PDE-Constrained Optimal Control for Static Heating

Yue Wu*

May 13, 2024

Contents

1 Introduction **1**
1.1 Background . . . 1
1.2 Objective . . . 2

2 The continuous problem **2**
2.1 Derivation . . . 2
2.2 Well-posedness . . . 3

3 The discrete problem **3**
3.1 Derivation . . . 3
3.2 Well-posedness . . . 4
3.3 A priori error estimates . . . 4
3.3.1 Estimates in the energy norm . . . 4
3.3.2 Improved estimates . . . 5
3.4 Numerical methods . . . 6

4 Numerical examples **7**
4.1 Example 1 . . . 7
4.2 Example 2 . . . 9

5 Conclusion **11**

References **11**

1 Introduction

1.1 Background

In various applications, ranging from electronic devices to industrial processes, controlling temperature distributions is essential for optimizing performance and efficiency. Under the Fourier law of heat transfer, the basic equations for the temperature distribution at equilibrium states are

$$\boldsymbol{q} = -\boldsymbol{A}\nabla u, \tag{1}$$

$$\nabla \cdot \boldsymbol{q} = f, \tag{2}$$

where u is the temperature, $\boldsymbol{A}$ is the thermal conductivity coefficient tensor, $\boldsymbol{q}$ is the thermal heat flux vector, and f is the heat source. Equation (1) describes the linear relation between the heat flux and the gradient of temperature, and equation (2) results from the conservation of internal energy.

Let's consider the case where the domain of interest Ω is bounded. In order to close the problem we have to introduce boundary conditions. Let's consider a simple case where the temperature outside of this region is given, which can be interpreted as the Dirichlet boundary condition

$$u|_{\partial\Omega} = g. \tag{3}$$

*Division of Applied Mathematics, Brown University, Providence, RI 02912, USA (yue_wu3@brown.edu).

Now, combining equations (1), (2) and (3) gives the static heat equation

$$\begin{cases} -\nabla \cdot (\boldsymbol{A}\nabla u) = f, & \text{in } \Omega, \\ u = g, & \text{on } \partial\Omega. \end{cases} \tag{4}$$

Suppose that we have a targeted subdomain $\Omega_0 \subset \Omega$ and a given targeted temperature u_0, our goal is to minimize a combination of the L^2 error between u and u_0 on Ω_0 plus the cost of control, which can be modeled as the L^2 norm of f, under the constraint given by (4). Then the problem is formulated as

$$\begin{aligned} &\text{find } f \text{ that minimizes } J[u, f], \\ &\text{subject to } \begin{cases} -\nabla \cdot (\boldsymbol{A}\nabla u) = f, & \text{in } \Omega, \\ u = g, & \text{on } \partial\Omega, \end{cases} \end{aligned} \tag{5}$$

where

$$J[u, f] := \frac{1}{2} \|u - u_0\|^2_{L^2(\Omega_0)} + \frac{\gamma}{2} \|f\|^2_{L^2(\Omega)}.$$

Here, $\gamma > 0$ is a weighting parameter, $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain, and $\Omega_0 \subset \Omega$ is a subdomain.

For simplicity, we assume $g = 0$ and $\boldsymbol{A} = \boldsymbol{I}$.

1.2 Objective

The primary goal of this project is to address the PDE-constrained optimal control problem derived from the model of optimal heating strategy. The specific objective is to solve the optimal external heat source f that minimizes the error on the targeted region plus the controlling cost by MFEM. This project will teach me how to integrate knowledge from both optimization and finite element methods to solve realistic problems and will enhance my understanding in scientific programming using MFEM.

2 The continuous problem

2.1 Derivation

If we rewrite the heat equation constraint (4) in (5) into its weak form and treat u also as an unknown variable, then problem (5) becomes

$$\begin{aligned} &\text{find } (u, f) \in H^1_0(\Omega) \times L^2(\Omega) \text{ that minimizes } J[u, f], \\ &\text{subject to: } \forall v \in H^1_0(\Omega), (\nabla u, \nabla v)_\Omega = (f, v)_\Omega. \end{aligned} \tag{6}$$

In order for the minimization problem (6) to be well-defined, we should require at least that $u_0 \in L^2(\Omega_0)$ and $f \in L^2(\Omega)$.

The well-posedness of the constrained minimization problem (6) is easy to check via classical calculus of variations [4, Section 8].

By the method of Lagrange multipliers, the equivalent weak form of (6) is

$$\begin{aligned} &\text{find } (u, f, z) \in H^1_0(\Omega) \times L^2(\Omega) \times H^1_0(\Omega) \text{ s.t. } \forall (v, \varphi, w) \in H^1_0(\Omega) \times L^2(\Omega) \times H^1_0(\Omega), \\ &\begin{cases} a((u, f), (v, \varphi)) + b((v, \varphi), z) = F((v, \varphi)), \\ b((u, f), w) = G(w), \end{cases} \end{aligned} \tag{7a}$$

where

$$a((u, f), (v, \varphi)) := (u, v)_{\Omega_0} + \gamma(f, \varphi)_\Omega, \tag{7b}$$

$$b((u, f), w) := (\nabla u, \nabla w)_\Omega - (f, w)_\Omega, \tag{7c}$$

$$F((v, \varphi)) := (u_0, v)_{\Omega_0}, \tag{7d}$$

$$G(w) := 0. \tag{7e}$$

We call equations (7) the continuous equations. Here and below, we choose the discrete 2-norm for the product space, i.e.

$$\|(u, f)\|_{H^1_0(\Omega) \times L^2(\Omega)} := \left(\|u\|^2_{H^1(\Omega)} + \|f\|^2_{L^2(\Omega)} \right)^{\frac{1}{2}}.$$

2.2 Well-posedness

By the Babuška–Brezzi (BB) theory, to prove the well-posedness and a priori estimates for the solutions of (7), it suffices to check the BB conditions for the bilinear forms a and b. It's easy to verify that $\|a\| \leqslant \max\{1, \gamma\}$, $\|b\| \leqslant 1$, $\|F\| \leqslant \|u_0\|_{L^2(\Omega_0)}$, and $\|G\| = 0$.

