SpezialForschungsBereich F 32

Karl–Franzens Universität Graz
Technische Universität Graz
Medizinische Universität Graz

OPTIMAL CONTROL FOR THE ELLIPTIC SYSTEM WITH POLYGONAL STATE CONSTRAINTS

K. Kunisch   L. Kewei   L. Xiliang

SFB-Report No. 2009–039   August 2009

A–8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the
Austrian Science Fund (FWF)

SFB sponsors:

- **Austrian Science Fund** (FWF)
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**

Das Land Steiermark

Stadt GRAZ Wissenschaft

OPTIMAL CONTROL FOR THE ELLIPTIC SYSTEM WITH POLYGONAL STATE CONSTRAINTS

KUNISCH KARL, KEWEI LIANG, AND XILIANG LU

ABSTRACT. This work is devoted to stationary optimal control problems with polygonal constraints on the components of the state. Existence of Lagrange multipliers, of different regularity, is verified for the cases with and without Slater condition holding. For the numerical realization a semi-smooth Newton method is proposed for an appropriately chosen family of regularized problems. The asymptotic behavior of the regularized problem class is studied, and numerical feasibility of the method is shown.

1. INTRODUCTION

In recent years a significant amount of attention was paid to open loop optimal control problems governed by partial differential equations with state and/or control constraints, see for example [3, 4, 6, 8, 10, 11, 12] and further references cited there. Most of this work, especially in the context of efficient numerical methods, considered the case of so-called simple or box constraints. In contrast, the focus in this paper is on polygonal constraints. Thus we consider a system of second order elliptic differential equations

$$\Lambda\vec{y} = \vec{u}, \text{ in } \Omega, \quad \vec{y}|_{\partial\Omega} = \vec{0}. \tag{1.1}$$

with vector-valued state-variable $\vec{y}$ and control variable $\vec{u}$. For simplicity here we consider only stationary problems posed on a bounded domain $\Omega \subset \mathbb{R}^n$ $(n = 2, 3)$ with Lipschitz continuous boundary $\partial\Omega$. As cost we take the quadratic functional

$$J(\vec{y}, \vec{u}) = \frac{1}{2}\|\vec{y} - \vec{y}_d\|^2 + \frac{\alpha}{2}\|\vec{u}\|^2.$$

where $\alpha > 0$ and $\vec{y}_d$ are fixed. Consider then the following optimal control problem:

Problem 1.1.

$$\min J(\vec{y}, \vec{u}), \quad \text{such that (1.1) holds and } M\vec{y}(\mathbf{x}) \le \vec{\psi}, \text{ for all } \mathbf{x} \in \Omega$$

.

Here $\vec{\psi} \in \mathbb{R}^m$, M is matrix of appropriate dimension and $M\vec{y}(\mathbf{x}) \le \vec{\psi}$ describes a point-wise polygonal bound that needs to be satisfied by the state $\vec{y}$ of the system (1.1).

Date: May 5, 2009.

Key words and phrases. optimal control, state constraints, semi-smooth Newton method, polygonal constraints, Salter condition.

Work of K.K. in part supported by Austrian Science Foundation under SFB032, "Mathematical Optimization and Applications in Biomedical Sciences".

To solve Problem 1.1 a Lagrangian approach will be used. It is by now well-known that the Lagrange multiplier corresponding to the state constraint is only a measure, see [3, 4]. Therefore a penalized formulation is introduced. Specifically, for $\gamma > 0$ we consider the family of regularized problems

Problem 1.2.

$$\min J(\vec{y}, \vec{u}) + \frac{\gamma}{2}\|(M\vec{y} - \vec{\psi})^+\|^2 \quad \text{such that (1.1) holds,}$$

where $(M\vec{y} - \vec{\psi})^+$ is defined coordinate-wise: $(M\vec{y} - \vec{\psi})_i^+ = \max(M_i^T\vec{y} - \psi_i, 0)$, for $i = 1, ..., m$.

Throughout $\|\cdot\|$ denotes the L^2 on Ω of appropriate dimensions and $\|\cdot\|_2$ will stand for the norm in H^2.

There are three instances which involve a dimension concept in this paper. First, there is the spatial domain Ω and we assume that it is of dimension 2 or 3. The dimensional of the system in particular the dimension of $\vec{y}$ is chosen to be 2. This is mainly for transparency of the proofs and higher dimensions can be treated by the same techniques. Finally there is the image space of M, which we assume to be of dimension m. This is the number of constraints on the state $\vec{y}$. Throughout it will be essential that the regularity and dimension assumptions are such that for the state of the system we have $\vec{y} \in C(\Omega, \mathbb{R}^2)$ for any admissible control $\vec{u}$.

2. Preliminaries and Problem Setting

2.1. **Problem Setting.** We define the spaces $W = H^2 \cap H_0^1$ and $\overrightarrow{W} = (H^2 \cap H_0^1)^2$. Let $C(\bar{\Omega})$ be the space of continuous functions on $\bar{\Omega}$, endowed with the maximum norm and denote by $C_0(\Omega)$ the subspace of $C(\bar{\Omega})$ with vanishing trace on the boundary. It is known that $\overrightarrow{W} \hookrightarrow (C_0(\Omega))^2 = C_0(\Omega, \mathbb{R}^2)$ for $1 \le n \le 3$.

