A mixed finite element method using a biorthogonal system for optimal control problems governed by a biharmonic equation

Bishnu P. Lamichhane[1] Neela Nataraj[2]
Deepanshu Verma[3]

(Received 30 January 2023; revised 10 October 2023)

Abstract

In this article, we consider an optimal control problem governed by a biharmonic equation with clamped boundary conditions. We use the Ciarlet–Raviart formulation combined with a biorthogonal system to obtain an efficient numerical scheme. We discuss the a priori error analysis and present results of the numerical experiments that validate the theoretical estimates.

DOI:10.21914/anziamj.v64.17961, Published 2023-11-01, as part of the Proceedings of the 20th Biennial Computational Techniques and Applications Conference. ISSN 1445-8810. (Print two pages per sheet of paper.)

Contents

1 Introduction

Let Ω be a bounded and convex domain in $\mathbb{R}^2$ and $\partial\Omega$ be the boundary of Ω. Consider the distributed optimal control problem governed by the biharmonic plate problem defined by

$$\inf_{u \in U_{ad}} \mathcal{K}(y, u) := \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{1}{2}\|\Delta y\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} \tag{1a}$$

$$\text{subject to} \quad \Delta^2 y = u + f \text{ in } \Omega \quad \text{and} \quad y|_{\partial\Omega} = \left.\frac{\partial y}{\partial n}\right|_{\partial\Omega} = 0\,. \tag{1b}$$

Here the unknowns y and u denote the displacement and control, respectively, y_d is the given observation for y, $\alpha > 0$ is a fixed regularization parameter, f is the given load function in $L^2(\Omega)$. For given $u_a, u_b \in \mathbb{R} \cup \{\pm\infty\}$ and $u_a \leqslant u_b$, a non-empty, convex, and bounded admissible set of controls is defined by

$$U_{ad} = \{u \in L^2(\Omega) : u_a \leqslant u(x) \leqslant u_b \text{ almost everywhere in } \Omega\} \subset L^2(\Omega)\,. \tag{2}$$

Biharmonic plate problems have many applications, for example, thin plates and beams [3], fluid flow [4] and phase separation of binary mixtures [10]. Optimal control problems governed by a biharmonic operator [9] are both interesting and challenging. Some approximation approaches for the optimal control problems governed by fourth order partial differential equations are the mixed finite element methods [2], interior penalty method [5], and collocation method [1].

In this article, we utilize a combination of the Ciarlet–Raviart mixed formulation [2] and an approach based on a biorthogonal system [6] to approximate the state and adjoint variables in the optimalility system. The biorthogonal system approach offers a significant advantage: it renders the cost of solving a biharmonic equation comparable with that of solving a Poisson equation.

Throughout the article, standard notions of Lebesgue and Sobolev spaces and their norms are employed [3]. For $s > 0$, the standard norms and semi-norms on $H^s(\Omega)$ space (resp. $W^{s,p}(\Omega)$) are denoted by $\|\cdot\|_s$ and $|\cdot|_s$ (resp. $\|\cdot\|_{s,p}$ and $|\cdot|_{s,p}$). The norm in the space $L^2(\Omega)$ is denoted by $\|\cdot\|$ and the standard inner product on $L^2(\Omega)$ space (resp. $H^s(\Omega)$) is denoted by $(\cdot\,,\cdot)$ (resp. $(\cdot\,,\cdot)_{s,\Omega}$). The spaces $H^1_0(\Omega)$ and $H^2_0(\Omega)$ have also standard definitions [3]. The notation $a \lesssim b$ implies $a \leqslant Cb$, where C is a generic constant that is independent of the mesh-size.

2 Mixed formulation

Following Ciarlet [3], we first recast the biharmonic problem (1b) as a minimisation problem

$$J(y) = \inf_{v \in H^2_0(\Omega)} J(v) =: \frac{1}{2}\int_\Omega |\Delta v|^2 \, dx - \int_\Omega (f+u)v \, dx \,. \tag{3}$$

Let $Q := H^1_0(\Omega)$, $M := L^2(\Omega)$ and $\boldsymbol{W} := Q \times M$. Let $\boldsymbol{W}$ be equipped with the inner product defined by $((y,\sigma),(v,\tau))_{\boldsymbol{W}} := (\nabla y, \nabla v) + (\sigma,\tau)$, and let $\|\cdot\|_{\boldsymbol{W}}$ denote the norm induced by the inner product. Let $S := H^1(\Omega)$.

Introduce a new unknown $\tau = \Delta v$ to recast (3) as the minimisation problem [3]

$$\begin{aligned} \mathcal{J}(y,\sigma) &= \inf_{(v,\tau)\in\mathcal{W}} \mathcal{J}(v,\tau) =: \frac{1}{2}\int_\Omega |\tau|^2\,dx - \int_\Omega (f+u)v\,dx \quad \text{with} \\ \mathcal{W} &= \left\{(v,\tau)\in \boldsymbol{W} : \int_\Omega (\nabla v\cdot\nabla q + \tau q)\,dx = 0 \text{ for all } q\in S\right\}, \end{aligned} \tag{4}$$

where the integral constraint in the above definition of $\mathcal{W}$ is obtained by multiplying $\tau = \Delta v$ by $q \in S$, and then performing an integration by parts. The saddle point formulation of this minimization problem seeks $((y,\sigma),\phi) \in \boldsymbol{W}\times S$ such that

$$\begin{aligned} a((y,\sigma),(v,\tau)) + b((v,\tau),\phi) &= \ell(v) \quad \text{for all } (v,\tau)\in\boldsymbol{W}, \\ b((y,\sigma),\psi) &= 0 \quad \text{for all } \psi\in S, \end{aligned} \tag{5}$$

where

$$\begin{aligned} a((y,\sigma),(v,\tau)) &= \int_\Omega \sigma\tau \quad dx, \quad b((y,\sigma),\psi) = \int_\Omega \sigma\psi\,dx + \int_\Omega \nabla y\cdot\nabla\psi\,dx, \\ \ell(v) &= \int_\Omega (f+u)v\,dx. \end{aligned}$$