By direct computation,

$$\begin{aligned}
\beta := & \inf_{w \in H_0^1(\Omega)} \sup_{(u,f) \in H_0^1(\Omega) \times L^2(\Omega)} \frac{|b((u,f),w)|}{\|(u,f)\|_{H_0^1(\Omega)\times L^2(\Omega)} \|w\|_{H^1(\Omega)}} \\
\geqslant & \inf_{w \in H_0^1(\Omega)} \frac{|b((w,-w),w)|}{\|(w,-w)\|_{H_0^1(\Omega)\times L^2(\Omega)} \|w\|_{H^1(\Omega)}} \\
= & \inf_{w \in H_0^1(\Omega)} \frac{\|w\|^2_{H^1(\Omega)}}{\|(w,-w)\|_{H_0^1(\Omega)\times L^2(\Omega)} \|w\|_{H^1(\Omega)}} \\
\geqslant & \frac{1}{\sqrt{2}},
\end{aligned}$$

so the bilinear form b satisfies the BB inf-sup condition.

Actually,

$$\begin{aligned}
\ker(B) &= \left\{(u,f) \in H_0^1(\Omega) \times L^2(\Omega) : \forall v \in H_0^1(\Omega), (\nabla u, \nabla v)_\Omega = (f,v)_\Omega\right\} \\
&= \left\{(u,f) \in H_0^1(\Omega) \times L^2(\Omega) : -\Delta u = f \text{ weakly in } H^{-1}(\Omega)\right\}.
\end{aligned}$$

Then by elliptic estimates,

$$\begin{aligned}
\alpha := & \inf_{(u,f)\in\ker(B)} \sup_{(v,\varphi)\in\ker(B)} \frac{|a((u,f),(v,\varphi))|}{\|(u,f)\| \, \|(v,\varphi)\|} \\
\geqslant & \inf_{(u,f)\in\ker(B)} \frac{|a((u,f),(u,f))|}{\|(u,f)\|^2} \\
\geqslant & \inf_{(u,f)\in\ker(B)} \frac{\gamma \|f\|^2_{L^2(\Omega)}}{\|u\|^2_{H^1(\Omega)} + \|f\|^2_{L^2(\Omega)}} \\
\geqslant & \frac{\gamma}{C_\Omega + 1}.
\end{aligned}$$

This implies that a is coercive on $\ker(B)$. Therefore, $\forall (v,\varphi) \in \ker(B)$, if $\forall (u,\varphi) \in \ker(B)$, $a((u,f),(v,\varphi)) = 0$, then $(v,\varphi) = 0$.

Hence, problem (7) is well-posed according to the Babuška–Brezzi theorem [3, Theorem 49.13], and the exact solution satisfies the a priori estimates,

$$\begin{aligned}
\|(u,f)\|_{H^1(\Omega)\times L^2(\Omega)} &\leqslant c_1 \|F\| + c_2 \|G\|, \\
\|z\|_{H^1(\Omega)} &\leqslant c_3 \|F\| + c_4 \|G\|,
\end{aligned}$$

where $c_1 = \frac{1}{\alpha}$, $c_2 = \frac{1}{\beta}(1 + \frac{\|a\|}{\alpha})$, $c_3 = \frac{1}{\beta}(1 + \frac{\|a\|}{\alpha})$, and $c_4 = \frac{\|a\|}{\beta^2}(1 + \frac{\|a\|}{\beta})$.

3 The discrete problem

3.1 Derivation

Let's discretize the problem using a conforming Galerkin method. Let $V_h \subset H_0^1(\Omega)$, $W_h \subset L^2(\Omega)$ and $Q_h \subset H_0^1(\Omega)$ be the (non trivial) finite element spaces. The discretized problem for (7) is

$$\begin{aligned}
&\text{find } (u_h, f_h, z_h) \in V_h \times W_h \times Q_h \text{ s.t. } \forall (v_h, \varphi_h, w_h) \in V_h \times W_h \times Q_h, \\
&\begin{cases} a((u_h,f_h),(v_h,\varphi_h)) + b((v_h,\varphi_h),z_h) = F((v_h,\varphi_h)), \\ b((u_h,f_h),w_h) = G(w_h). \end{cases}
\end{aligned} \tag{8}$$

We define the mappings $A_h : V_h \times W_h \mapsto V_h \times W_h$ and $B_h : V_h \times W_h \mapsto Q_h$ by

$$\begin{aligned}
&\forall (v_h,\varphi_h) \in V_h \times W_h, (A_h(u_h,f_h),(v_h,\varphi_h)) = a((u_h,f_h),(v_h,\varphi_h)), \\
&\forall w_h \in Q_h, (B_h(u_h,f_h),w_h) = b((u_h,f_h),w_h).
\end{aligned}$$

Their upper bounds are controlled, i.e. $\|A_h\| \leqslant \|a\|$ and $\|B_h\| \leqslant \|b\|$.

3.2 Well-posedness

Assume that $Q_h \subset V_h$. By Poincaré's inequality,

$$\begin{aligned}\beta_h :=& \inf_{w_h\in Q_h}\sup_{(u_h,f_h)\in V_h\times W_h}\frac{|b((u_h,f_h),w_h)|}{\|(u_h,f_h)\|\,\|w_h\|_{H^1(\Omega)}}\\ \geqslant& \inf_{w_h\in Q_h}\frac{|b((w_h,0),w_h)|}{\|(w_h,0)\|\,\|w_h\|_{H^1(\Omega)}}\\ =& \inf_{w_h\in Q_h}\frac{\|\nabla w_h\|^2_{L^2(\Omega)}}{\|w_h\|^2_{H^1(\Omega)}}\\ \geqslant& C_\Omega.\end{aligned}$$

Notice that

$$\ker(B_h) = \{(u_h,f_h)\in V_h\times W_h : \forall v_h\in Q_h, (\nabla u_h,\nabla v_h)_\Omega = (f_h,v_h)_\Omega\}.$$