We assume that for every $\vec{u} \in L^2(\Omega, \mathbb{R}^2)$ system (1.1) has a unique solution $\vec{y} \in \overrightarrow{W}$ satisfying

$$\|\vec{y}\|_2 \le C\|\vec{u}\|. \tag{2.1}$$

We also consider a polygonal domain $D \subset \mathbb{R}^2$ which is defined as the intersection of m half spaces $M_i^T\vec{y} \le \psi_i$, $i = 1, ..., m$. Define the matrix

$$M : \mathbb{R}^2 \to \mathbb{R}^m$$

and the vector $\psi \in \mathbb{R}^m$ by

$$M = \begin{pmatrix} M_1^T \\ \vdots \\ M_m^T \end{pmatrix}, \quad \vec{\psi} = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_m \end{pmatrix}, \tag{2.2}$$

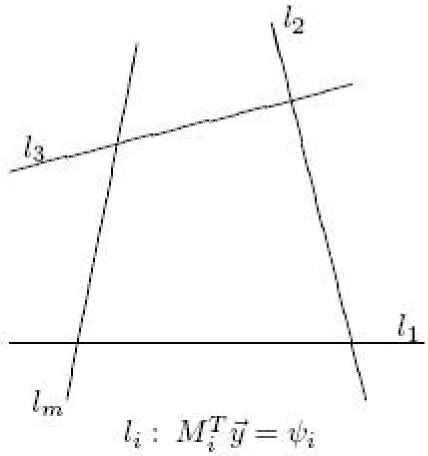
the polygon can be expressed as $D = \{\vec{y} : M\vec{y} \le \vec{\psi}\}$, see Figure 2.1.

Remark 2.1. Note that the dimensions of the composite mapping $\mathbf{x} \to M\vec{y}(\mathbf{x})$ is given by

$$\mathbf{x} \in \Omega \in \mathbb{R}^{2(3)} \to \vec{y}(\mathbf{x}) \in \mathbb{R}^2 \to M\vec{y}(\mathbf{x}) \in \mathbb{R}^m.$$

As already remarked, considering $\vec{y}(\mathbf{x}) \in \mathbb{R}^2$ simplifies the explanation, but it is not essential. The results can be extended to $\vec{y}(\mathbf{x}) \in \mathbb{R}^l$, $l > 2$.

FIGURE 2.1. Polygonal Constraint



2.2. **Measure Theory.** We review some basic results in measure theory. By Riesz's representation theorem, the regular Borel measures on the compact set $\overline{\Omega}$ can be identified with the dual space of $C(\overline{\Omega})$ (see [5]). An analogous result holds for vector valued regular Borel measures. We shall not distinguish in notation these two concepts, i.e. we use $\lambda(U)$ to denote the measure of a Borel measurable set U and we use the duality pairing $\langle \lambda, f \rangle_{C^*,C}$ for $f \in C(\overline{\Omega})$ to represent a measure. The positivity of a measure can equivalently be checked by following relation:

$$\lambda \geq 0 \Leftrightarrow \langle \lambda, f \rangle_{C^*,C} \geq 0, \quad \text{for all } f \in C(\bar{\Omega}), \text{ with } f \geq 0.$$

Given $\lambda \in C^*$, the measure can be defined as follows (see [5]). For any open subset $U \in \Omega$,

$$\lambda(U) = \sup_{f \in C_0(\Omega), 0 \leq f \leq \chi_U} \langle \lambda, f \rangle_{C^*,C}. \tag{2.3}$$

Subsequently the property of the regular Borel measure allows us to extend the above calculation to any Borel measurable set A:

$$\lambda(A) = \inf_{A \subset U, U \text{ is open}} \lambda(U).$$

The following relation can be obtained by the above characterization (see [5]). Suppose that f is a continuous function, then

$$\begin{aligned} &\chi_A \leq f \Rightarrow \lambda(A) \leq \langle \lambda, f \rangle_{C^*,C}, \text{ for all measurable sets } A, \\ &0 \leq f \leq \chi_A \Rightarrow \lambda(A) \geq \langle \lambda, f \rangle_{C^*,C}, \text{ for all compact sets } A. \end{aligned} \tag{2.4}$$

3. FIRST ORDER OPTIMALITY CONDITION

Let K be the following set

$$K = \{\vec{z} \in C(\overline{\Omega}, \mathbb{R}^m) : \vec{z}(\mathbf{x}) \leq \vec{\psi}, \text{ for all } \mathbf{x} \in \overline{\Omega}\}, \tag{3.1}$$

and let I_K be the indicator functional of K:

$$I_K(\vec{z}) = \begin{cases} +\infty, & \vec{z} \notin K \\ 0, & \vec{z} \in K. \end{cases} \tag{3.2}$$

It is clear that K is a closed convex set in $C(\overline{\Omega}, \mathbb{R}^m)$ and therefore the indicator functional I_K is convex.

We denote the map $\vec{u} \to \vec{y}(\vec{u})$ by T, where $T : L^2(\Omega, \mathbb{R}^2) \to C_0(\Omega, \mathbb{R}^2)$. Hence cost functional $J(\vec{y}, \vec{u})$ can be equivalently represented by the reduced functional

$$\hat{J}(\vec{u}) = J(T\vec{u}, \vec{u}),$$

and 1.1 can be rewritten in the following equivalent way:

$$\inf_{\vec{u} \in L^2(\Omega, \mathbb{R}^2)} \hat{J}(\vec{u}) + I_K(MT\vec{u}). \tag{3.3}$$

Theorem 3.1. *There exists a unique solution for Problem 1.1.*

Proof: We can easily check that $\hat{J} + I_K \circ MT$ is a lower semi-continuous convex functional. Then standard arguments imply the existence of an optimal solution. Strictly convexity implies uniqueness of the solution. □

3.1. **Optimality System.** Using the notation of subdifferential calculus (ref. [2]), $\vec{u}^*$ is a solution to (3.3) if and only if

$$0 \in \partial(\hat{J}(\vec{u}^*) + I_k(MT\vec{u}^*)).$$

Assumption 3.2. $\vec{\psi} > 0$.

