\cdot\nabla\phi^h+r\phi^h,v^h)\\&=\epsilon(\nabla\eta,\nabla v^h)+\sigma((I-P_H)\nabla\eta,(I-P_H)\nabla v^h)+(\vec{b}\cdot\nabla\eta+r\eta,v^h)\\&-\sigma((I-P_H)\nabla y,(I-P_H)\nabla v^h),\ \forall v^h\in Y^h.\end{aligned}\tag{4.11}$$

We now let $v^h=\phi^h$ and rearrange (4.11) to get

$$\begin{aligned}&\epsilon\|\nabla\phi^h\|^2+\sigma\|((I-P_H)\nabla\phi^h\|^2+((r-\frac{1}{2}\nabla.\vec{b})\phi^h,\phi^h)\\&=\epsilon(\nabla\eta,\nabla\phi^h)+\sigma((I-P_H)\nabla\eta,(I-P_H)\nabla\phi^h)+(\vec{b}\cdot\nabla\eta+r\eta,\phi^h)\\&-\sigma((I-P_H)\nabla y,(I-P_H)\nabla\phi^h).\end{aligned}\tag{4.12}$$

Making use of the coercivity assumption on the left-hand side of (4.12), we have

$$\begin{aligned}&\epsilon\|\nabla\phi^h\|^2+\sigma\|((I-P_H)\nabla\phi^h\|^2+\beta(\phi^h,\phi^h)\\&=\epsilon(\nabla\eta,\nabla\phi^h)+\sigma((I-P_H)\nabla\eta,(I-P_H)\nabla\phi^h)+(\vec{b}\cdot\nabla\eta+r\eta,\phi^h)\\&-\sigma((I-P_H)\nabla y(u),(I-P_H)\nabla\phi^h).\end{aligned}\tag{4.13}$$

We will take the absolute value of the right-hand side of (4.13) and treat each term separately. The first term is bounded as:

$$|\epsilon(\nabla\eta,\nabla\phi^h)|\leq\frac{\epsilon}{2}\|\nabla\eta\|^2+\frac{\epsilon}{2}\|\nabla\phi^h\|^2,$$

which is obtained by applying Cauchy–Schwartz and Young's inequalities. For the next term, we have

$$|\sigma((I-P_H)\nabla\eta,(I-P_H)\nabla\phi^h)|\leq\sigma\|(I-P_H)\nabla\eta\|^2+\frac{\sigma}{4}\|(I-P_H)\nabla\phi^h\|^2.$$

For the inconsistency error term, we obtain

$$|\sigma((I-P_H)\nabla y(u),(I-P_H)\nabla\phi^h)|\leq\sigma\|(I-P_H)\nabla y(u)\|^2+\frac{\sigma}{4}\|(I-P_H)\nabla\phi^h\|^2.$$

Thus, we are left with only the $(\vec{b}\cdot\nabla\eta + r\eta, \phi^h)$. Making use of integration by parts, boundedness of $\nabla.\vec{b}$ and Poincare Friedrich's inequality will give:

$$|(\vec{b}\cdot\nabla\eta,\phi^h)| = |-((\nabla.\vec{b})\eta,\phi^h)| \leq C\beta^{-1}\|\nabla\eta\|^2 + \frac{\beta}{4}\|\phi^h\|^2$$

and

$$|(r\eta,\phi^h)| \leq \|r\|_\infty \|\eta\| \left\|\phi^h\right\| \leq C\beta^{-1}\|\nabla\eta\|^2 + \frac{\beta}{4}\|\phi^h\|^2.$$

Combining all the bounds and collecting the ϕ^h terms on the left-hand side of (4.13) we have

$$\begin{aligned}&\frac{\epsilon}{2}\|\nabla\phi^h\|^2 + \frac{\sigma}{2}\|((I-P_H)\nabla\phi^h\|^2 + \frac{\beta}{2}\left\|\phi^h\right\|^2 \leq \\ &C\{(\epsilon+\beta^{-1})\|\nabla(\eta)\|^2 + \sigma\|(I-P_H)\nabla(\eta)\|^2 + \sigma\|(I-P_H)\nabla y(u)\|^2\}.\end{aligned} \tag{4.14}$$

Multiplying both sides of (4.14) by 2 and application of the triangle inequality give the desired result now. □

We state a similar bound for the adjoint equation in the next lemma.

Lemma 4.2 *Let p and $p^h(u)$ be solutions of (2.5) and (4.9), respectively. Then we have*

$$\begin{aligned}&\epsilon\|\nabla(p(u)-p^h(u))\|^2 + \sigma\|((I-P_H)\nabla(p(u)-p^h(u))\|^2 + \beta\left\|(p(u)-p^h(u))\right\|^2 \leq \\ &C\{(\epsilon+\beta^{-1})(\|\nabla(p(u)-\tilde{p})\|^2 + \beta^{-2}\|\nabla(y(u)-\tilde{y})\|^2) + \sigma\|(I-P_H)\nabla(p(u)-\tilde{p})\|^2 \\ &+\beta^{-2}\sigma\|(I-P_H)\nabla(y(u)-\tilde{y})\|^2 + \sigma\|(I-P_H)\nabla p(u)\|^2 + \beta^{-2}\sigma\|(I-P_H)\nabla y(u)\|^2\},\end{aligned}$$

where $\tilde{p}$ is the best approximation of $p(u)$ in Y^h and $\tilde{y}$ is the best approximation of $y(u)$ in Y^h.