If $V_h \subset Q_h$, by the conforming finite element method for the Poisson equation, $\forall (u_h,f_h)\in\ker(B_h)$, $\|u_h\|^2_{H^1(\Omega)} \leqslant C_\Omega\|f_h\|^2_{L^2(\Omega)}$. Thus

$$\begin{aligned}\alpha_h :=& \inf_{(u_h,f_h)\in\ker(B_h)}\sup_{(v_h,\varphi_h)\in\ker(B_h)}\frac{|a((u_h,f_h),(v_h,\varphi_h))|}{\|(u_h,f_h)\|\,\|(v_h,\varphi_h)\|}\\ \geqslant& \inf_{(u_h,f_h)\in\ker(B_h)}\frac{|a((u_h,f_h),(u_h,f_h))|}{\|(u_h,f_h)\|^2}\\ \geqslant& \inf_{(u_h,f_h)\in\ker(B_h)}\frac{\gamma\|f_h\|^2_{L^2(\Omega)}}{\|u_h\|^2_{H^1(\Omega)}+\|f_h\|^2_{L^2(\Omega)}}\\ =& \inf_{(u_h,f_h)\in\ker(B_h)}\frac{\gamma}{\frac{\|u_h\|^2_{H^1(\Omega)}}{\|f_h\|^2_{L^2(\Omega)}}+1}\\ \geqslant& \frac{\gamma}{C_\Omega+1}.\end{aligned}$$

In summary, if $V_h = Q_h$, then the discrete problem (8) is well-posed by [3, Proposition 50.1]. Moreover, the discrete solution satisfies the a priori estimates,

$$\begin{aligned}\|(u_h,f_h)\|_{H^1(\Omega)\times L^2(\Omega)} &\leqslant c_1\|F\| + c_2\|G\|,\\ \|z_h\|_{H^1(\Omega)} &\leqslant c_3\|F\| + c_4\|G\|,\end{aligned}$$

where $c_1 = \frac{1}{\alpha_h}$, $c_2 = \frac{1}{\beta_h}(1+\frac{\|a\|}{\alpha_h})$, $c_3 = \frac{1}{\beta_h}(1+\frac{\|a\|}{\alpha_h})$, and $c_4 = \frac{\|a\|}{\beta_h^2}(1+\frac{\|a\|}{\beta_h})$.

3.3 A priori error estimates

3.3.1 Estimates in the energy norm

Assume $V_h = Q_h$. It's easy to check that the mapping $\Pi_h : (u,f)\in H^1_0(\Omega)\times L^2(\Omega)\mapsto (u_h,0)\in V_h\times W_h$ defined by

$$\forall w_h\in V_h, (\nabla u_h,\nabla w_h)_\Omega = (\nabla u,\nabla w_h)_\Omega - (f,w_h)_\Omega$$

is the Fortin operator for b. Moreover, $\|\Pi_h\|\leqslant C_\Omega$. By [3, Corollary 50.5],

$$\begin{aligned}&\|(u-u_h,f-f_h)\|_{H^1_0(\Omega)\times L^2(\Omega)}\\ \leqslant& c'_{1,h}\inf_{(v_h,\varphi_h)\in V_h\times W_h}\|(u-v_h,f-\varphi_h)\|_{H^1_0(\Omega)\times L^2(\Omega)} + c_{2,h}\inf_{w_h\in V_h}\|z-w_h\|_{H^1(\Omega)},\\ &\|z-z_h\|_{H^1_0(\Omega)}\\ \leqslant& c'_{3,h}\inf_{(v_h,\varphi_h)\in V_h\times W_h}\|(u-v_h,f-\varphi_h)\|_{H^1_0(\Omega)\times L^2(\Omega)} + c_{4,h}\inf_{w_h\in V_h}\|z-w_h\|_{H^1(\Omega)},\end{aligned}$$

where

$$c'_{1,h} = (1+\frac{\|a\|}{\alpha_h})(1+\|\Pi_h\|)\leqslant C(1+C\max\{1,\gamma^{-1}\}),$$

$$\begin{aligned}
c_{2,h} &= \frac{\|b\|}{\alpha_h} \leqslant (1+C_\Omega)\gamma^{-1} \leqslant C\gamma^{-1}, \\
c'_{3,h} &= c'_{1,h}\frac{\|a\|}{\beta_h} \leqslant C(1+\max\left\{\gamma^{-1},\gamma\right\}), \\
c_{4,h} &= 1 + \frac{\|b\|}{\beta_h} + c_{2,h}\frac{\|a\|}{\beta_h} \leqslant C(1+\max\left\{1,\gamma^{-1}\right\}).
\end{aligned}$$

Suppose $\{\mathcal{T}_h\}_h$ is a family of shape-regular conforming triangulations of Ω. Assume we take $V_h = Q_h \supset \mathcal{P}^k(\mathcal{T}_h) \cap H_0^1(\Omega)$ and $W_h \supset \mathcal{P}^l(\mathcal{T}_h) \cap L^2(\Omega)$ for some integers $k \geqslant 1$ and $l \geqslant 0$. By the best approximation error [2, Theorem 22.14 and 11.13], for the exact solution $(u,f,z) \in H^{\tilde{k}_1+1}(\Omega) \times H^{\tilde{l}_0+1}(\Omega) \times H^{\tilde{k}_2+1}(\Omega)$, the errors will satisfy

$$\begin{aligned}
&\|u-u_h\|_{H^1(\Omega)} + \|f-f_h\|_{L^2(\Omega)} \lesssim c'_{1,h}(h^{k_1}|u|_{H^{k_1+1}} + h^{l_0+1}|f|_{H^{l_1+1}}) + c_{2,h}h^{k_2}|z|_{H^{k_2+1}}, \\
&\|z-z_h\|_{H^1(\Omega)} \lesssim c'_{3,h}(h^{k_1}|u|_{H^{k_1+1}} + h^{l_0+1}|f|_{H^{l_0+1}}) + c_{4,h}h^{k_2}|z|_{H^{k_2+1}},
\end{aligned}$$

where $k_1 = \min\left\{k,\tilde{k}_1\right\}$, $k_2 = \min\left\{k,\tilde{k}_2\right\}$, and $l_0 = \min\left\{l,\tilde{l}_0\right\}$. In conclusion, for each fixed $\gamma > 0$,

$$\begin{aligned}
&\|u-u_h\|_{H^1(\Omega)} + \|f-f_h\|_{L^2(\Omega)} + \|z-z_h\|_{H^1(\Omega)} \\
\lesssim& h^{k_1}|u|_{H^{k_1+1}} + h^{l_0+1}|f|_{H^{l_0+1}} + h^{k_2}|z|_{H^{k_2+1}} \\
\lesssim& C(u,f,z)h^{\min\{k_1,k_2,l_0+1\}}.
\end{aligned} \tag{9}$$