With Assumption 3.2 holding the Slater condition is satisfied, i.e.:

$$\text{there exists } \vec{u}, \quad \text{such that} \quad MT\vec{u} \in \text{int}(K).$$

In fact, for the choice $\vec{u} = 0$, we have $\vec{y} = T\vec{u} = 0$, and hence $M\vec{y} = 0$, is an interior point of K.

Remark 3.1. The assumption $\vec{\psi} > 0$ ensures that the Salter condition is satisfied. It implies that $\mathbf{0}$ is the interior point of the polygon $D \subset \mathbb{R}^m$. In practice this assumption may not be true, e.g. $\mathbf{0}$ is a vertex of D in our numerical example in Section 5. The following convex analysis approach relies on the Slater condition and thus it is not applicable to the case where we only have $\vec{\psi} \geq 0$. For this case, however, we can still study the existence of Lagrange multiplier, a in a slightly weaker sense, namely with $\vec{\lambda}^* \in W^*(\Omega, \mathbb{R}^m)$. This will be considered in Proposition 3.4.

If assumption 3.2 is satisfied, then (ref. [9])

$$0 \in \partial\hat{J}(\vec{u}^*) + T^* M^T \partial I_k(MT\vec{u}^*).$$

Hence we can find $\vec{\lambda}^* \in \partial I_k(MT\vec{u}^*)$ such that $\vec{\lambda}^* \in C^*(\overline{\Omega}, \mathbb{R}^m)$ and

$$0 \in \partial\hat{J}(\vec{u}^*) + T^* M^T \vec{\lambda}^*.$$

Since $\hat{J}(\vec{u}^*) = \frac{1}{2}\|T\vec{u}^* - \vec{y}_d\|^2 + \frac{\alpha}{2}\|\vec{u}^*\|^2$, we have

(3.4)

$$0 = (T^*(T\vec{u}^* - \vec{y}_d) + \alpha\vec{u}^* + T^* M^T \vec{\lambda}^*, \vec{v}) = (\alpha\vec{u}^* - \vec{p}^*, \vec{v}), \quad \text{for all } \vec{v} \in L^2(\Omega, \mathbb{R}^2),$$

where the adjoint variable $\vec{p}^*$ is given by $\vec{p}^* = -T^*(T\vec{u}^* - \vec{y}_d) - T^*M^T\vec{\lambda}^*$. Hence

$$
\begin{aligned}
(\vec{p}^*, \vec{v}) &= (-T^*M^T\vec{\lambda}^* - T^*(T\vec{u}^* - \vec{y}_d), \vec{v}), \quad \text{for all } \vec{v} \in L^2(\Omega, \mathbb{R}^2), \\
(\vec{p}^*, \Lambda\vec{z}) &= -\langle \vec{\lambda}^*, M\vec{z}\rangle_{C^*,C} - (\vec{y}^* - \vec{y}_d, \vec{z}), \quad \text{for all } \vec{z} \in \overrightarrow{W}.
\end{aligned}
\tag{3.5}
$$

Lastly $\vec{\lambda}^* \in \partial I_k(MT\vec{u}^*)$ is equivalent to

$$
M\vec{y}^* \in K, \quad \langle \vec{\lambda}^*, \vec{z} - M\vec{y}^*\rangle_{C^*,C} \le 0, \quad \text{for all } \vec{z} \in K.
$$

By a standard argument, this is also equivalent to

$$
\vec{\lambda}^* \ge 0, \quad M\vec{y}^* \le \vec{\psi}, \quad \langle \vec{\lambda}^*, M\vec{y}^* - \vec{\psi}\rangle_{C^*,C} = 0. \tag{3.6}
$$

Combining these arguments we have

Theorem 3.3. *If assumption 3.2 is satisfied, then there exists an optimal solution $\vec{u}^*$, $\vec{y}^*$ and an associated dual variable $\vec{p}^* \in L^2(\Omega, \mathbb{R}^2)$, and a Lagrange multiplier $\vec{\lambda}^* \in C^*(\overline{\Omega}, \mathbb{R}^m)$, such that the first order optimality system*

$$
\begin{cases}
\Lambda\vec{y}^* = \vec{u}^*, \\
(\vec{p}^*, \Lambda\vec{z}) + \langle \vec{\lambda}^*, M\vec{z}\rangle_{C^*,C} = (\vec{y}_d - \vec{y}^*, \vec{z}), \quad \textit{for all } \vec{z} \in \overrightarrow{W}, \\
\alpha\vec{u}^* = \vec{p}^*, \\
\vec{\lambda}^* \ge 0, \quad M\vec{y}^* \le \vec{\psi}, \quad \langle \vec{\lambda}^*, M\vec{y}^* - \vec{\psi}\rangle_{C^*,C} = 0,
\end{cases}
\tag{3.7}
$$

holds.