The existence and uniqueness of the solution of mixed formulation (5) are established by Ciarlet [3] under regularity assumptions on y. Using (5), the optimal control problem (1) is rewritten as

$$\inf_{(y,\sigma,u)\in\boldsymbol{W}\times U_{ad}} \frac{1}{2}\|y-y_d\|^2 + \frac{1}{2}\|\sigma\|^2 + \frac{\alpha}{2}\|u\|^2, \tag{6}$$

subject to equation (5).

It is well-known [8, 9] that the convex control problem (6) has a unique solution $((\bar{y},\bar{\sigma}),\bar{\phi},\bar{u}) \in \boldsymbol{W}\times S\times U_{ad}$. The Karush–Kuhn–Tucker optimality conditions [9] lead to the problem of finding

$$((\bar{y},\bar{\sigma}),\bar{\phi},\bar{u},(\bar{p},\bar{\chi}),\bar{\eta}) \in \boldsymbol{X} := \boldsymbol{W}\times S\times U_{ad}\times\boldsymbol{W}\times S,$$

such that for all $(v,\tau) \in \mathbf{W}$, $w \in U_{ad}$, and $\psi \in S$,

$$a((\bar{y},\bar{\sigma}),(v,\tau)) + b((v,\tau),\bar{\phi}) = (f+\bar{u},v)\,, \tag{7a}$$
$$b((\bar{y},\bar{\sigma}),\psi) = 0\,, \tag{7b}$$
$$(\alpha\bar{u}+\bar{p}, w-\bar{u}) \geqslant 0\,, \tag{7c}$$
$$a((\bar{p},\bar{\chi}),(v,\tau)) + b((v,\tau),\bar{\eta}) = (\bar{y}-y_d,v) + (\bar{\sigma},\tau)\,, \tag{7d}$$
$$b((\bar{p},\bar{\chi}),\psi) = 0\,. \tag{7e}$$

Note that, for almost every $x \in \Omega$, the optimal control $\bar{u}$ in (7c) has the representation $\bar{u}(x) = P_{[u_a,u_b]}(-\bar{p}/\alpha)$ where the projection operator

$$P_{[a,b]}(f(x)) = \max(a, \min(b, f(x)))\,.$$

3 Finite element discretisation and a priori error analysis

Consider a quasi-uniform and shape-regular triangulation $\mathcal{T}_h$ of the polygonal domain Ω, where $\mathcal{T}_h$ consists of triangles or parallelograms. Let

$$V_h = \{v_h \in C^0(\bar{\Omega}) : v_h|_T \in \mathcal{P}_1(T)\,, T \in \mathcal{T}_h\}\,,$$

be the H^1-conforming linear finite element space, and $Q_h := V_h \cap H^1_0(\Omega)$. Let $\{\rho_1, \rho_2, \ldots, \rho_n\}$ be the finite element basis for the space V_h. Then we construct another piecewise polynomial space M_h, whose basis $\{\mu_1, \mu_2, \ldots, \mu_n\}$ is constructed in such a way that the basis functions of V_h and M_h satisfy the biorthogonality relation

$$\int_\Omega \mu_i \phi_j \, dx = c_i \delta_{ij}\,, \quad c_j \neq 0\,, \quad 1 \leqslant i, \quad j \leqslant n\,,$$

where $n := \dim M_h = \dim V_h$, and c_j is chosen as proportional to the area of support of ϕ_j [6]. Basis functions of M_h are also local and constructed on a reference element. Let $\mathbf{W}_h := Q_h \times M_h$. Working with a biorthogonal system for V_h and M_h, the matrix corresponding to the bilinear form $\int_\Omega \mu_h q_h \, dx$ for $\mu_h \in M_h$ and $q_h \in V_h$ is a diagonal matrix. Then the cost of solving the biharmonic equation is almost the same as solving a Poisson problem.

Define the space of piecewise constants

$$U_h = \{u_h \in U_{ad} : u_h|_T \in \mathcal{P}_0(T)\,, T \in \mathcal{T}_h\} \subset U_{ad}\,.$$

For all $(v_h, \tau_h) \in \boldsymbol{W}_h$, $w_h \in U_h$, and $\psi_h \in V_h$, the discrete optimal control problem corresponding to (7) seeks

$$((\bar{y}_h, \bar{\sigma}_h), \bar{\phi}_h, \bar{u}_h, (\bar{p}_h, \bar{\chi}_h), \bar{\eta}_h) \in \mathbf{X}_h := \boldsymbol{W}_h \times V_h \times U_h \times \boldsymbol{W}_h \times V_h\,,$$

such that

$$a((\bar{y}_h, \bar{\sigma}_h), (v_h, \tau_h)) + b((v_h, \tau_h), \bar{\phi}_h) = (f + \bar{u}_h, v_h)\,, \tag{8a}$$
$$b((\bar{y}_h, \bar{\sigma}_h), \psi_h) = 0\,, \tag{8b}$$
$$(\alpha \bar{u}_h + \bar{p}_h, w_h - \bar{u}_h) \geqslant 0\,, \tag{8c}$$
$$a((\bar{p}_h, \bar{\chi}_h), (v_h, \tau_h)) + b((v_h, \tau_h), \bar{\eta}_h) = (\bar{y}_h - y_d, v_h) + (\bar{\sigma}_h, \tau_h)\,, \tag{8d}$$
$$b((\bar{p}_h, \bar{\chi}_h), \psi_h) = 0\,. \tag{8e}$$

The bilinear form $a((\cdot\,,\cdot), (\cdot\,,\cdot))$ is coercive [6]. That is, there exists a positive constant $\alpha_0 > 0$ such that

$$a((v_h, \tau_h), (v_h, \tau_h)) \geqslant \alpha_0 (|v_h|_1^2 + \|\tau_h\|^2) \quad \text{for all } (v_h, \tau_h) \in \operatorname{Ker} B_h\,. \tag{9}$$

We now define a few projection operators for use in later analysis.