Proof As in other error proofs, we begin by subtracting (4.9) from (2.5) and split the error $p(u)-p^h(u) = p(u)-\tilde{p}-(p^h(u)-\tilde{p}) = \eta-\phi^h$ with $\tilde{p}$ being the best approximation of $p(u)$ in Y^h. Now, with the test function choice of ϕ^h, our error equation takes the form

$$\begin{aligned}&\epsilon\|\nabla\phi^h\|^2 + \sigma\|(I-P_H)\nabla\phi^h\|^2 - (\nabla.(\vec{b}\phi^h),\phi^h) + (r\phi^h,\phi^h) \\ &= \epsilon(\nabla\eta,\nabla\phi^h) + \sigma((I-P_H)\nabla\eta,(I-P_H)\nabla\phi^h) - (\nabla.(\vec{b}\eta),\phi^h) + (r\eta,\phi^h) \\ &-\sigma((I-P_H)\nabla p(u),(I-P_H)\nabla\phi^h) - (y(u)-y^h(u),\phi^h).\end{aligned}$$

We can easily obtain the bounds of the terms on the right-hand side of (4.15). To list them:

$$\begin{aligned}|\epsilon(\nabla\eta,\nabla\phi^h) + \sigma((I-P_H)\nabla\eta,(I-P_H)\nabla\phi^h)| &\leq \epsilon\|\nabla\eta\|^2 + \frac{\epsilon}{4}\|\nabla\phi^h\|^2 + \sigma\|(I-P_H)\nabla\eta\|^2 \\ &+ \frac{\sigma}{4}\|(I-P_H)\nabla\phi^h\|^2\end{aligned}$$

$$|\sigma((I-P_H)\nabla p(u),(I-P_H)\nabla\phi^h)| \leq \sigma\|(I-P_H)\nabla p(u)\|^2 + \frac{\sigma}{4}\|(I-P_H)\nabla\phi^h\|^2.$$

$$|(y(u)-y^h(u),\phi^h)| \leq C\beta^{-1}\|y(u)-y^h(u)\|^2 + \frac{\beta}{6}\|\phi^h\|^2$$

The bound for the term $\beta^{-1}\|y-y^h(u)\|^2$ is obtained by using the previous lemma. To proceed, we have

$$|-(\nabla.(\vec{b}\eta),\phi^h)| \leq C\{\beta^{-1}\left\|\nabla\eta\right\|^2 + \frac{\beta}{6}\|\phi^h\|^2\}$$

due to the boundedness of the terms $\nabla.\vec{b}$. Similar to the previous proof, we have

$$|(r\eta,\phi^h)| \leq \|r\|_\infty \|\eta\| \left\|\phi^h\right\| \leq C\beta^{-1}\|\nabla\eta\|^2 + \frac{\beta}{6}\|\phi^h\|^2.$$

Rearranging the left-hand side of (4.15) by using the coercivity assumption, combining all obtained bounds and the triangle inequality yield the desired result. □

In the following two lemmas, we obtain estimates between the discrete and auxiliary state and adjoint variables.

Lemma 4.3 *Let y^h and $y^h(u)$ be solutions of (3.10) and (4.8), respectively. Then there holds*

$$\epsilon\left\|\nabla(y^h-y^h(u))\right\|^2 + \sigma\left\|(I-P_H)\nabla(y^h-y^h(u))\right\|^2 \leq \frac{1}{2\beta}\left\|u-u^h\right\|^2 \tag{4.15}$$

Proof We subtract (4.8) from (3.10) to get

$$\begin{aligned}&\epsilon(\nabla(y^h-y^h(u)),\nabla v^h) + \sigma((I-P_H)\nabla(y^h-y^h(u)),(I-P_H)\nabla v^h)\\&+(\vec{b}\cdot\nabla(y^h-y^h(u)),v^h) + (r(y^h-y^h(u)),v^h) = (u^h-u,v^h),\ \forall v^h\in Y^h.\end{aligned}$$

Now we choose $v^h = y^h - y^h(u)$ and use integration by parts and the coercivity assumption to get

$$\begin{aligned}&\epsilon\left\|\nabla(y^h-y^h(u))\right\|^2 + \sigma\left\|(I-P_H)\nabla(y^h-y^h(u))\right\|^2 + \beta\left\|y^h-y^h(u)\right\|^2\\&\leq (u^h-u, y^h-y^h(u)).\end{aligned}$$

Then we use Young's inequality for the right-hand-side term:

$$\begin{aligned}&\epsilon\left\|\nabla(y^h-y^h(u))\right\|^2 + \sigma\left\|(I-P_H)\nabla(y^h-y^h(u))\right\|^2 + \frac{\beta}{2}\left\|y^h-y^h(u)\right\|^2\\&\leq \frac{1}{2\beta}\left\|u^h-u\right\|^2.\end{aligned}$$