Remark 3.2. The dual variable $\vec{p}^*$ has more regularity than $L^2(\Omega, \mathbb{R}^2)$. The second equation of the optimality system has the form

$$
\Lambda^T\vec{p}^* = \vec{y}_d - \vec{y}^* - M^T\vec{\lambda}^*.
$$

This is an elliptic equation where the right hand side is given by a Radon measure. From [13], we have $\vec{p}^* \in W^{1,p}(\Omega, \mathbb{R}^2)$ for any $1 \le p < 2$. By the Sobolev embedding theorem this implies that, $\vec{p}^* \in L^q(\Omega, \mathbb{R}^2)$, for all $q < \infty$. Then use of standard regularity theory for elliptic equations implies that the regularity of the state can be improved from $\vec{y}^*$ from $H^2 \cap H_0^1(\Omega, \mathbb{R}^2)$ to $W^{2,q} \cap H_0^1(\Omega, \mathbb{R}^2)$ for any $q < \infty$. Here we use the first and the third equation in (3.8).

Proposition 3.4. *Let $\vec{\psi} \ge 0$. Then there exists an optimal solution $\vec{u}^*$, $\vec{y}^*$ and an associated dual variable $\vec{p}^* \in L^2(\Omega, \mathbb{R}^2)$, and a Lagrange multiplier $\vec{\lambda}^* \in W^*(\Omega, \mathbb{R}^m)$, which satisfy the first order optimality system*

$$
\begin{cases}
\Lambda\vec{y}^* = \vec{u}^*, \\
(\vec{p}^*, \Lambda\vec{z}) + \langle \vec{\lambda}^*, M\vec{z}\rangle_{W^*,W} = (\vec{y}_d - \vec{y}^*, \vec{z}), \quad \textit{for all } \vec{z} \in \overrightarrow{W}, \\
\alpha\vec{u}^* = \vec{p}^*, \\
\vec{\lambda}^* \ge 0, \quad M\vec{y}^* \le \vec{\psi}, \quad \langle \vec{\lambda}^*, M\vec{y}^* - \vec{\psi}\rangle_{W^*,W} = 0.
\end{cases}
\tag{3.8}
$$

The proof is given in the Appendix. It depends on a regularization technique which is developed in Section 4. The assumption $\vec{\psi} > 0$ is not essential in most part of this paper, which include the convergence of the solution to Problem 1.2, super-linear convergence of the semi-smooth method. But we need to assume $\vec{\psi} > 0$ in Section 3.2, to ensure the uniqueness of Lagrange multiplier.

3.2. **Uniqueness of Optimal Solution.** From Theorem 3.1, we know that the optimal solution pair $(\vec{u}^*, \vec{y}^*)$ is unique, and hence $\vec{p}^*$ is unique. The main task is to prove the uniqueness of the Lagrange multiplier $\vec{\lambda}^*$. We already know that the Lagrange multiplier $\vec{\lambda}^* \in C^*(\overline{\Omega}, \mathbb{R}^m)$, and we henceforth investigate some of its properties. Firstly it is clear that the linear complementary condition (3.6) holds componentwise, i.e.

$$\lambda_i^* \geq 0, \quad M_i^T \vec{y}^* \leq \psi_i, \quad \langle \lambda_i^*, M_i^T \vec{y}^* - \psi_i \rangle_{C^*,C} = 0, \; i = 1, ..., m. \tag{3.9}$$

Lemma 3.5. *There exists a decomposition Ω_i, $i = 1, ..., m$ of Ω, which satisfies*

$$\cup_i \Omega_i = \Omega, \; \Omega_i \textit{ is open}, \; \lambda_j^*(\Omega_i) = 0, \textit{ for all } j \neq i, i+1 \textit{ (where we let m+1 be 1)}.$$

Proof: From the definition, the function $\vec{y}^*$ maps from $\overline{\Omega}$ to the polygonal domain D. It is a continuous, hence uniformly continuous. More precisely, for any given positive constant δ, there exists another positive constant ϵ, such that

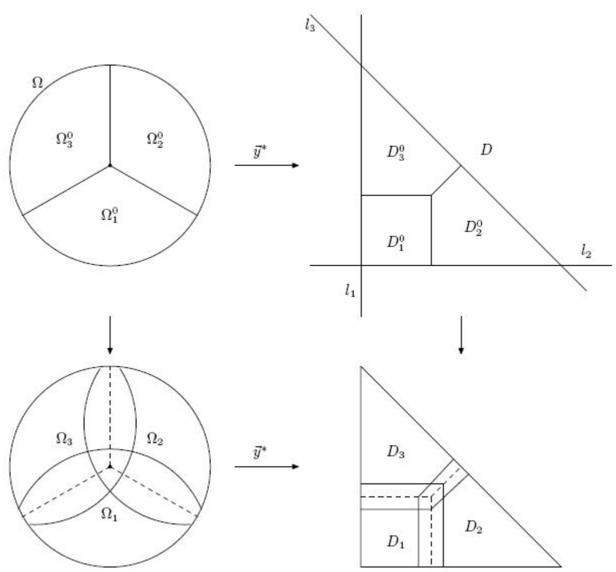
$$|\mathbf{x}_1 - \mathbf{x}_2| \leq \epsilon \Rightarrow |\vec{y}^*(\mathbf{x}_1) - \vec{y}^*(\mathbf{x}_2)| \leq \frac{\delta}{2}.$$

We can decompose the polygon D into m parts, such that every part D_i^o has at least distance $\delta > 0$ to the constraint which with index different from $i, i+1$. Then define the subdomains Ω_i^o, Ω_i and D_i as

$$\Omega_i^o = \{\mathbf{x} : \vec{y}^*(\mathbf{x}) \in D_i^o\},$$
$$\Omega_i = \{\mathbf{x} \in \Omega : \text{dist}(\mathbf{x}, \Omega_i^o) < \epsilon\},$$
$$D_i = \vec{y}^*(\Omega_i).$$

For this kind of partition, please refer to the figure 3.2.