Definition 1 (Projections). The L^2 projections $\Pi_{0h} : L^2(\Omega) \to M_h$ and $\Pi_h : L^2(\Omega) \to U_h$ are defined by

$$(\Pi_{0h} v, \phi_h) = (v, \phi_h) \quad \text{for all } \phi_h \in V_h \text{ and } v \in L^2(\Omega)\,,$$
$$(\Pi_h v, u_h) = (v, u_h) \quad \text{for all } u_h \in U_h \text{ and } v \in L^2(\Omega).$$

The H^1 projection $\Pi_{1h} : H^1(\Omega) \to Q_h$ is defined by

$$(\Pi_{1h} v, q_h)_1 = (v, q_h)_1 \quad \text{for all } q_h \in Q_h \text{ and } v \in H^1(\Omega)\,.$$

The following lemma establishes an approximation property of Ritz projection [4, Chapter III], which is used to establish an a priori error estimate.

Lemma 2 (Ritz projection). *Let* $k \in \mathbb{N}$ *and* $r \in \mathbb{R}$ *be such that* $k \geqslant 1$ *and* $2 \leqslant r \leqslant \infty$ *. Let* $R_h : H_0^1(\Omega) \to Q_h$ *be the Ritz projection defined by*

$$\int_\Omega \nabla(R_h w - w) \cdot \nabla v_h \, dx = 0 \quad \text{for all } v_h \in Q_h .$$

Then, for all $w \in W^{k+1,r}(\Omega) \cap H_0^1(\Omega)$ *:*

$$\sup_{v_h \in V_h} \frac{\int_\Omega \nabla(w - R_h w) \cdot \nabla v_h \, dx}{\|v_h\|} \lesssim h^{k-\frac{1}{2}-\frac{1}{r}} \|w\|_{k+1,r} .$$

Let

$$\begin{aligned} \operatorname{Ker} B &= \{(v,\tau) \in \mathbf{W} : \; b((v,\tau),\phi)) = 0 \,, \phi \in S\} \quad \text{and} \\ \operatorname{Ker} B_h &= \{(v_h,\tau_h) \in \mathbf{W}_h : \; b((v_h,\tau_h),\phi_h)) = 0 \,, \phi_h \in V_h\} . \end{aligned}$$

For all $(v_h,\tau_h) \in \mathbf{W}_h$ and $\psi_h \in Q_h$, an auxiliary problem seeks

$$((y_h(\bar{u}),\sigma_h(\bar{u})),\phi_h(\bar{u}),\bar{u},(p_h(\bar{u}),\chi_h(\bar{u})),\eta_h(\bar{u})) \in \mathbf{X}_h ,$$

such that

$$a((y_h(\bar{u}),\sigma_h(\bar{u})),(v_h,\tau_h)) + b((v_h,\tau_h),\phi_h(\bar{u})) = (f+\bar{u},v_h) , \tag{10a}$$
$$b((y_h(\bar{u}),\sigma_h(\bar{u})),\psi_h) = 0 , \tag{10b}$$
$$a((p_h(\bar{u}),\chi_h(\bar{u})),(v_h,\tau_h)) + b((v_h,\tau_h),\eta_h(\bar{u})) = (\bar{y}-y_d,v_h) + (\bar{\sigma},\tau_h) , \tag{10c}$$
$$b((p_h(\bar{u}),\chi_h(\bar{u})),\psi_h) = 0 . \tag{10d}$$

We now prove the main result of the article that establishes an a priori error estimate for the mixed finite element approximation of the optimal control problem.

Theorem 3. *Let*

$$((\bar{y},\bar{\sigma}),\bar{\phi},\bar{u},(\bar{p},\bar{\chi}),\bar{\eta}) \in \mathbf{X} \quad \text{and} \quad ((\bar{y}_h,\bar{\sigma}_h),\bar{\phi}_h,\bar{u}_h,(\bar{p}_h,\bar{\chi}_h),\bar{\eta}_h) \in \mathbf{X}_h$$

be the solutions of (7) *and* (8)*, respectively. Under the extra regularity assumptions*

$$\bar{y},\bar{p} \in W^{2,\infty}(\Omega) \cap H_0^2(\Omega) \quad \text{and} \quad \bar{\phi},\bar{\eta} \in H^1(\Omega)\,,$$

it holds that

$$\|\bar{u}-\bar{u}_h\| \lesssim h(\|\bar{\phi}\|_1 + \|\bar{\eta}\|_1)\,, \tag{11a}$$

$$\|(\bar{y}-\bar{y}_h,\bar{\sigma}-\bar{\sigma}_h)\|_W \lesssim h\left(|\bar{y}|_2 + |\bar{\sigma}|_1 + \|\bar{\phi}\|_1 + \|\bar{\eta}\|_1\right) + h^{\frac{1}{2}}\|\bar{y}\|_{2,\infty}\,, \tag{11b}$$

$$\|(\bar{p}-\bar{p}_h,\bar{\chi}-\bar{\chi}_h)\|_W \lesssim h\left(|\bar{y}|_2 + |\bar{\chi}|_1 + \|\bar{\eta}\|_1 + h^{\frac{1}{2}}\|\bar{\phi}\|_1\right) + h^{\frac{1}{2}}\|\bar{p}\|_{2,\infty}\,. \tag{11c}$$