□

Lemma 4.4 *Let p^h and $p^h(u)$ be solutions of (3.11) and (4.9), respectively. Then there holds*

$$\epsilon\left\|\nabla(p^h-p^h(u))\right\|^2 + \sigma\left\|(I-P_H)\nabla(p^h-p^h(u))\right\|^2 \leq \frac{1}{4\beta^2}\left\|u-u^h\right\|^2. \tag{4.16}$$

Proof We follow the same procedure as in Lemma 4.3. Firstly, we subtract (4.9) from (3.11) to get

$$\begin{aligned}\epsilon(\nabla(p^h-p^h(u)),\nabla w^h) + \sigma((I-P_H)\nabla(p^h-p^h(u)),(I-P_H)\nabla w^h)&\\-(\nabla.(\vec{b}(p^h-p^h(u))),w^h) + (r(p^h-p^h(u)),w^h)&\\= (y^h-y^h(u),w^h),\ \forall w^h\in Y^h,&\end{aligned}$$

Now we choose $w^h = p^h - p^h(u)$ and use integration by parts, equation (3.16), and the coercivity assumption to get

$$\begin{aligned}&\epsilon \left\|\nabla(p^h - p^h(u))\right\|^2 + \sigma \left\|(I - P_H)\nabla(p^h - p^h(u))\right\|^2 + \frac{\beta}{2}\left\|p^h - p^h(u)\right\|^2\\ &\leq (y^h - y^h(u), p^h - p^h(u)).\end{aligned}$$

Then Young's inequality implies that

$$\begin{aligned}&\epsilon \left\|\nabla(p^h - p^h(u))\right\|^2 + \sigma \left\|(I - P_H)\nabla(p^h - p^h(u))\right\|^2 + \beta\left\|p^h - p^h(u)\right\|^2\\ &\leq \frac{1}{2\beta}\left\|y^h - y^h(u)\right\|^2.\end{aligned}$$

Now we make use of Lemma 4.3 to get the desired result. □

We need the following result concerning the derivative of the reduced cost function (4.1).

Lemma 4.5 *The first derivative of the reduced cost function for the continuous and the discrete cases satisfies*

$$\|j'(u)(\delta) - j'_h(u)(\delta)\| \leq \left\|p(u) - p^h(u)\right\| \|\delta\| \quad \forall u, \delta \in U. \tag{4.17}$$

Proof The result is obtained by using Eqs. (4.3) and (4.6) directly. □

The following lemma gives the error estimate for the control variable u.

Lemma 4.6 *Let (u, y) and (u^h, y^h) be solutions to (1.1–1.2) and (3.8–3.9), respectively. Then we have*

$$\left\|u - u^h\right\| \leq \|u - \tilde{u}\| + \frac{1}{\alpha}\left\|p(u) - p^h(u)\right\|, \quad \tilde{u} \in U. \tag{4.18}$$

Proof Let u and u^h be solutions to continuous and discrete control problems, respectively. We choose an arbitrary $\tilde{u}$ from U. We write

$$u - u^h = u - \tilde{u} + \tilde{u} - u^h. \tag{4.19}$$

From Eq. (4.7), we have

$$\begin{aligned}\alpha\left\|\tilde{u} - u^h\right\|^2 &\leq j''_h(\tilde{u})(\tilde{u} - u^h, \tilde{u} - u^h)\\ &= j'_h(u)(\tilde{u} - u^h) - j'_h(u^h)(\tilde{u} - u^h).\end{aligned}$$

Since u and u^h are optimal solutions, then

$$j'_h(u^h)(\tilde{u} - u^h) = 0 = j'(u)(\tilde{u} - u^h).$$

Then

$$\begin{aligned}\alpha\left\|\tilde{u} - u^h\right\|^2 &\leq j'_h(u)(\tilde{u} - u^h) - j'(u)(\tilde{u} - u^h)\\ &\leq \left\|p(u) - p^h(u)\right\| \left\|\tilde{u} - u^h\right\|.\end{aligned} \tag{4.20}$$

Finally,

$$\left\|\tilde{u} - u^h\right\| \quad \leq \quad \frac{1}{\alpha}\left\|p(u) - p^h(u)\right\|. \tag{4.21}$$

□

Corollary 4.1 *The error in state variable y satisfies:*

$$\begin{aligned}
&\epsilon\|\nabla(y - y^h)\|^2 + \sigma\|(I - P_H)\nabla(y - y^h)\|^2 \leq \\
&C\epsilon \inf_{\tilde{y}\in Y}\{\|\nabla(y - \tilde{y})\|^2 + \sigma\|(I - P_H)\nabla(y - \tilde{y})\|^2 + \sigma\|(I - P_H)\nabla y\|^2 \\
&+ \left\|u - u^h\right\|^2\}.
\end{aligned} \tag{4.22}$$