3.3.2 Improved estimates

We use the Oden–Nitsche technique to improve the error estimates. Let

$$t((u,f,z),(v,\varphi,w)) := a((u,f),(v,\varphi)) + b((v,\varphi),z) + b((u,f),w)$$

and consider the following auxiliary problem,

$$\begin{aligned}
&\text{find } (\tilde{v},\tilde{\varphi},\tilde{w}) \in H_0^1(\Omega)\times L^2(\Omega)\times H_0^1(\Omega) \text{ s.t. } \forall(\tilde{u},\tilde{f},\tilde{z}) \in H_0^1(\Omega)\times L^2(\Omega)\times H_0^1(\Omega), \\
&t((\tilde{u},\tilde{f},\tilde{z}),(\tilde{v},\tilde{\varphi},\tilde{w})) = (u-u_h,\tilde{u})_\Omega + (z-z_h,\tilde{z})_\Omega.
\end{aligned} \tag{10}$$

The well-posedness of the auxiliary problem (10) follows from that of the original problem (7). Moreover, we can deduce (in the weak sense)

$$\begin{aligned}
&-\Delta\tilde{w} = (u-u_h) - 1_{\Omega_0}\tilde{v}, \\
&-\Delta\tilde{v} - \tilde{\varphi} = (z-z_h) + \tilde{\varphi}, \\
&\gamma\tilde{\varphi} = \tilde{w}.
\end{aligned}$$

Using the elliptic regularity theory (for Lipschitz domain) and a priori estimates for the auxiliary problem (10), we obtain

$$\begin{aligned}
&\|\tilde{w}\|_{H^2(\Omega)} \lesssim \|u-u_h\|_{L^2(\Omega)} + \|\tilde{v}\|_{L^2(\Omega_0)} \lesssim \|u-u_h\|_{L^2(\Omega)} + \|z-z_h\|_{L^2(\Omega)}, \\
&\|\tilde{v}\|_{H^2(\Omega)} \lesssim \|z-z_h\|_{L^2(\Omega)} + \|\tilde{\varphi}\|_{L^2(\Omega)} \lesssim \|u-u_h\|_{L^2(\Omega)} + \|z-z_h\|_{L^2(\Omega)}, \\
&\|\tilde{\varphi}\|_{H^1(\Omega)} = \frac{1}{\gamma}\|\tilde{w}\|_{H^1(\Omega)} \lesssim \|u-u_h\|_{L^2(\Omega)} + \|z-z_h\|_{L^2(\Omega)}.
\end{aligned}$$

Choose $(\tilde{u},\tilde{f},\tilde{z}) = (u-u_h, f-f_h, z-z_h)$ in (10) and use the equations (7) and (8). For any $(\tilde{v}_h,\tilde{\varphi}_h,\tilde{w}_h) \in V_h \times W_h \times V_h$,

$$\begin{aligned}
&\|u-u_h\|^2_{L^2(\Omega)} + \|z-z_h\|^2_{L^2(\Omega)} \\
=&(u-u_h,\tilde{u})_\Omega + (z-z_h,\tilde{z})_\Omega \\
=&t((\tilde{u},\tilde{f},\tilde{z}),(\tilde{v},\tilde{\varphi},\tilde{w})) \\
=&t((\tilde{u},\tilde{f},\tilde{z}),(\tilde{v}-\tilde{v}_h,\tilde{\varphi}-\tilde{\varphi}_h,\tilde{w}-\tilde{w}_h)) \\
=&(\tilde{u},\tilde{v}-\tilde{v}_h)_{\Omega_0} + (\nabla(\tilde{v}-\tilde{v}_h),\nabla\tilde{z})_\Omega + (\nabla\tilde{u},\nabla(\tilde{w}-\tilde{w}_h))_\Omega - (\tilde{\varphi}-\tilde{\varphi}_h,\tilde{z})_\Omega
\end{aligned}$$

$$
\begin{aligned}
&+(\gamma\tilde{f},\tilde{\varphi}-\tilde{\varphi}_h)_\Omega-(\tilde{f},\tilde{w}-\tilde{w}_h)_\Omega\\
\lesssim&\left(\|\tilde{u}\|_{H^1(\Omega)}+\left\|\tilde{f}\right\|_{L^2(\Omega)}+\|\tilde{z}\|_{H^1(\Omega)}\right)\left(\|\tilde{v}-\tilde{v}_h\|_{H^1(\Omega)}+\|\tilde{\varphi}-\tilde{\varphi}_h\|_{L^2(\Omega)}+\|\tilde{w}-\tilde{w}_h\|_{H^1(\Omega)}\right)\\
\lesssim&C(u,f,z)h^{\min\{k_1,k_2,l_0+1\}}\left(\|\tilde{v}-\tilde{v}_h\|_{H^1(\Omega)}+\|\tilde{\varphi}-\tilde{\varphi}_h\|_{L^2(\Omega)}+\|\tilde{w}-\tilde{w}_h\|_{H^1(\Omega)}\right).
\end{aligned}
$$