Proof:

Step 1 (Proof of (11a)) Using the L^2 projection $\Pi_h u$ and (8c) we obtain

$$\begin{aligned}(\bar{p}_h+\alpha\bar{u}_h,\bar{u}-\bar{u}_h) &= (\bar{p}_h+\alpha\bar{u}_h,\bar{u}-\Pi_h\bar{u}) + (\bar{p}_h+\alpha\bar{u}_h,\Pi_h\bar{u}-\bar{u}_h)\\ &\geqslant (\bar{p}_h+\alpha\bar{u}_h,\bar{u}-\Pi_h\bar{u})\,.\end{aligned} \tag{12}$$

Elementary algebra with (7c) and (12) shows

$$\begin{aligned}\alpha\|\bar{u}-\bar{u}_h\|^2 \leqslant &-(\bar{p}_h+\alpha\bar{u}_h,\bar{u}-\Pi_h\bar{u}) - (\bar{p}-p_h(\bar{u}),\bar{u}-\bar{u}_h)\\ &+(\bar{p}_h-p_h(\bar{u}),\bar{u}-\bar{u}_h) =: T_1+T_2+T_3\,,\end{aligned} \tag{13}$$

where T_1, T_2 and T_3 denote the first, second and third term on the right-hand side of (13). The Cauchy–Schwarz inequality estimates the term T_2 and hence we focus on T_1 and T_3 below.

The orthogonality of Π_h and elementary algebra reveal

$$T_1 = -(\bar{p}+\alpha\bar{u},\bar{u}-\Pi_h\bar{u}) + (\bar{p}-\bar{p}_h+\alpha(\bar{u}-\bar{u}_h),\bar{u}-\Pi_h\bar{u})$$

$$= -(\bar p - \Pi_h \bar p + \alpha(\bar u - \Pi_h \bar u), \bar u - \Pi_h \bar u) + (\bar p - \bar p_h + \alpha(\bar u - \bar u_h), \bar u - \Pi_h \bar u)$$
$$= -(\bar p - \Pi_h \bar p, \bar u - \Pi_h \bar u) + (\bar p - \bar p_h, \bar u - \Pi_h \bar u) = -(\bar p - \Pi_h \bar p, \bar u - \Pi_h \bar u)$$
$$+ (\bar p - p_h(\bar u), \bar u - \Pi_h \bar u) + (p_h(\bar u) - \bar p_h, \bar u - \Pi_h \bar u)\,, \tag{14}$$

with the term $p_h(\bar u)$ included in the last step. Subtract (8d) and (10c) (resp. (8e) and (10d)) and choose $(v_h, \tau_h) = (\bar y_h - y_h(\bar u), \bar\sigma_h - \sigma_h(\bar u))$ to obtain

$$a((\bar p_h - p_h(\bar u), \bar\chi_h - \chi_h(\bar u)), (\bar y_h - y_h(\bar u), \bar\sigma_h - \sigma_h(\bar u)))$$
$$= (\bar y_h - \bar y, \bar y_h - y_h(\bar u)) + (\bar\sigma_h - \bar\sigma, \bar\sigma_h - \sigma_h(\bar u))\,.$$

A similar manipulation with (7a) and (10a) (resp. (7b) and (10b)) yields

$$a((\bar y_h - y_h(\bar u), \bar\sigma_h - \sigma_h(\bar u)), (\bar p_h - p_h(\bar u), \bar\chi_h - \chi_h(\bar u)))$$
$$= (\bar u_h - \bar u, \bar p_h - p_h(\bar u))\,.$$

The symmetry of $a(\cdot\,,\cdot)$ shows that the right-hand side terms in the last two displayed relations are equal. This with elementary algebra reveals

$$T_3 = (\bar y - \bar y_h, \bar y_h - y_h(\bar u)) + (\bar\sigma - \bar\sigma_h, \bar\sigma_h - \sigma_h(\bar u))$$
$$= (\bar y - y_h(\bar u), \bar y_h - y_h(\bar u)) + (\bar\sigma - \sigma_h(\bar u), \bar\sigma_h - \sigma_h(\bar u))$$
$$- \|\bar y_h - y_h(\bar u)\|^2 - \|\bar\sigma_h - \sigma_h(\bar u)\|^2\,. \tag{15}$$

The expressions in (14) and (15) plus the Cauchy–Schwarz inequality in (13) yield

$$\alpha\|\bar u - \bar u_h\|^2 + \|\bar y_h - y_h(\bar u)\|^2 + \|\bar\sigma_h - \sigma_h(\bar u)\|^2$$
$$\leqslant \big(\|\bar p - \Pi_h \bar p\| + \|\bar p - p_h(\bar u)\|$$
$$+ \|p_h(\bar u) - \bar p_h\|\big)\|\bar u - \Pi_h \bar u\| + \|\bar p - p_h(\bar u)\|\,\|\bar u - \bar u_h\|$$
$$+ \|\bar y - y_h(\bar u)\|\,\|\bar y_h - y_h(\bar u)\| + \|\bar\sigma - \sigma_h(\bar u)\|\,\|\bar\sigma_h - \sigma_h(\bar u)\|\,. \tag{16}$$