Proof The corollary is the combination of Lemma (4.1) and Lemma (4.3). □

Corollary 4.2 *The error in adjoint state variable p satisfies:*

$$\begin{aligned}
&\epsilon\|\nabla(p - p^h)\|^2 + \sigma\|(I - P_H)\nabla(p - p^h)\|^2 \leq \\
&C\epsilon \inf_{\tilde{y},\tilde{p}\in Y}\{\|\nabla(p - \tilde{p})\|^2 + \sigma\|(I - P_H)\nabla(p - \tilde{p})\|^2 \\
&+\|\nabla(y - \tilde{y})\|^2 + \sigma\|(I - P_H)\nabla(y - \tilde{y})\|^2 + \sigma\|(I - P_H)\nabla y\|^2 \\
&+\sigma\|(I - P_H)\nabla p\|^2 + \left\|u - u^h\right\|^2\}.
\end{aligned} \tag{4.23}$$

Proof The proof is just a combination of the results of Lemma (4.2) and Lemma (4.4). □

Remark 4.1 *In order to define the error orders in terms of mesh width h, we will define approximation spaces. We make a choice for stabilization parameter σ and construct a relation between σ, H, and h. Here we are given the fine scale mesh h and by equilibrating the orders of convergence, appropriate values for the mesh scale H and parameter σ are chosen. That is, the error is optimal for $\sigma H^{2s} = h^{2s}$. For instance, let us consider the case for $s = 1$ and use finite element pairs, which are given below explicitly along with the choice of $L^H = \nabla Y^H$:*

$$\begin{aligned}
Y^h &= \{y^h \in Y : y^h|_\Delta \in P_1(\Delta), \forall\Delta \in \tau^h\} \quad \textit{and} \\
U^h &= \{y^h \in U : y^h|_\Delta \in P_1(\Delta), \forall\Delta \in \tau^h\}.
\end{aligned}$$

For details of these choices, we refer to [15].

We are now in a position to state approximation results. We give corollaries for each variable. We first assume that y, u, p are sufficiently smooth, before stating the approximation results.

Corollary 4.3 *The control variable u satisfies*

$$\|u - u^h\| \cong \mathcal{O}(h).$$

Proof Making use of approximation assumptions (3.1),(3.2), and (3.3) in Lemma 4.6 and considering Remark 4.1 along with property (3.4), we get

$$\|u-u^h\| \leq C(u)h^2 + C(\alpha^{-1},y,p)\sqrt{\epsilon(h^2+\sigma h^2+\sigma H^2)}. \tag{4.24}$$

According to Remark 4.1, we can choose $(\sigma,H)=(h,h^{1/2})$. Putting these selections into (4.24) completes the proof. □

Corollary 4.4 *The adjoint state variable p satisfies*

$$\epsilon\|\nabla(p-p^h)\|^2+\sigma\|(I-P_H)\nabla((p-p^h))\|^2 \cong \mathcal{O}(h^2).$$

Proof The proof is similar to the previous case, which is stated for u. □

Corollary 4.5 *The state variable y satisfies*

$$\epsilon\|\nabla(y-y^h)\|^2+\sigma\|(I-P_H)\nabla(y-y^h)\|^2 \cong \mathcal{O}(h^2).$$

Proof The proof is similar to the previous cases, which are stated for u and p. □

As seen through these corollaries, the error in each case is optimal.

Remark 4.2 *We note that discretize-then-optimize and optimize-then-discretize approaches commute. If we apply the discretize-then-optimize approach we get the commutativity of (OD) and (DO). Indeed, we can define the discrete Lagrangian as:*

$$\begin{aligned} L^h(y^h,u^h,p^h) &:= \frac{1}{2}(y^h-y_d,y^h-y_d) \ +\frac{\alpha}{2}(u^h,u^h)-\epsilon(\nabla y^h,\nabla p^h) \\ &-\sigma((I-P_H)\nabla y^h,(I-P_H)\nabla p^h)-(\vec{b}\cdot\nabla y^h+ry^h,p^h)+(f+u^h,p^h). \end{aligned} \tag{4.25}$$

We let $\nabla_{y^h}L^h=0$ and $\nabla_{u^h}L^h=0$ to get the first-order discrete optimality conditions. By standard theory and integration by parts, one can obtain the same discrete adjoint scheme as:

$$\begin{aligned} \epsilon(\nabla p^h,\nabla w^h)+\sigma((I-P_H)\nabla p^h,(I-P_H)\nabla w^h)-(\vec{b}\cdot\nabla p^h+(\nabla.\vec{b})p^h,w^h)+(rp^h,w^h) \\ =(y^h-y_d,w^h),\ \forall w^h\in Y^h. \end{aligned}$$

We note that the stabilization term does not break the commutativity of (OD) and (DO).

5. Pointwise control constraints

In this section, we consider some pointwise constraints on the control variable. If we impose some conditions on the control u as:

$$a\leq u\leq b,$$

then we say that a box constraint holds on the control variable. In this case, the control variable u is searched for in an admissible set $Q_{ad}\subseteq L^2(\Omega)$ such as

$$Q_{ad}:=\{u\in L^2(\Omega)|\, a\leq u\leq b \text{ a.e. in } \Omega\},$$

where a and b are real numbers. The inequality (2.4) is written as

$$(\alpha u + p, w - u) \geq 0 \quad \forall\, w \in Q_{ad}. \tag{5.1}$$

This inequality can be equivalently formulated [17] as

$$u = \Pi_{[a,b]}(-\frac{1}{\alpha}p), \tag{5.2}$$

where the projection Π is defined as

$$\Pi_{[a,b]}(f(x)) := \max(a, \min(b, f(x))). \tag{5.3}$$

This projection also satisfies the Lipschitz continuity property:

$$\left\| \Pi_{[a,b]}(\frac{1}{\alpha}p) - \Pi_{[a,b]}(\frac{1}{\alpha}\bar{p}) \right\| \leq \frac{1}{\alpha} \left\| p - \bar{p} \right\|. \tag{5.4}$$