Take the infimum w.r.t. $(\tilde{v}_h,\tilde{\varphi}_h,\tilde{w}_h)$. By the best approximation error for V_h $(k\geqslant 1)$ and W_h $(l\geqslant 0)$,

$$
\begin{aligned}
&\|u-u_h\|^2_{L^2(\Omega)}+\|z-z_h\|^2_{L^2(\Omega)}\\
\lesssim&C(u,f,z)h^{\min\{k_1,k_2,l_0+1\}}\left(h\,|\tilde{v}|_{H^2(\Omega)}+h\,|\tilde{\varphi}|_{H^1(\Omega)}+h\,|\tilde{w}|_{H^2(\Omega)}\right)\\
\lesssim&C(u,f,z)h^{\min\{k_1+1,k_2+1,l_0+2\}}\left(\|u-u_h\|_{L^2(\Omega)}+\|z-z_h\|_{L^2(\Omega)}\right).
\end{aligned}
$$

Therefore,

$$
\|u-u_h\|_{L^2(\Omega)}+\|z-z_h\|_{L^2(\Omega)}\lesssim C(u,f,z)h^{\min\{k_1+1,k_2+1,l_0+2\}}.
$$

Now, let Π_{W_h} be the L^2-projection onto W_h. By (7) and (8), $\forall\varphi_h\in W_h$,

$$
\gamma^{-1}(\Pi_{W_h}z-z_h,\varphi_h)_\Omega=\gamma^{-1}(z-z_h,\varphi_h)_\Omega=(f-f_h,\varphi_h)_\Omega=(\Pi_{W_h}f-f_h,\varphi_h)_\Omega.
$$

By duality,

$$
\|\Pi_{W_h}f-f_h\|_{L^2(\Omega)}\leqslant\gamma^{-1}\|\Pi_{W_h}z-z_h\|_{L^2(\Omega)}\leqslant\gamma^{-1}\|z-z_h\|_{L^2(\Omega)}+\gamma^{-1}\|z-\Pi_{W_h}z\|_{L^2(\Omega)},
$$

so

$$
\begin{aligned}
\|f-f_h\|_{L^2(\Omega)}&\leqslant\|f-\Pi_{W_h}f\|_{L^2(\Omega)}+\|\Pi_{W_h}f-f_h\|_{L^2(\Omega)}\\
&\leqslant\|f-\Pi_{W_h}f\|_{L^2(\Omega)}+\gamma^{-1}\|z-z_h\|_{L^2(\Omega)}+\gamma^{-1}\|z-\Pi_{W_h}z\|_{L^2(\Omega)}\\
&\lesssim h^{l_0+1}|f|_{H^{l_0+1}}+C(u,f,z)h^{\min\{k_1+1,k_2+1,l_0+2\}}+C(z)h^{\min\left\{\tilde{k}_2+1,l+1\right\}}\\
&\lesssim C(u,f,z)h^{\min\{k_1+1,k_2+1,l_0+1\}}.
\end{aligned}
$$

So far, we have obtained the improved estimates,

$$
\|u-u_h\|_{L^2(\Omega)}+\|z-z_h\|_{L^2(\Omega)}\lesssim C(u,f,z)h^{\min\{k_1+1,k_2+1,l_0+2\}},\tag{11a}
$$

$$
|u-u_h|_{H^1(\Omega)}+|z-z_h|_{H^1(\Omega)}\lesssim C(u,f,z)h^{\min\{k_1,k_2,l_0+1\}},\tag{11b}
$$

$$
\|f-f_h\|_{L^2(\Omega)}\lesssim C(u,f,z)h^{\min\{k_1+1,k_2+1,l_0+1\}}.\tag{11c}
$$

3.4 Numerical methods

The matrix problem corresponding to the discrete problem (8) has the double saddle point form

$$
\boldsymbol{K}_h\begin{pmatrix}[u_h]\\ [f_h]\\ [z_h]\end{pmatrix}=\boldsymbol{F}_h,\tag{12}
$$

where

$$
\boldsymbol{K}_h=\begin{pmatrix}\boldsymbol{A}_h & \boldsymbol{0} & \boldsymbol{B}_h^T\\ \boldsymbol{0} & \gamma\boldsymbol{C}_h & -\boldsymbol{D}_h^T\\ \boldsymbol{B}_h & -\boldsymbol{D}_h & \boldsymbol{0}\end{pmatrix}.
$$

Here, $\boldsymbol{A}_h$ is merely symmetric positive semi-definite since $\Omega_0\subset\Omega$, and possibly $\boldsymbol{C}_h\neq\boldsymbol{D}_h$ because $W_h\neq Q_h$, so a lot of existing solvers and preconditioners in the literature for this problem cannot be used.

According to the discrete well-posedness result, we choose $V_h=Q_h$ for the discrete problems. For simplicity, we choose

$$
\begin{aligned}
V_h&=Q_h=\mathcal{P}^k(\mathcal{T}_h)\cap H^1_0(\Omega),\\
W_h&=\mathcal{P}^l(\mathcal{T}_h)\cap L^2(\Omega).
\end{aligned}
$$

as our finite element spaces, where $k \in \mathbb{N}_+$ and $l \in \mathbb{N}$.

In this setting, $\boldsymbol{B}_h$ is the stiffness matrix on $V_h \times V_h$, which is symmetric positive-definite and easy to invert using the preconditioned gradient descent method, and $\boldsymbol{C}_h$ is block-diagonal thus admitting an explicit inverse. Now

$$\boldsymbol{P}_h \boldsymbol{K}_h \boldsymbol{P}_h^T = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \boldsymbol{B}_h^T \\ \mathbf{0} & \gamma \boldsymbol{C}_h + \boldsymbol{D}_h^T \boldsymbol{B}_h^{-T} \boldsymbol{A}_h \boldsymbol{B}_h^{-1} \boldsymbol{D}_h & \mathbf{0} \\ \boldsymbol{B}_h & \mathbf{0} & \mathbf{0} \end{pmatrix},$$

where

$$\boldsymbol{P}_h = \begin{pmatrix} \boldsymbol{I} & \mathbf{0} & -\frac{1}{2} \boldsymbol{A}_h \boldsymbol{B}_h^{-1} \\ \boldsymbol{D}_h^T \boldsymbol{B}_h^{-T} & \boldsymbol{I} & -\boldsymbol{D}_h^T \boldsymbol{B}_h^{-T} \boldsymbol{A}_h \boldsymbol{B}_h^{-1} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I} \end{pmatrix}.$$