FIGURE 3.2. Partition to the Domain and the Polygonal



Clearly Ω_i are open subsets and $\Omega = \cup_i \Omega_i$. By the above definition and from the uniform continuity of $\vec{y}^*$ on $\overline{\Omega}$, we have that every D_i has at least distance $\frac{\delta}{2}$

to the constraint whose index is not $i, i+1$. In fact, without loss of generality, we only consider the case $i = 1$. Since D_1 has strict positive distance at least $\frac{\delta}{2}$ to the constraints with index $i \neq 1, 2$, we have

$$\max_{i \neq 1,2} \sup_{\mathbf{x} \in \Omega_1} (M_i \vec{y}^* - \psi_i) \leq -\frac{\delta}{2} < 0.$$

Let $f_i = \psi_i - M_i \vec{y}^*$. Then $f_i \geq \frac{\delta}{2} \chi_{\Omega_1}$ and from (2.4) we have

$$0 \leq \lambda_i^*(\Omega_1) \leq \frac{\delta}{2} \langle \lambda_i^*, f_i \rangle_{C^*,C} = 0, \text{ for all } i > 2.$$

□

Lemma 3.6. *There exists an open set $\Omega_0 \subset \Omega$ and a function $f_0 \in C_0(\Omega)$, such that*
1. dist$(\Omega_0, \partial\Omega) > 0$, and $\vec{\lambda}^*(\overline{\Omega} \backslash \Omega_0) = 0$,
2. for all $\vec{\phi} \in C(\overline{\Omega}, \mathbb{R}^m)$, we have

$$\langle \vec{\lambda}^*, \vec{\phi} \rangle_{C^*,C} = \langle \vec{\lambda}^*, f_0 \vec{\phi} \rangle_{C^*,C}.$$

Proof: By assumption 3.2, and since $\vec{y}^*$ satisfies a homogenous boundary condition, $M\vec{y}^* - \vec{\psi}$ has strictly negative values on $\partial\Omega$, i.e. there exists a positive constant δ such that

$$M_i \vec{y}^*(\mathbf{x}) - \psi_i \leq -\delta < 0, \text{ for all } i = 1, ..., m, \text{ and } \mathbf{x} \in \partial\Omega.$$

Since $M\vec{y}^* - \vec{\psi}$ is uniformly continuous on $\overline{\Omega}$, there exists a positive constant ϵ, such that

$$|\mathbf{x}_1 - \mathbf{x}_2| \leq \epsilon \Rightarrow |g_i(\mathbf{x}_1) - g_i(\mathbf{x}_2)| \leq \frac{\delta}{2}, \text{ for all } i,$$

where $g_i(\mathbf{x}) = M_i \vec{y}^*(\mathbf{x}) - \psi_i$. We define the open set

$$\Omega_0 = \{\mathbf{x} : \mathbf{x} \in \Omega, \text{dist}(\mathbf{x}, \partial\Omega) > \epsilon\}. \tag{3.10}$$

Uniform continuity of the function $M\vec{y}^* - \vec{\psi}$ implies that

$$M_i \vec{y}^*(\mathbf{x}) - \psi_i \leq -\frac{\delta}{2} < 0, \text{ for all } i, \text{ for all } \mathbf{x} \in \overline{\Omega} \backslash \Omega_0.$$

The linear complementary condition (3.9) implies that $\lambda_i(\overline{\Omega} \backslash \Omega_0) = 0$. We define the function f_0 by

$$f_0(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in \Omega_0 \\ \frac{1}{\epsilon} \text{dist}(\mathbf{x}, \partial\Omega) & \mathbf{x} \in \overline{\Omega} \backslash \Omega_0. \end{cases}$$

Clearly f_0 is continuous and for all $\vec{\phi} \in C(\overline{\Omega}, \mathbb{R}^m)$, we have

$$\langle \vec{\lambda}^*, \vec{\phi} \rangle_{C^*,C} = \int_{\overline{\Omega}} \vec{\phi} \, d\vec{\lambda}^* = \int_{\Omega_0} \vec{\phi} \, d\vec{\lambda}^* = \int_{\overline{\Omega}} f_0 \vec{\phi} \, d\vec{\lambda}^* = \langle \vec{\lambda}^*, f_0 \vec{\phi} \rangle_{C^*,C}.$$

□

We are now prepared for the proof of uniqueness of the Lagrange multiplier. Suppose that there exist two measures $\vec{\lambda}^1$ and $\vec{\lambda}^2$ satisfying the optimality system. By Lemma 3.6,

$$\vec{\lambda}^1(\overline{\Omega} \backslash \Omega_0) = \vec{\lambda}^2(\overline{\Omega} \backslash \Omega_0) = \vec{0}.$$

Since $\vec{\lambda}^1$ and $\vec{\lambda}^2$ are regular Borel measure, we only need to show that they coincide on any open subset $U \subset \Omega_0$. By the next Lemma, the problem can be further reduced.