Theorem 5.1 *Let (u, y) and (u^h, y^h) be solutions to (1.1–1.2) and (3.8–3.9), respectively. Then*

$$\left\| u - u^h \right\| \leq \frac{1}{\alpha} \left\| p - p^h \right\|. \tag{5.5}$$

Proof Since $u = \Pi_{[a,b]}(-\frac{1}{\alpha}p)$ and $u^h = \Pi_{[a,b]}(-\frac{1}{\alpha}p^h)$ then we use the Lipschitz continuity property of the projection Π to get

$$\left\| u - u^h \right\| = \left\| \Pi_{[a,b]}(-\frac{1}{\alpha}p) - \Pi_{[a,b]}(-\frac{1}{\alpha}p^h) \right\| \leq \frac{1}{\alpha} \left\| p - p^h \right\|.$$

□

Thus, by a similar argument as in the unconstraint case, the error in the optimal control is obtained.

6. Numerical applications

In this section, we perform some numerical tests showing the efficiency of the VMS method. We use a gradient descent-type algorithm to solve the optimization problem. We also proved in the numerical convergence test that the stabilization does not degenerate the order of the error. All computations are carried out with the finite element software package Freefem++ [8].

6.1. Numerical example for smaller ϵ

As a first numerical application, we study in the domain $(0,1)^2$ with a mesh resolution of 32×32. We choose the parameters as $\epsilon = 10^{-6}$, $\alpha = 0.001$, $\vec{b} = (1,0)^T$, and $r = 0$ [7]. We let $y_d = \sin(\pi x_1)\sin(\pi x_2)$ and $f = 0$. Furthermore, we choose the stabilization parameter σ and coarse mesh size H as explained in Remark 4.1.

In Figures 1 and 2, we compare the optimal control and optimal state solutions for both stabilized and unstabilized cases. One can easily see the efficiency of the stabilization through comparison of these figures for both control and state solutions. As the stabilized solutions are observed to be smooth and acceptable, unstabilized solutions blow up and oscillations are easily determined by a rough look. The effect of the stabilization on optimal control problem is trivial through this simple test case even for a relatively fine mesh of resolution 32×32. We do not include the results for coarser meshes since the unstabilized solutions are very hard to obtain.

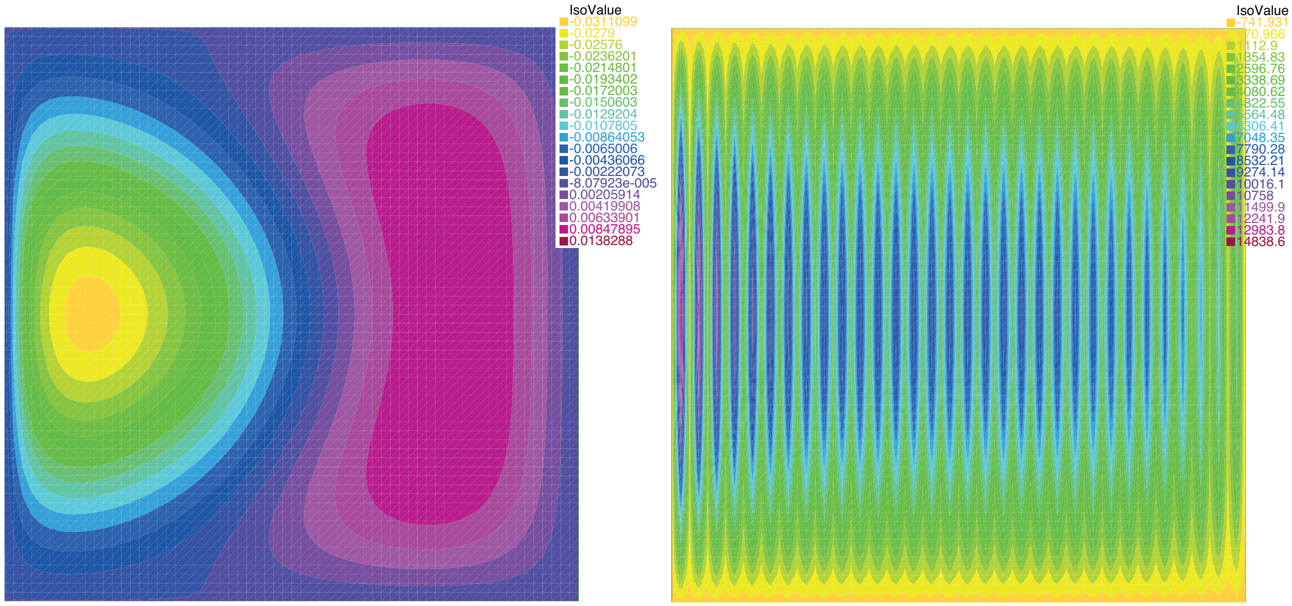

Figure 1. Comparison of optimal control solutions (first numerical test): stabilized (up) and unstabilized (down).