Hence the system matrix in (12) is symmetric and indefinite. We use the MINRES solver with a simple block-diagonal preconditioner

$$\boldsymbol{M}_h = \operatorname{diag}\left(\boldsymbol{B}_h, \gamma \boldsymbol{C}_h, \boldsymbol{B}_h\right).$$

4 Numerical examples

4.1 Example 1

We consider the following exact solution on $\Omega_0 = \Omega = [0, \pi]^2$.

$$\begin{aligned} u(x, y) &= \sin(k_1 x) \sin(k_2 y), \\ f(x, y) &= \left(k_1^2 + k_2^2\right) u(x, y), \\ z(x, y) &= \gamma f(x, y), \\ u_0(x, y) &= \left(\gamma (k_1^2 + k_2^2)^2 + 1\right) u(x, y). \end{aligned}$$

Here, $k_1, k_2 \in \mathbb{N}_+$ and $\gamma > 0$ can be arbitrary.

In our tests, we choose the parameters $k_1 = k_2 = 1$ and $\gamma = 1$. In order to verify our improved estimates (11), we focus on two combinations of polynomial orders, $k = l \in \mathbb{N}_+$ and $k = l + 1 \in \mathbb{N}_+$. The coarsest mesh is shown in Figure 1. The convergence history for all the variables in these two cases are shown in Tables 1 to 5 and Tables 6 to 10 respectively.

We see from the results that the convergence rates for both u and z in both L^2- and H^1-norms are optimal for $k = l$ or $k = l + 1$, and the convergence rate for f in L^2-norm is $\min\{k + 1, l + 1\}$. This agrees with our improved estimates (11) well. Moreover, the errors of both u and z are almost identical in these two cases.

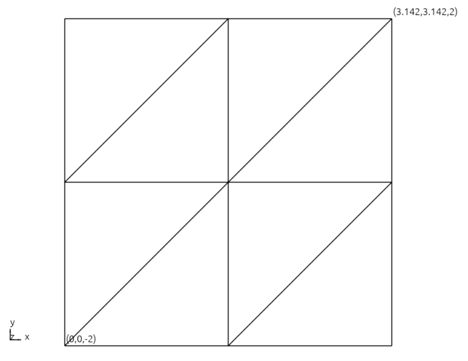


Figure 1: Coarsest mesh

Table 1: Convergence history of $\|u - u_h\|_{L^2}$ with $k = l$

refine	$k = 1$	order	$k = 2$	order	$k = 3$	order
0	9.672e-01	-	1.214e-01	-	1.634e-02	-
1	3.536e-01	1.45	1.370e-02	3.15	8.958e-04	4.19
2	9.766e-02	1.86	1.567e-03	3.13	5.100e-05	4.13
3	2.505e-02	1.96	1.896e-04	3.05	3.051e-06	4.06
4	6.302e-03	1.99	2.349e-05	3.01	5.514e-07	2.47
5	1.578e-03	2.00	2.944e-06	3.00	5.833e-07	-0.08

Table 2: Convergence history of $|u - u_h|_{H^1}$ with $k = l$

refine	$k = 1$	order	$k = 2$	order	$k = 3$	order
0	1.579e+00	-	4.712e-01	-	1.016e-01	-
1	8.607e-01	0.88	1.296e-01	1.86	1.324e-02	2.94
2	4.351e-01	0.98	3.340e-02	1.96	1.655e-03	3.00
3	2.180e-01	1.00	8.420e-03	1.99	2.060e-04	3.01
4	1.090e-01	1.00	2.110e-03	2.00	2.576e-05	3.00
5	5.452e-02	1.00	5.277e-04	2.00	4.189e-06	2.62

Table 3: Convergence history of $\|f - f_h\|_{L^2}$ with $k = l$

refine	$k = 1$	order	$k = 2$	order	$k = 3$	order
0	1.282e+00	-	1.489e-01	-	2.925e-02	-
1	3.482e-01	1.88	2.215e-02	2.75	1.753e-03	4.06
2	8.721e-02	2.00	2.932e-03	2.92	1.016e-04	4.11
3	2.177e-02	2.00	3.724e-04	2.98	6.257e-06	4.02
4	5.440e-03	2.00	4.682e-05	2.99	1.616e-06	1.95
5	1.359e-03	2.00	5.996e-06	2.97	1.658e-06	-0.04

Table 4: Convergence history of $\|z - z_h\|_{L^2}$ with $k = l$

refine	$k = 1$	order	$k = 2$	order	$k = 3$	order
0	1.282e+00	-	1.489e-01	-	2.925e-02	-
1	3.482e-01	1.88	2.215e-02	2.75	1.754e-03	4.06
2	8.721e-02	2.00	2.932e-03	2.92	1.015e-04	4.11
3	2.177e-02	2.00	3.724e-04	2.98	6.082e-06	4.06
4	5.440e-03	2.00	4.677e-05	2.99	1.635e-06	1.90
5	1.360e-03	2.00	5.859e-06	3.00	1.701e-06	-0.06

Table 5: Convergence history of $|z - z_h|_{H^1}$ with $k = l$

refine	$k = 1$	order	$k = 2$	order	$k = 3$	order
0	3.072e+00	-	9.391e-01	-	2.030e-01	-
1	1.696e+00	0.86	2.592e-01	1.86	2.647e-02	2.94
2	8.666e-01	0.97	6.680e-02	1.96	3.310e-03	3.00
3	4.355e-01	0.99	1.684e-02	1.99	4.121e-04	3.01
4	2.180e-01	1.00	4.219e-03	2.00	5.156e-05	3.00
5	1.090e-01	1.00	1.055e-03	2.00	7.958e-06	2.70