Lemma 3.7. *If two measures λ^1 and λ^2 coincide on any open subset $U \subset \Omega_i$, $i = 1, ..., m$, then $\lambda^1(V) = \lambda^2(V)$, for any open set $V \in \Omega_0$.*

Proof: The proof is based on the inclusion-exclusion principle. □

From Lemma 3.7, it is enough to show for any open subset $U \in \Omega_i$, $\vec{\lambda}^1(U) = \vec{\lambda}^2(U)$. Without loss of generality, we only consider the case $i = 1$. Lemma 3.5 implies that

$$\lambda_i^1(U) = \lambda_i^2(U) = 0, \text{ for all } i > 2. \tag{3.11}$$

For any $\vec{z} \in \overrightarrow{W}$, by (3.8), we have

$$\langle M^T \vec{\lambda}^1, \vec{z} \rangle_{C^*,C} = \langle M^T \vec{\lambda}^2, \vec{z} \rangle_{C^*,C},$$

which implies that

$$\langle M_{1,2}^T \vec{\lambda}_{1,2}^1, \vec{z} \rangle_{C^*,C} = \langle M_{1,2}^T \vec{\lambda}_{1,2}^2, \vec{z} \rangle_{C^*,C},$$

where the square matrix $M_{1,2}$ and the vectors $\vec{\lambda}_{1,2}^1$, $\vec{\lambda}_{1,2}^2$ are given by

$$M_{1,2} = \begin{pmatrix} M_1^T \\ M_2^T \end{pmatrix}, \quad \vec{\lambda}_{1,2}^1 = \begin{pmatrix} \lambda_1^1 \\ \lambda_2^1 \end{pmatrix}, \quad \vec{\lambda}_{1,2}^2 = \begin{pmatrix} \lambda_1^2 \\ \lambda_2^2 \end{pmatrix}.$$

Since the line segments 1 and 2 are not parallel (they intersector at vertex 1), the matrix $M_{1,2}$ is invertible. Since the space $\overrightarrow{W}$ is dense in $C_0(\Omega, \mathbb{R}^2)$, equation (2.3) implies that

$$M_{1,2}^T \vec{\lambda}_{1,2}^1(U) = M_{1,2}^T \vec{\lambda}_{1,2}^2(U).$$

Hence $\vec{\lambda}^1(U) = \vec{\lambda}^2(U)$ is obtained by the invertibility of the matrix $M_{1,2}$. Combining the results in this subsection, we have

Theorem 3.8. *There exists a unique solution $(\vec{y}^*, \vec{u}^*, \vec{p}^*, \vec{\lambda}^*)$ satisfying the optimality system (3.8).*

Remark 3.3. *Theorem 3.3 only gives first order necessary condition. Together with Theorem 3.8 this optimality system (3.8) provides also a sufficient condition.*

4. Semi-Smooth Newton Method

The optimal control Problem 1.2 without constraints on the state admits an optimal solution which satisfies the following first order optimality system:

$$\begin{cases} \Lambda \vec{y}_\gamma = \vec{u}_\gamma, \\ (\vec{p}_\gamma, \Lambda \vec{z}) + (\vec{\lambda}_\gamma, M\vec{z}) = (\vec{y}_d - \vec{y}_\gamma, \vec{z}), \quad \text{for all } \vec{z} \in \overrightarrow{W}, \\ \alpha \vec{u}_\gamma = \vec{p}_\gamma, \\ \vec{\lambda}_\gamma = \gamma(M\vec{y}_\gamma - \psi)^+. \end{cases} \tag{4.1}$$

We will use a semi-smooth Newton method to solve system (4.1). Before, however, we consider the asymptotic behavior as the penalty parameter γ tends to infinity.

4.1. **Convergence with respect to γ.** In this subsection, we study the convergence when the parameter $\gamma \to +\infty$.

Theorem 4.1. *Let $(\vec{u}^*, \vec{y}^*, \vec{p}^*, \vec{\lambda}^*)$ and $(\vec{u}_\gamma, \vec{y}_\gamma, \vec{p}_\gamma, \vec{\lambda}_\gamma)$ satisfy the optimality systems (3.8) and (4.1) respectively. Then we have*

$$\begin{aligned} \vec{y}_\gamma &\to \vec{y}^* \quad in\, \overrightarrow{W}, \\ \vec{p}_\gamma &\to \vec{p}^* \quad in\, L^2(\Omega, \mathbb{R}^2), \\ \vec{\lambda}_\gamma &\rightharpoonup \vec{\lambda}^* \quad in\; W^*(\Omega, \mathbb{R}^m). \end{aligned}$$

Proof: Since $\vec{\lambda}_\gamma \geq 0$ and $M\vec{y}^* - \vec{\psi} \leq 0$, it follows that

$$(\vec{\lambda}_\gamma, M(\vec{y}_\gamma - \vec{y}^*)) = (\vec{\lambda}_\gamma, M\vec{y}_\gamma - \vec{\psi}) - (\vec{\lambda}_\gamma, M\vec{y}^* - \psi) \geq \frac{1}{\gamma}\|\vec{\lambda}_\gamma\|^2. \tag{4.2}$$

From (4.1), we deduce that

$$\alpha(\Lambda\vec{y}_\gamma, \Lambda(\vec{y}_\gamma - \vec{y}^*)) = (\vec{p}_\gamma, \Lambda(\vec{y}_\gamma - \vec{y}^*)) = (\vec{y}_d - \vec{y}_\gamma, \vec{y}_\gamma - \vec{y}^*) - (\vec{\lambda}_\gamma, M(\vec{y}_\gamma - \vec{y}^*)). \tag{4.3}$$