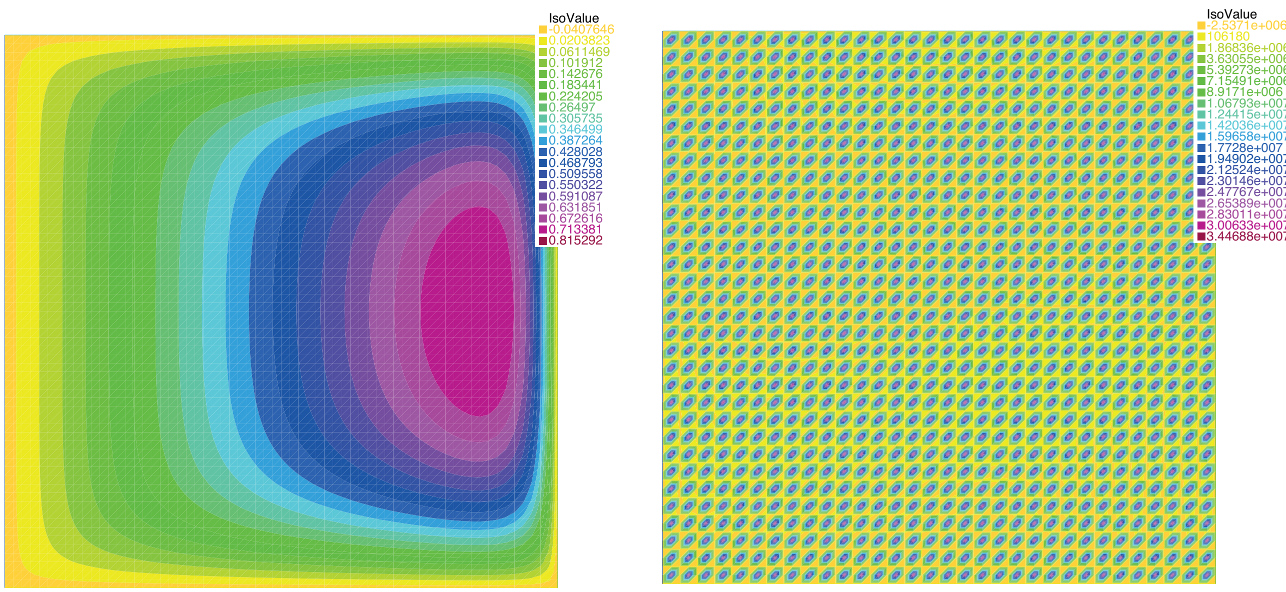

Figure 2. Comparison of optimal state solutions (first numerical test): stabilized(up) and unstabilized(down)

6.2. Solution with parabolic and exponential boundary layers

In this numerical example, we again study in the domain $(0,1)^2$ this time with a mesh resolution of 64×64. We choose the parameters as $\epsilon = 10^{-8}$, $\alpha = 0.001$, $\vec{b} = (1,0)^T$ and $r = 0$ [14]. We let $y_d = \sin(\pi x_1)\sin(\pi x_2)$ and $f = 0$. Dirichlet boundary data are taken to be 0 in this case. Again, we choose the stabilization parameter σ and coarse mesh size H as explained in Remark 4.1. We first carry out the test with standard Galerkin FEM and also apply the method in order the compare the effect. The solution of this problem is known to have parabolic boundary layers at $y = 0$ and $y = 1$ and an exponential boundary layer at $x = 1$. The solutions of state and adjoint state variables for stabilized and stabilized methods can be seen in the figures below.