Now, use (7), (8), and (10) with the approximation properties of $\Pi_{1h}\bar\phi$ and $\Pi_{1h}\bar\eta$ to get

$$\alpha_0\|(\bar p - p_h(\bar u), \bar\chi - \chi_h(\bar u))\|_W \lesssim \|\bar\eta - \Pi_{1h}\bar\eta\| \lesssim h\|\bar\eta\|_1\,, \tag{17}$$

$$\alpha_0 \|(\bar{y} - y_h(\bar{u}), \bar{\sigma} - \sigma_h(\bar{u}))\|_W \lesssim \|\bar{\phi} - \Pi_{1h}\bar{\phi}\| \lesssim h\|\bar{\phi}\|_1 \,.$$

Similarly, we have

$$\begin{aligned}\alpha_0 \|(\bar{p}_h - p_h(\bar{u}), \bar{\chi}_h - \chi_h(\bar{u}))\|_W \lesssim \|(\bar{y}_h - y_h(\bar{u}), \bar{\sigma}_h - \sigma_h(\bar{u}))\|_W \\ + \|(\bar{y} - y_h(\bar{u}), \bar{\sigma} - \sigma_h(\bar{u}))\|_W \,.\end{aligned} \tag{18}$$

Substitute (17) and (18) in (16). Use of Young's inequality and $\sum_i {a_i}^2 \leqslant (\sum_i a_i)^2$ where $a_i \geqslant 0$ for all i, concludes the proof of (11a).

Step 2 (Proof of (11b)) For $w_h := R_h\bar{y}$, with R_h defined in Lemma 2, and $\xi_h \in M_h$ defined by

$$(\xi_h, q_h) + (\nabla w_h, \nabla q_h) = 0 \quad \text{for all } q_h \in V_h \,,$$

we obtain $(w_h, \xi_h) \in \operatorname{Ker} B_h$. Hence $(\bar{y}_h - w_h, \bar{\sigma}_h - \xi_h) \in \operatorname{Ker} B_h$. The coercivity of $a(\cdot\,,\cdot)$ on $\operatorname{Ker} B_h$ reveals

$$\alpha_0 \|(\bar{y}_h - w_h, \bar{\sigma}_h - \xi_h)\|_W \leqslant \sup_{(v_h,\psi_h)\in \operatorname{Ker} B_h} \frac{a((\bar{y}_h - w_h, \bar{\sigma}_h - \xi_h), (v_h, \psi_h))}{\|(v_h, \psi_h)\|_W} \,. \tag{19}$$

For all $(v_h, \psi_h) \in \operatorname{Ker} B_h$, elementary algebra plus (7a) and (8a) show

$$\begin{aligned}a((\bar{y}_h - w_h, \bar{\sigma}_h - \xi_h), (v_h, \psi_h)) &= a((\bar{y} - w_h, \bar{\sigma} - \xi_h), (v_h, \psi_h)) \\ &\quad + a((\bar{y}_h - \bar{y}, \bar{\sigma}_h - \bar{\sigma}), (v_h, \psi_h)) \\ &= a((\bar{y} - w_h, \bar{\sigma} - \xi_h), (v_h, \psi_h)) \\ &\quad + b((v_h, \psi_h), \bar{\phi}) + (\bar{u}_h - \bar{u}, v_h) \,.\end{aligned} \tag{20}$$

For all $v_h \in V_h$, the definition of the H^1 projection operator Π_{1h} shows

$$\int_\Omega \nabla v_h \cdot \nabla(\bar{\phi} - \Pi_{1h}\bar{\phi})\, dx = -\int_\Omega v_h(\bar{\phi} - \Pi_{1h}\bar{\phi})\, dx \,. \tag{21}$$

As $(v_h, \psi_h) \in \operatorname{Ker} B_h$, $b((v_h, \psi_h), \Pi_{1h}\bar{\phi}) = 0$. This and (21) yield

$$b((v_h, \psi_h), \bar{\phi}) = b((v_h, \psi_h), \bar{\phi} - \Pi_{1h}\bar{\phi})$$

$$= -\int_\Omega \nu_h(\bar\phi - \Pi_{1h}\bar\phi)\,dx + \int_\Omega \psi_h(\bar\phi - \Pi_{1h}\bar\phi)\,dx\,.$$

Thus, using this in (20) and (19) yields

$$\begin{aligned}\alpha_0\|(\bar y_h - w_h, \bar\sigma_h - \xi_h)\|_W &\lesssim \sup_{(\nu_h,\xi_h)\in \operatorname{Ker} B_h} \frac{a((\bar y - w_h, \bar\sigma - \xi_h),(\nu_h,\psi_h))}{\|(\nu_h,\psi_h)\|_W} \\ &\quad + \|\bar\phi - \Pi_{1h}\bar\phi\| + \|\bar u - \bar u_h\| \\ &\lesssim \|a\|\,\|(\bar y - w_h, \bar\sigma - \xi_h)\|_W \\ &\quad + \|\bar\phi - \Pi_{1h}\bar\phi\| + \|\bar u - \bar u_h\|\,.\end{aligned}$$

Applying the triangle inequality reveals

$$\begin{aligned}\|(\bar y - \bar y_h, \bar\sigma - \bar\sigma_h)\|_W &\leqslant \|(\bar y - w_h, \bar\sigma - \xi_h)\|_W + \|(w_h - \bar y_h, \xi_h - \bar\sigma_h)\|_W \\ &\lesssim (1 + \alpha_0^{-1}\|a\|)\|(\bar y - w_h, \bar\sigma - \xi_h)\|_W \\ &\quad + \alpha_0^{-1}\|\bar\phi - \Pi_{1h}\bar\phi\| + \alpha_0^{-1}\|\bar u - \bar u_h\|\,.\end{aligned}$$