Inequalities (4.2) and (4.3) lead to

$$\alpha\|\Lambda\vec{y}_\gamma\|^2 - \alpha(\Lambda\vec{y}_\gamma, \Lambda\vec{y}^*) \leq -\|\vec{y}_\gamma\|^2 + (\vec{y}_\gamma, \vec{y}^*) + (\vec{y}_d, \vec{y}_\gamma - \vec{y}^*) - \frac{1}{\gamma}\|\vec{\lambda}_\gamma\|^2.$$

This implies uniform boundedness of $\|\vec{y}_\gamma\|_2^2 + \frac{1}{\gamma}\|\vec{\lambda}_\gamma\|^2$ with respect to $\gamma \geq 1$. Hence $\|\vec{p}_\gamma\| + \|M^T\lambda_\gamma\|_{C^*}$ is also uniformly bounded. After passage to a subsequence,

$$\begin{aligned} \vec{y}_\gamma &\rightharpoonup \hat{y} \quad \text{in}\, \overrightarrow{W}, \\ \vec{p}_\gamma &\rightharpoonup \hat{p} \quad \text{in}\, L^2(\Omega, \mathbb{R}^2). \end{aligned}$$

From the definition $\vec{\lambda}_\gamma = \gamma(M\vec{y}_\gamma - \vec{\psi})^+$, it can be shown that

$$\|(M\vec{y}_\gamma - \vec{\psi})^+\| = \frac{1}{\gamma^2}\|\vec{\lambda}_\gamma\|^2 \to 0.$$

Since $\vec{y}_\gamma \to \hat{y}$ in $C(\overline{\Omega}, \mathbb{R}^2)$, we have $(M\vec{y}_\gamma - \vec{\psi})^+ \to (M\hat{y} - \vec{\psi})^+$ in $C(\overline{\Omega}, \mathbb{R}^m)$, and hence

$$\|(M\vec{y}_\gamma - \vec{\psi})^+\| \to \|(M\hat{y} - \vec{\psi})^+\| = 0.$$

This implies $M\hat{y} \leq \vec{\psi}$. Using (4.2) we have the following estimate:

$$\begin{aligned} \alpha\|\Lambda(\vec{y}_\gamma - \vec{y}^*)\|^2 &= (\vec{p}_\gamma - \vec{p}^*, \Lambda(\vec{y}_\gamma - \vec{y}^*)) \\ &= \langle \vec{y}_\gamma - \vec{y}^*, -\vec{y}_\gamma + \vec{y}^* + M^T(\vec{\lambda}^* - \vec{\lambda}_\gamma)\rangle_{C^*,C} \\ &= -\|\vec{y}_\gamma - \vec{y}^*\|^2 + \langle \vec{\lambda}^* - \vec{\lambda}_\gamma, M(\vec{y}_\gamma - \vec{y}^*)\rangle_{C^*,C} \\ &\leq -\|\vec{y}_\gamma - \vec{y}^*\|^2 - \frac{1}{\gamma}\|\vec{\lambda}_\gamma\|^2 + \langle \vec{\lambda}^*, M(\vec{y}_\gamma - \vec{y}^*)\rangle_{C^*,C}. \end{aligned} \tag{4.4}$$

Hence

$$\alpha\|\Lambda(\vec{y}_\gamma - \vec{y}^*)\|^2 + \|\vec{y}_\gamma - \vec{y}^*\|^2 + \frac{1}{\gamma}\|\vec{\lambda}_\gamma\|^2 \leq \langle \vec{\lambda}^*, M(\vec{y}_\gamma - \vec{y}^*)\rangle_{C^*,C} = \langle \vec{\lambda}^*, M\vec{y}_\gamma - \vec{\psi}\rangle_{C^*,C}.$$

Taking the limit $\gamma \to \infty$ and using that $M\hat{y} \le \vec{\psi}$, we have $\hat{y} = \vec{y}^*$ and

$$\begin{aligned} &\vec{y}_\gamma \to \vec{y}^* \quad \text{in}\, \overrightarrow{W},\\ &\vec{p}_\gamma \to \vec{p}^* \quad \text{in}\, (L^2)^2,\\ &\frac{1}{\gamma}\|\vec{\lambda}_\gamma\|^2 \to 0. \end{aligned}$$

Now we consider convergence of the Lagrange multipliers. Since

$$M^T(\vec{\lambda}_\gamma - \vec{\lambda}^*) = \Lambda^T(\vec{p}^* - \vec{p}_\gamma) + (\vec{y}^* - \vec{y}_\gamma),$$

the convergence results for $\vec{y}_\gamma$ and $\vec{p}_\gamma$ imply that

$$M^T\vec{\lambda}_\gamma \to M^T\vec{\lambda}^* \quad \text{in}\, \overrightarrow{W}^*. \tag{4.5}$$

Moreover, we also have

$$\begin{aligned} &\langle \vec{\lambda}_\gamma, \vec{\psi}\rangle_{C^*,C} = \langle \vec{\lambda}_\gamma, \vec{\psi} - M\vec{y}_\gamma\rangle_{C^*,C} + \langle \vec{\lambda}_\gamma, M\vec{y}_\gamma\rangle_{C^*,C}\\ &= -\frac{1}{\gamma}\|\vec{\lambda}_\gamma\|^2 + \langle M^T\vec{\lambda}_\gamma, \vec{y}_\gamma\rangle_{C^*,C} \to \langle M^T\vec{\lambda}^*, \vec{y}^*\rangle_{C^*,C} = \langle \vec{\lambda}^*, \vec{\psi}\rangle_{C^*,C}. \end{aligned}$$