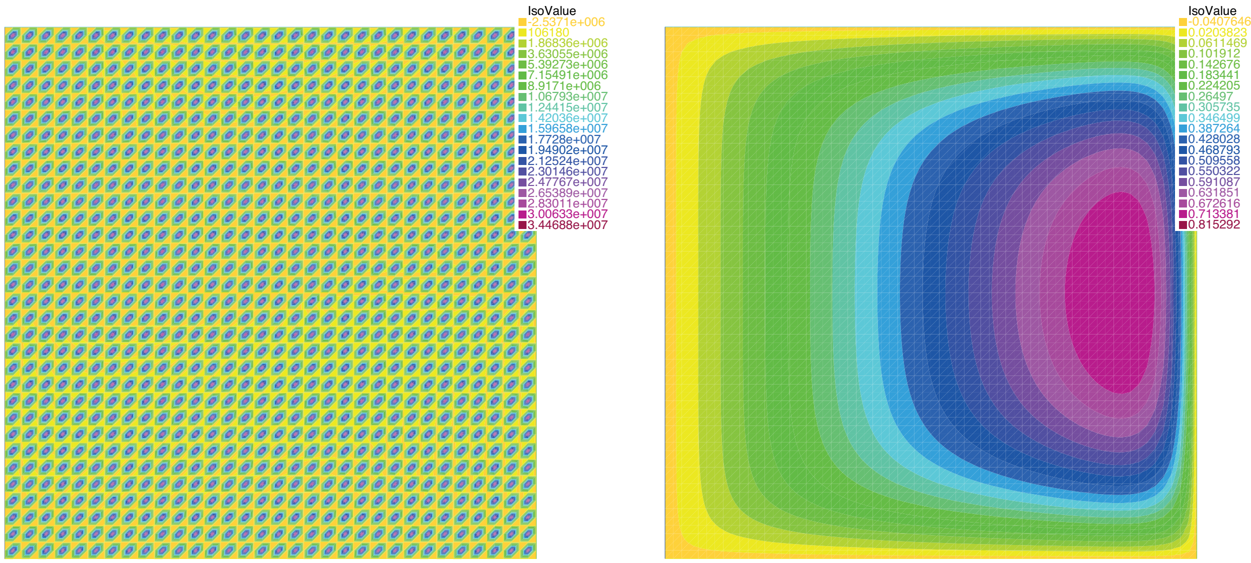

Figure 3. Comparison of optimal state solutions (second numerical test): unstabilized (up) and stabilized (down).

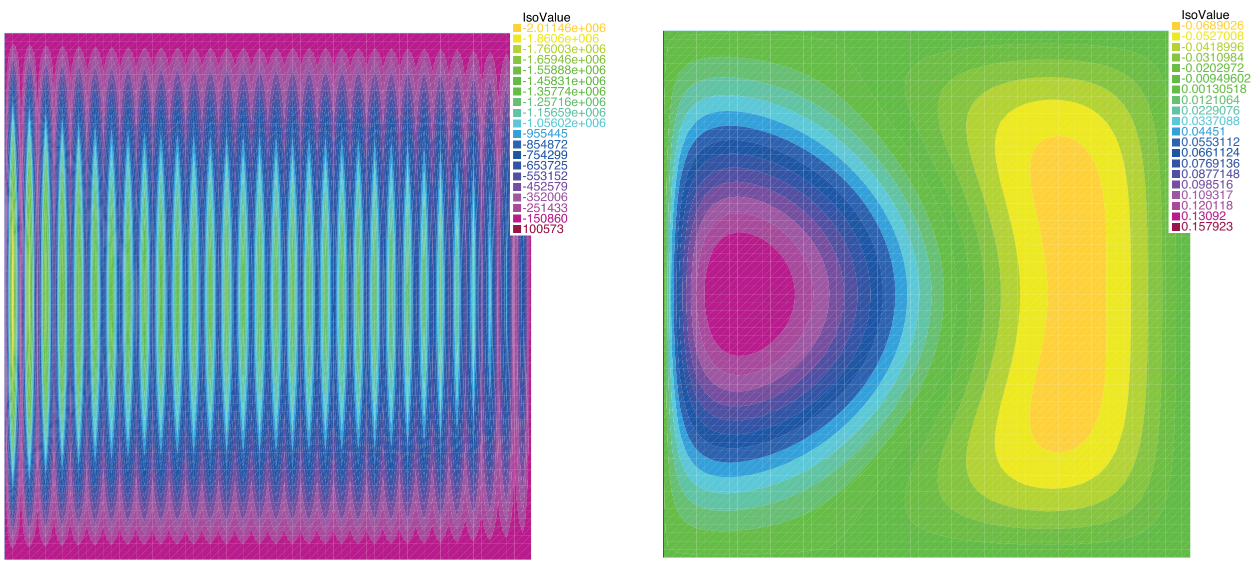

Figure 4. Comparison of optimal adjoint state solutions (second numerical test): unstabilized (up) and stabilized (down).

As could be directly deduced from the Figures 3 and 4, both state and adjoint state solutions blow up for the unstabilized case. When the stabilization applies, both variables could be captured in solutions. However, there are still strong oscillations at the parabolic layer and a more suitable stabilization technique might be applied in order to capture better solution data such as SOLD methods [14]. Still, this numerical example shows the promise of the stabilization we apply by revealing the layers that naturally should occur due to the problem's nature.

6.3. Numerical convergence study

In this subsection, we show that the theoretical orders of the errors are also obtained through a numerical simulation. We prove that the addition of the extra stabilization term in means of (VMS) does not degenerate order. We again study in the domain $(0,1)^2$ with different mesh resolutions. The problem parameters are chosen as $\vec{b} = (1,0)^T$ and $r = 1$. We study for two different ϵ and give the results of both. The following smooth solution is used to compute the orders of convergence:

$$y = \exp(-0.5)\sin(2\pi x_1)\sin(2\pi x_2)$$
$$p = \frac{1}{2}\exp(-0.5)\sin(2\pi x_1)\sin(2\pi x_2)$$
$$u = \frac{1}{\alpha}p$$

The corresponding source functions y_d and f are chosen so that the given solutions satisfy the equations (2.2)–(2.3).

Furthermore, the stabilization parameter σ and coarse mesh size H are chosen to satisfy the relation given in Remark 4.1. We pick $(\sigma, H) = (h, h^{1/2})$ here.