The term $\|\bar u - \bar u_h\|$ is estimated in Step 1, and now we estimate the terms $\|(\bar y - w_h, \bar\sigma - \xi_h)\|_W$ and $\|\bar\phi - \Pi_{1h}\bar\phi\|$ [6]. The definition of $\|\cdot\|_W$ and the triangle inequality show

$$\|(\bar y - w_h, \bar\sigma - \xi_h)\|_W \lesssim |\bar y - w_h|_1 + \|\bar\sigma - \Pi_{0h}\bar\sigma\| + \|\Pi_{0h}\bar\sigma - \xi_h\|\,. \qquad (22)$$

First, we note that the approximation property of the Ritz Projection R_h yields $|\bar y - w_h|_1 \lesssim h|\bar y|_2$ when $\bar y \in H^2(\Omega)$. Moreover, the approximation property of M_h [7] yields $\|\nu - \Pi_{0h}\nu_h\| \lesssim h|\nu|_1$ for $\nu \in H^1(\Omega)$. Now, we estimate the last term on the right of (22). Since $(w_h, \xi_h) \in \operatorname{Ker} B_h$ and $(\bar y, \bar\sigma) \in \operatorname{Ker} B$, we have

$$\int_\Omega (\nabla(\bar y - \xi_h)\cdot\nabla q_h + (\bar\sigma - \xi_h)q_h)\,dx = 0\,, \quad q_h \in V_h\,. \qquad (23)$$

Then

$$\|\xi_h - \Pi_h\bar\sigma\| \lesssim \sup_{q_h\in Q_h\setminus\{0\}} \frac{\int_\Omega(\xi_h - \Pi_h\bar\sigma)q_h\,dx}{\|q_h\|} \lesssim \sup_{q_h\in Q_h\setminus\{0\}} \frac{\int_\Omega(\xi_h - \bar\sigma)q_h\,dx}{\|q_h\|}$$

$$\lesssim \sup_{q_h \in Q_h \setminus \{0\}} \frac{\int_\Omega \nabla(\bar{y} - w_h) \cdot \nabla q_h \, dx}{\|q_h\|},$$

where we have used (23) in the last step. Since w_h is the Ritz projection of $\bar{y}$ onto V_h, the final result follows by using Lemma 2 with $r \to \infty$. The proof for the adjoint estimates follows exactly as above.

♠

4 Algebraic formulation

The biorthogonal system helps to statically condense out all auxiliary state and adjoint variables [6] and leads to a reduced system. We rewrite the variational inequality and use the primal-dual active set strategy [9] to solve the arising system. The algebraic system arising out of (8) is derived first.

Choosing test functions $\tau_h = 0$ and $v_h = 0$ in, successively, (8a), (8b), (8e) and (8d) lead to

$$\begin{aligned}
\int_\Omega \nabla\bar{\phi}_h \cdot \nabla v_h \, dx - \int_\Omega \bar{u}_h v_h \, dx = \int_\Omega f v_h \, dx, \quad v_h \in Q_h, \\
\int_\Omega \bar{\sigma}_h \tau_h \, dx + \int_\Omega \bar{\phi}_h \tau_h \, dx = 0, \quad \tau_h \in M_h, \\
\int_\Omega \nabla\bar{y}_h \cdot \nabla\psi_h \, dx + \int_\Omega \bar{\sigma}_h \psi_h \, dx = 0, \quad \psi_h \in V_h, \\
\int_\Omega \nabla\bar{\eta}_h \cdot \nabla v_h \, dx - \int_\Omega \bar{y}_h v_h \, dx = -\int_\Omega y_d v_h \, dx, \quad v_h \in Q_h, \\
\int_\Omega \bar{\chi}_h \tau_h \, dx + \int_\Omega \bar{\eta}_h \tau_h \, dx - \int_\Omega \bar{\sigma}_h \tau_h \, dx = 0, \quad \tau_h \in M_h, \\
\int_\Omega \nabla\bar{p}_h \cdot \nabla\psi_h \, dx + \int_\Omega \bar{\chi}_h \psi_h \, dx = 0, \quad \psi_h \in V_h.
\end{aligned} \tag{24}$$

Recall that $\{\rho_1, \rho_2, \dots, \rho_n\}$ and $\{\mu_1, \mu_2, \dots, \mu_n\}$ are, respectively, the finite element basis functions for V_h and M_h. Let $\{\rho_1, \rho_2, \dots, \rho_m\}$ denote the basis for Q_h, where $n-m$ denotes the number of boundary nodes. Let $\{\theta_1, \theta_2, \cdots, \theta_{\text{NT}}\}$ denote the basis for U_h, where NT denotes the number of triangles in the triangulation. Let the solution of (8) be

$$((\bar{y}_h, \bar{\sigma}_h), \bar{\phi}_h, \bar{u}_h, (\bar{p}_h, \bar{\chi}_h), \bar{\eta}_h) \in \boldsymbol{W}_h \times V_h \times U_h \times \boldsymbol{W}_h \times V_h \,,$$

and the data u_a and u_b be represented as

$$\begin{aligned} \bar{y}_h &= \sum_{i=1}^{m} \bar{y}_i \rho_i \,, & \bar{\sigma}_h &= \sum_{i=1}^{n} \bar{\sigma}_i \mu_i \,, & \bar{\phi}_h &= \sum_{i=1}^{n} \bar{\phi}_i \rho_i \,, \\ \bar{p}_h &= \sum_{i=1}^{m} \bar{p}_i \rho_i \,, & \bar{\chi}_h &= \sum_{i=1}^{n} \bar{\chi}_i \mu_i \,, & \bar{\eta}_h &= \sum_{i=1}^{n} \bar{\eta}_i \rho_i \,, \\ \bar{u}_h &= \sum_{i=1}^{\text{NT}} \bar{u}_i \theta_i \,, & U_a &= \sum_{i=1}^{\text{NT}} u_a \theta_i \,, & U_b &= \sum_{i=1}^{\text{NT}} u_b \theta_i \,. \end{aligned}$$