Table 6: Convergence history of $\|u - u_h\|_{L^2}$ with $k = l + 1$

refine	$k = 1$	order	$k = 2$	order	$k = 3$	order
0	1.151e+00	-	1.407e-01	-	1.715e-02	-
1	4.284e-01	1.43	1.491e-02	3.24	9.194e-04	4.22
2	1.183e-01	1.86	1.619e-03	3.20	5.151e-05	4.16
3	3.034e-02	1.96	1.913e-04	3.08	3.060e-06	4.07
4	7.632e-03	1.99	2.355e-05	3.02	5.514e-07	2.47
5	1.911e-03	2.00	2.946e-06	3.00	5.833e-07	-0.08

Table 7: Convergence history of $|u - u_h|_{H^1}$ with $k = l + 1$

refine	$k = 1$	order	$k = 2$	order	$k = 3$	order
0	1.721e+00	-	4.775e-01	-	1.023e-01	-
1	8.939e-01	0.95	1.300e-01	1.88	1.326e-02	2.95
2	4.398e-01	1.02	3.342e-02	1.96	1.656e-03	3.00
3	2.186e-01	1.01	8.421e-03	1.99	2.061e-04	3.01
4	1.091e-01	1.00	2.110e-03	2.00	2.576e-05	3.00
5	5.453e-02	1.00	5.277e-04	2.00	4.189e-06	2.62

Table 8: Convergence history of $\|f - f_h\|_{L^2}$ with $k = l + 1$

refine	$k = 1$	order	$k = 2$	order	$k = 3$	order
0	1.934e+00	-	4.793e-01	-	1.068e-01	-
1	8.684e-01	1.16	1.244e-01	1.95	1.373e-02	2.96
2	4.169e-01	1.06	3.125e-02	1.99	1.730e-03	2.99
3	2.063e-01	1.01	7.817e-03	2.00	2.167e-04	3.00
4	1.029e-01	1.00	1.954e-03	2.00	2.715e-05	3.00
5	5.142e-02	1.00	4.886e-04	2.00	3.772e-06	2.85

Table 9: Convergence history of $\|z - z_h\|_{L^2}$ with $k = l + 1$

refine	$k = 1$	order	$k = 2$	order	$k = 3$	order
0	1.239e+00	-	1.437e-01	-	2.918e-02	-
1	3.217e-01	1.95	2.190e-02	2.71	1.753e-03	4.06
2	7.894e-02	2.03	2.923e-03	2.91	1.016e-04	4.11
3	1.959e-02	2.01	3.721e-04	2.97	6.083e-06	4.06
4	4.887e-03	2.00	4.677e-05	2.99	1.635e-06	1.90
5	1.221e-03	2.00	5.859e-06	3.00	1.701e-06	-0.06

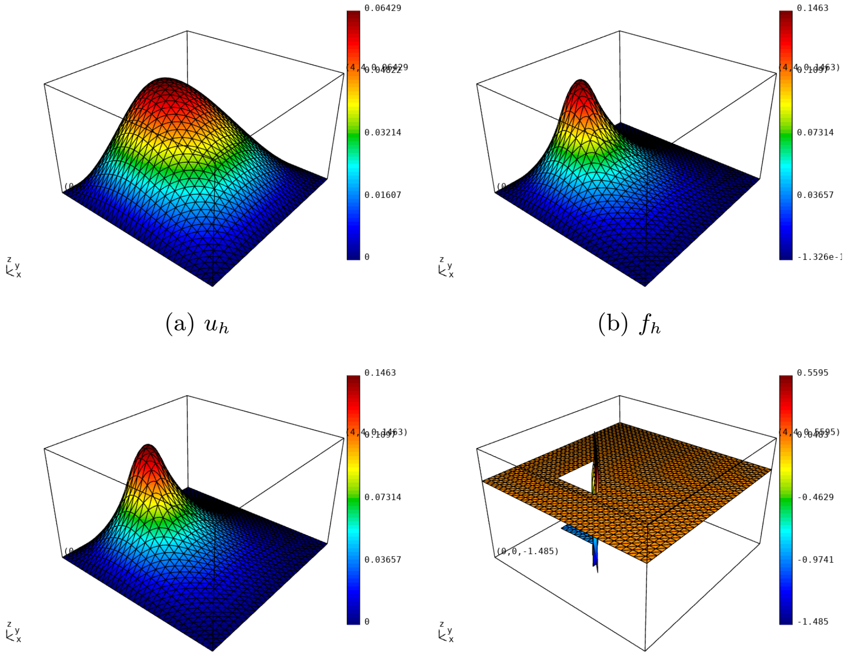
Table 10: Convergence history of $|z - z_h|_{H^1}$ with $k = l + 1$

refine	$k = 1$	order	$k = 2$	order	$k = 3$	order
0	3.094e+00	-	9.405e-01	-	2.030e-01	-
1	1.703e+00	0.86	2.592e-01	1.86	2.647e-02	2.94
2	8.678e-01	0.97	6.680e-02	1.96	3.310e-03	3.00
3	4.356e-01	0.99	1.684e-02	1.99	4.121e-04	3.01
4	2.180e-01	1.00	4.219e-03	2.00	5.156e-05	3.00
5	1.090e-01	1.00	1.055e-03	2.00	7.958e-06	2.70

4.2 Example 2

We consider the optimal temperature control of a strict subdomain, where $\Omega = [0, 4]^2$ and $\Omega_0 = \{(x, y) : 0 \leqslant (x-1), (y-1) \leqslant (x-1) + (y-1) \leqslant 1\}$. The target temperature on Ω_0 is set to $u_0(x, y) =$