In order to show that stabilization term does not degenerate the order of convergence, we first take the parameter $\epsilon = 1$. In Table 1, we present the orders of convergence for both the control and state variables. The numerical results confirm the theoretical expectations given in the previous section. The expected order of convergence is 1 for the L^2 norm of the control variable and for the H^1 norm of the state variable. Moreover, the computed cost function value is given with respect to different mesh sizes.

Table 1. Errors and rates of convergence for $\epsilon = 1$ (second numerical test).

h	$\|u-u^h\|$	Rate	$\|y-y^h\|_1$	Rate	$\|J(u^h,y^h)\|$
2^{-4}	2.69e-2		6.57e-1		74.60
2^{-5}	1.31e-2	1.03	3.42e-1	0.94	74.28
2^{-6}	6.51e-3	1.01	1.70e-1	1.00	74.12
2^{-7}	3.21e-3	1.02	8.50e-2	1.03	74.04
2^{-8}	1.60e-3	1.00	4.20e-2	1.02	74.00

In Table 2, we present the orders of convergence for both the control and state variables with $\epsilon = 0.1$. The expected order 1 is obtained a little bit later and for finer meshes. This situation occurs since the physical oscillations in the solutions begins to take place as ϵ gets smaller and becomes more dominant.

Table 2. Errors and rates of convergence for $\epsilon = 0.1$ (second numerical test).

h	$\|u-u^h\|$	Rate	$\|y-y^h\|_1$	Rate	$\|J(u^h,y^h)\|$
2^{-4}	8.90e-2		1.96e-0		1.66
2^{-5}	5.90e-2	0.60	1.21e-0	0.69	1.57
2^{-6}	3.50e-2	0.75	7.50e-1	0.70	1.50
2^{-7}	2.00e-2	0.81	4.10e-1	0.87	1.44
2^{-8}	1.01e-2	1.00	2.00e-1	1.03	1.41

7. Conclusion and outlook

In this work, we have studied the variational multiscale method for the optimal control problems governed by convection diffusion reaction equations. We have obtained the stability results for both the state and adjoint state variables. We have derived a priori error bounds for each variable and proved that the error is optimal in each one. In the numerical examples, we have shown the efficiency of the stabilization in the solutions of control and state variables. In future studies, we will consider the optimal control of time dependent and nonlinear flow problems.

References

[1] Adams RA. Sobolev Spaces. New York, NY, USA: Academic Press, 1975.

[2] Becker R, Vexler B. Optimal control of the convection-diffusion equation using stabilized finite element methods. Numer Math 2007; 106: 349-367.

[3] Braack M, Burman E, John V, Lube G. Stabilized finite element methods for the generalized Oseen problem. Comput Meth Appl Mech Engrg 2007; 196: 853-866.

[4] Çıbık A, Kaya S. A projection-based stabilized finite element method for steady-state natural convection problem. J Math Anal Appl 2011; 381: 469-484.

[5] Çıbık A, Yılmaz F. Brezzi-Pitkaranta stabilization and a priori error analysis for the Stokes Control. An International Journal of Optimization and Control: Theories and Applications 2017; 7: 75-82.

[6] Codina R. Comparison of some finite element methods for solving the diffusion-convection-reaction equation. Comput Meth Appl Mech Engrg 1998; 156: 185-210.

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

[8] Hecht N. New development in FreeFem++. J Numer Math 2012; 3: 251-265.

[9] Heinkenschloss M, Leykekhman D. Local Error Estimates for SUPG Solutions of Advection-Dominated Elliptic Linear-Quadratic Optimal Control Problems. Houston, TX, USA: Rice University CAAM Technical Report TR08-30, 2008.

[10] Heitman N. Subgridscale stabilization of time-dependent convection dominated diffusive transport. J Math Anal Appl 2007; 331: 38-50.

[11] Hinze M, Yan N, Zhou Z. Variational discretization for optimal control governed by convection dominated diffusion equations. J Comput Math 2009; 27: 237-253.

[12] John V, Kaya S, Layton W. A two-level variational multiscale method for convection-diffusion equations. Comput Meth Appl Mech Engrg 2005; 195: 4594-4603.

[13] John V, Kaya S, Kindl A. Finite element error analysis for a projection-based variational multiscale method with nonlinear eddy viscosity. J Math Anal Appl 2008; 344: 627-641.

[14] John V, Knobloch P, Kindl A. On spurious at layers diminishing (SOLD) methods for convection-diffusion equations: Part I- A review. Comput Meth Appl Mech Engrg 2007; 196: 2197-2215.

[15] Layton WJ. A connection between subgrid scale eddy viscosity and mixed methods. Appl Math Comput 2002; 133: 147-157.

[16] Lions JL. Optimal control of systems governed by partial differential equations. New York, NY, USA: Springer Verlag, 1971.

[17] Malanowski K. Convergence of approximations vs. regularity of solutions for convex, control constrained optimal control problems. Appl Math 1982; 8: 69-95.

[18] Sun T. Discontinuous Galerkin finite element method with interior penalties for convection diffusion optimal control problem. Int J Numer Anal Model 2010; 7: 87-107.

[19] Zhou Z, Yan N. The local discontinuous Galerkin method for optimal control problem governed by convection-diffusion equations. International Journal of Numerical Analysis & Modeling 2010; 7: 681-699.