Let

$$\begin{aligned} &\vec{y} = (\bar{y}_i)_{i=1}^{m} \,, \quad \vec{\sigma} = (\bar{\sigma}_i)_{i=1}^{n} \,, \quad \vec{\phi} = (\bar{\phi}_i)_{i=1}^{n} \,, \quad \vec{p} = (\bar{p}_i)_{i=1}^{m} \,, \\ &\vec{\chi} = (\bar{\chi}_i)_{i=1}^{n} \,, \quad \vec{\eta} = (\bar{\eta}_i)_{i=1}^{n} \quad \text{and} \quad \vec{u} = (\bar{u}_i)_{i=1}^{\text{NT}} \,. \end{aligned}$$

Define the matrices

$$\begin{aligned} &\mathtt{A} = \left(\int_\Omega \nabla\rho_i \cdot \nabla\rho_j \quad dx \right)_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant m}} , \quad \mathtt{B} = \left(\int_\Omega \theta_i \rho_j \quad dx \right)_{\substack{1\leqslant i\leqslant \text{NT} \\ 1\leqslant j\leqslant m}} , \\ &\mathtt{M} = \left(\int_\Omega \mu_i \mu_j \quad dx \right)_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant n}} , \quad \mathtt{D} = \left(\int_\Omega \rho_i \mu_j \quad dx \right)_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant n}} , \\ &\mathtt{G} = \left(\int_\Omega \rho_i \rho_j \quad dx \right)_{\substack{1\leqslant i\leqslant m \\ 1\leqslant j\leqslant m}} , \quad \mathtt{J} = \left(\int_\Omega \theta_i \theta_j \quad dx \right)_{\substack{1\leqslant i\leqslant \text{NT} \\ 1\leqslant j\leqslant \text{NT}}} , \\ &\mathtt{E} := \alpha^{-1} \mathtt{J}^{-1} (\mathtt{I} - \mathbf{X}^a - \mathbf{X}^b) \,, \quad \vec{f} = \left(\int_\Omega f \rho_j \right)_{1\leqslant j\leqslant m} , \quad \vec{y}_d = \left(\int_\Omega y_d \rho_j \right)_{1\leqslant j\leqslant m} . \end{aligned}$$

In the final line, $\mathbf{X}^a$ and $\mathbf{X}^b$ are matrices of sizes $NT \times NT$ and are the discrete analogues of the characteristic functions that correspond to the active sets (Tröltzsch [9] provides more details), and I is the identity matrix of size $NT \times NT$. Note that D and J are diagonal matrices.

The matrix form corresponding to (24) is

$$\begin{bmatrix} \begin{bmatrix} 0 & 0 & A^T \\ 0 & M & D \\ A & D^T & 0 \end{bmatrix} & \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \begin{matrix} -B^T \\ 0 \\ 0 \end{matrix} \\ \begin{matrix} -G & 0 & 0 \\ 0 & -M & 0 \\ 0 & 0 & 0 \end{matrix} & \begin{bmatrix} 0 & 0 & A^T \\ 0 & M & D \\ A & D^T & 0 \end{bmatrix} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \end{matrix} & \begin{matrix} EB & 0 & 0 \end{matrix} & I \end{bmatrix} \begin{bmatrix} \vec{y} \\ \vec{\sigma} \\ \vec{\phi} \\ \vec{p} \\ \vec{\chi} \\ \vec{\eta} \\ \vec{u} \end{bmatrix} = \begin{bmatrix} \vec{f} \\ 0 \\ 0 \\ -\vec{y}_d \\ 0 \\ 0 \\ \mathbf{X}^a u_a + \mathbf{X}^b u_b \end{bmatrix}.$$

Following Tröltzsch [9] verbatim, the variational inequality in (8c) is reformulated as an equation displayed in the last line of the above matrix. Since the matrix D is diagonal, we do the static condensation of unknowns $\vec{\sigma}$ and $\vec{\phi}$ (respectively $\vec{\chi}$ and $\vec{\eta}$) and arrive at the formulation

$$\begin{bmatrix} \mathcal{S} & 0 & -B^T \\ -G - \mathcal{S} & \mathcal{S} & 0 \\ 0 & EB & I \end{bmatrix} \begin{bmatrix} \vec{y} \\ \vec{p} \\ \vec{u} \end{bmatrix} = \begin{bmatrix} \vec{f} \\ -\vec{y}_d \\ \mathbf{X}^a u_a + \mathbf{X}^b u_b \end{bmatrix},$$

where $\mathcal{S} = A^T D^{-1} M (D^{-1})^T A$.