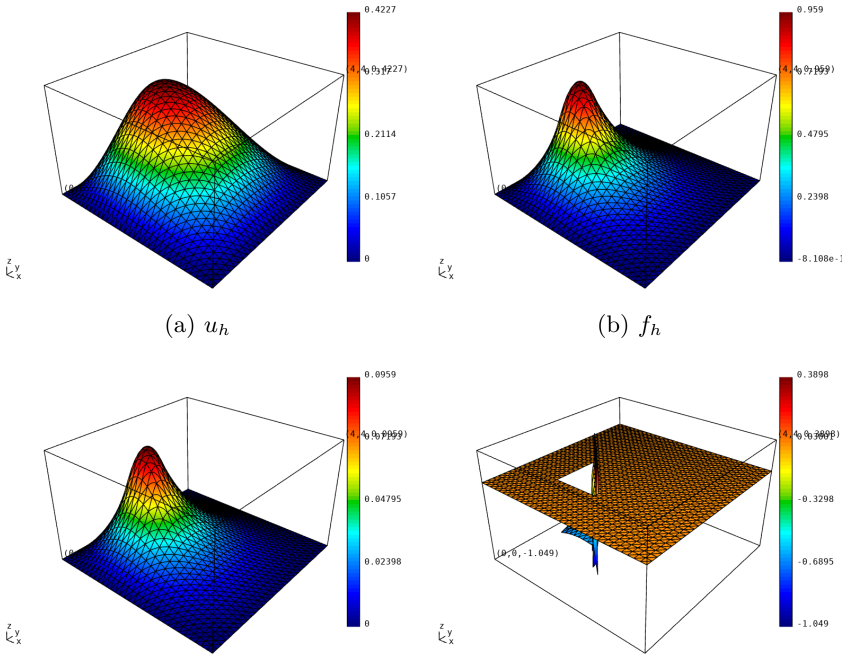
1. We present the numerical results with weighting parameter $\gamma = 10^{-0}, 10^{-1}, 10^{-2}$ in Figures 2 to 4 on the same mesh, with polynomial orders $k = l = 2$. It is evident that the smaller γ is, the better u approximates u_0 on Ω_0.



(a) u_h (b) f_h

(c) z_h (d) $(u_h - u_0)1_{\Omega_0}$

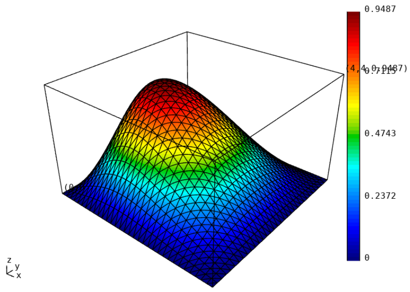
Figure 2: Numerical solution with $\gamma = 10^{-0}$



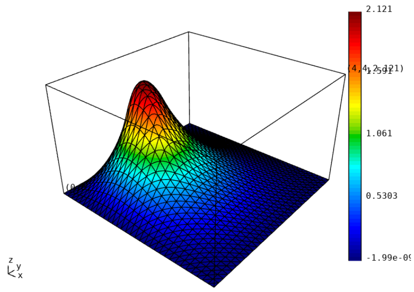
(a) u_h (b) f_h

(c) z_h (d) $(u_h - u_0)1_{\Omega_0}$

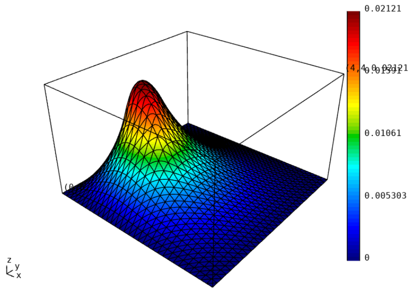
Figure 3: Numerical solution with $\gamma = 10^{-1}$



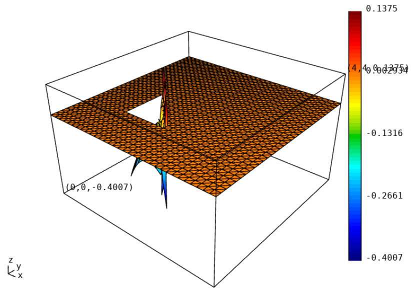
(a) u_h



(b) f_h



(c) z_h



(d) $(u_h - u_0)1_{\Omega_0}$

Figure 4: Numerical solution with $\gamma = 10^{-2}$

5 Conclusion

In this project, I studied the PDE-constrained optimal control model for static heating. The problem minimizes the L^2-error between the unknown u and target temperature u_0 on the subdomain $\Omega_0 \subset \Omega$ regularized by the control cost, which is modeled as the L^2-norm of f, under the static heat equation constraint. First, I derived the weak form of the optimality condition and proved its well-posedness. Then, I discretized it using a conforming Galerkin method, and proved the discrete well-posedness for a special combination of finite element spaces. Later, I proved then a priori error estimates in the energy norm, and improved it using the Oden–Nitsche technique. Finally, the validity of this numerical discretization is displayed by several examples.

References

[1] Roland Becker, Hartmut Kapp, and Rolf Rannacher. "Adaptive Finite Element Methods for Optimal Control of Partial Differential Equations: Basic Concept". In: *SIAM Journal on Control and Optimization* 39.1 (2000), pp. 113–132. DOI: 10.1137/S0363012999351097.

[2] Alexandre Ern and Jean-Luc Guermond. *Finite Elements I. Approximation and Interpolation*. Vol. 72. Texts in Applied Mathematics. Cham: Springer International Publishing, Feb. 2021. ISBN: 978-3-030-56341-7. DOI: 10.1007/978-3-030-56341-7.

[3] Alexandre Ern and Jean-Luc Guermond. *Finite Elements II. Galerkin Approximation, Elliptic and Mixed PDEs*. Vol. 73. Texts in Applied Mathematics. Cham: Springer International Publishing, Apr. 2021. ISBN: 978-3-030-56923-5. DOI: 10.1007/978-3-030-56923-5.

[4] Lawrence Craig Evans. *Partial Differential Equations*. 2nd ed. Vol. 19. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2010. ISBN: 978-0-8218-4974-3.

[5] Wei Gong and Ningning Yan. "Adaptive finite element method for elliptic optimal control problems: convergence and optimality". In: *Numerische Mathematik* 135.4 (Apr. 2017), pp. 1121–1170. ISSN: 0945-3245. DOI: 10.1007/s00211-016-0827-9.

[6] Fredi Tröltzsch. *Optimal Control of Partial Differential Equations. Theory, Methods and Applications*. Vol. 112. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2010. ISBN: 978-0-8218-4904-0. DOI: 10.1090/gsm/112.