5 Numerical results

We present a numerical example to validate the a priori estimates derived in Section 3. Consider the example of Gudi et al. [5] with the domain $\Omega = (0,1)^2$. The exact state and adjoint variables are chosen as $\bar{y} = \sin^2(\pi x)\sin^2(\pi y)$ and $\bar{p} = \sin^2(\pi x)\sin^2(\pi y)$, and the exact control as $\bar{u}(x) = \Pi_{[-750,-50]}(-\bar{p}(x)/\alpha)$ with $\alpha = 10^{-3}$. Then we compute $f = \Delta^2\bar{y} - \bar{u}$ and $y_d = \bar{y} - \Delta^2\bar{p} + \Delta^2\bar{y}$. The errors and order of convergence (OOC) of the numerical solutions are

Table 1: Errors and orders of convergence for the state variable.

h	$\|\bar{y}-\bar{y}_h\|$	OoC	$\|\bar{y}-\bar{y}_h\|_1$	OoC
2^{-2}	0.6273	—	0.6909	—
2^{-3}	0.3516	0.8354	0.3940	0.8101
2^{-4}	0.1403	1.3248	0.1764	1.1595
2^{-5}	0.0421	1.7365	0.0713	1.3071
2^{-6}	0.0112	1.9126	0.0313	1.1857
2^{-7}	0.0028	1.9831	0.0150	1.0674

Table 2: Errors and orders of convergence for the adjoint and control variables.

h	$\|\bar{p}-\bar{p}_h\|$	OoC	$\|\bar{p}-\bar{p}_h\|_1$	OoC	$\|\bar{u}-\bar{u}_h\|$	OoC
2^{-2}	0.6274	—	0.6910	—	0.6303	—
2^{-3}	0.3517	0.8352	0.3941	0.8100	0.3344	0.9145
2^{-4}	0.1404	1.3246	0.1765	1.1594	0.1323	1.3381
2^{-5}	0.0421	1.7364	0.0713	1.3073	0.0512	1.3678
2^{-6}	0.0112	1.9125	0.0313	1.1859	0.0225	1.1873
2^{-7}	0.0028	1.9831	0.0150	1.0675	0.0108	1.0619

tabulated in Tables 1 and 2. We see that approximations to both state and adjoint variables converge with almost order two in the L^2 norm and order one in the H^1 norm, whereas the control variable converges with order one in the L^2, thus confirming the theoretical estimates.

Acknowledgement

The first author gratefully acknowledges the hospitality provided by the Indian Institute of Technology, Mumbai during his visit in 2017 and 2023. This work was completed during those visits. The second author gratefully acknowledges the SERB POWER Fellowship SPF/2020/000019.

References

[1] L. Boudjaj, A. Naji, and F. Ghafrani. "Solving biharmonic equation as an optimal control problem using localized radial basis functions collocation method". In: *Eng. Anal. Bound. Elements* 107 (2019), pp. 208–217. DOI: 10.1016/j.enganabound.2019.07.007 (cit. on p. C47).

[2] W. Cao and D. Yang. "Ciarlet–Raviart mixed finite element approximation for an optimal control problem governed by the first biharmonic equation". In: *J. Comput. App. Math.* 233.2 (2009), pp. 372–388. DOI: 10.1016/j.cam.2009.07.039 (cit. on p. C47).

[3] P. G. Ciarlet. *The finite element method for elliptic problems*. Vol. 40. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2002. DOI: 10.1137/1.9780898719208. (Cit. on pp. C46, C47, C48).

[4] V. Girault and P.-A. Raviart. *Finite element methods for Navier–Stokes equations*. Vol. 5. Springer Series in Computational Mathematics. Springer-Verlag, 1986. DOI: 10.1007/978-3-642-61623-5 (cit. on pp. C46, C51).

[5] T. Gudi, N. Nataraj, and K. Porwal. "An interior penalty method for distributed optimal control problems governed by the biharmonic operator". In: *Comput. Math. App.* 68.12 (2014), pp. 2205–2221. DOI: 10.1016/j.camwa.2014.08.012 (cit. on pp. C47, C58).

[6] B. P. Lamichhane. "A mixed finite element method for the biharmonic problem using biorthogonal or quasi-biorthogonal systems". In: *J. Sci. Comput.* 46.3 (2011), pp. 379–396. DOI: 10.1007/s10915-010-9409-7. (Cit. on pp. C47, C49, C50, C55, C56).

[7] B. P. Lamichhane and E. Stephan. "A symmetric mixed finite element method for nearly incompressible elasticity based on biorthogonal systems". In: *Numer. Meth. Part. Diff. Eq.* 28 (2012), pp. 1336–1353. DOI: 10.1002/num.20683 (cit. on p. C55).

[8] J. L. Lions. *Optimal control of systems governed by partial differential equations*. Vol. 170. Die Grundlehren der mathematischen Wissenschaften. Springer-Verlag, New York-Berlin, 1971. URL: https://link.springer.com/book/9783642650260 (cit. on p. C48).

[9] F. Tröltzsch. *Optimal control of partial differential equations: Theory, methods and applications*. Vol. 112. Graduate Studies in Mathematics. American Mathematical Society, 2010. DOI: 10.1090/gsm/112. (Cit. on pp. C46, C48, C56, C58).

[10] G. N. Wells, E. Kuhl, and K. Garikipati. "A discontinuous Galerkin method for the Cahn–Hilliard equation". In: *J. Comput. Phys.* 218 (2006), pp. 860 –877. DOI: 10.1016/j.jcp.2006.03.010 (cit. on p. C46).

Author addresses

1. **Bishnu P. Lamichhane**, School of Information & Physical Sciences, University of Newcastle, Callaghan, NSW 2308, Australia
mailto:Bishnu.Lamichhane@newcastle.edu.au

2. **Neela Nataraj**, Department of Mathematics, Indian Institute of Technology Bombay, Powai, Mumbai, 400076, India
mailto:neela@math.iitb.ac.in

3. **Deepanshu Verma**, Department of Mathematics, Emory University, Atlanta, Georgia 30322, USA
mailto:dverma4@emory.edu