Since M is not a square matrix, the investigation of convergence of the Lagrange multipliers $\vec{\lambda}_\gamma$ requires extra care. We employe the same notation for Ω_i, $i = 0, 1, ..., m$ as in Section 3.2. Recall that

$$\vec{\lambda}_\gamma = \gamma(M\vec{y}_\gamma - \vec{\psi})^+. \tag{4.6}$$

Our goal is to show $\vec{\lambda}_\gamma \rightharpoonup \vec{\lambda}^*$ in $W^*(\Omega, \mathbb{R}^m)$, i.e.,

$$(\vec{\lambda}_\gamma, \vec{\phi}) \to \langle \vec{\lambda}^*, \vec{\phi}\rangle_{W^*,W}, \qquad \text{for all } \vec{\phi} \in W(\Omega, \mathbb{R}^m).$$

Without loss of generality, we only need to show

$$(\lambda_{\gamma,1}, \phi) \to \langle \lambda_1^*, \phi\rangle_{W^*,W}, \qquad \text{for all } \phi \in W. \tag{4.7}$$

By the construction of Ω_0 it follows that $\Omega_0 \Subset \Omega$. Using a partition of unity argument (c.f. [1]) there exist smooth functions $\xi_i \in C_0^\infty(\Omega)$ with support in Ω_i, and

$$1 = \sum_{i=1}^m \xi_i(\mathbf{x}), \qquad \text{for all } \mathbf{x} \in \Omega_0.$$

Since $\vec{\lambda}_\gamma|_{\overline{\Omega}\backslash\Omega_0} = 0$ and $\vec{\lambda}^*(\overline{\Omega}\backslash\Omega_0) = 0$, we have

$$\begin{aligned} &(\lambda_{\gamma,1}, \phi) = (\lambda_{\gamma,1}, \phi)_{\Omega_0} = \sum_i (\lambda_{\gamma,1}, \phi\xi_i)_{\Omega_0} = \sum_i (\lambda_{\gamma,1}, \phi\xi_i),\\ &\langle \lambda_1^*, \phi\rangle_{W^*,W} = \langle \lambda_1^*, \phi\rangle_{C^*,C} = \int_{\Omega_0} \phi\, d\lambda_1^*\\ &= \sum_i \int_{\Omega_0} \phi\xi_i\, d\lambda_1^* = \sum_i \langle \lambda_1^*, \phi\xi_i\rangle_{C^*,C} = \sum_i \langle \lambda_1^*, \phi\xi_i\rangle_{W^*,W}. \end{aligned}$$

Hence (4.7) follows by applying the following Lemma 4.2. □

Lemma 4.2. *Let $\phi \in H^2 \cap H_0^1$, with support in Ω_i. Then we have $(\lambda_{\gamma,1}, \phi) \to \langle \lambda_1^*, \phi\rangle_{W^*,W}$.*

Proof: If $i = 2, ..., m-1$, then for sufficient large γ, we have $(\lambda_{\gamma,1}, \phi) = 0 = \langle \lambda_1^*, \phi \rangle_{W^*,W}$. Let $i = 1$ ($i = m$ can be treated similarly) and consider any $\vec{f} \in \overrightarrow{W}$ which vanishes outside of Ω_1. Using the same notation $M_{1,2}$ as in Section 3.2, and a similar arguments as in Theorem 3.8, we have $(\lambda_\gamma)_j = 0$, for all $j > 2$. Taking the inverse of $M_{1,2}$ and choosing $\vec{f} = \phi M_{1,2}^{-1}\mathbf{e}$ (where $\mathbf{e} = (1,0)^T$), the convergence result (4.5) implies that

$$(\lambda_{\gamma,1}, \phi) \to \langle \lambda_1^*, \phi \rangle_{W^*,W}.$$

□

4.2. **Algorithm.** For convenience we recall the optimality system (4.1) for Problem 1.2:

$$\begin{cases} \alpha \Lambda \vec{y}_\gamma = \vec{p}_\gamma, \\ \Lambda^T \vec{p}_\gamma + M^T \vec{\lambda}_\gamma + \vec{y}_\gamma = \vec{y}_d, \\ \vec{\lambda}_\gamma = \gamma (M\vec{y}_\gamma - \psi)^+, \end{cases} \tag{4.8}$$

where $\gamma > 0$ is fixed. The semi-smooth Newton algorithm (or primal-dual active set algorithm) to solve the optimality system (4.8) is given in Algorithm 1 (ref. [7, 8]).

Algorithm 1 Primal - Dual Active Set Algorithm

1: Set k=0, initialize $\vec{y}^0$,

2: let $A_i^k = \{\mathbf{x} : M_i \vec{y}^k > \psi_i\}$, and set

$$M^k = \begin{pmatrix} M_1 \chi_{A_1^k} \\ \vdots \\ M_m \chi_{A_m^k} \end{pmatrix}, \quad \vec{\psi}^k = \begin{pmatrix} \psi_1 \chi_{A_1^k} \\ \vdots \\ \psi_m \chi_{A_m^k} \end{pmatrix},$$

3: solve for $(\vec{y}^{k+1}, \vec{p}^{k+1})$

$$\begin{cases} \alpha \Lambda^T \vec{y}^{k+1} = \vec{p}^{k+1}, \\ \Lambda \vec{p}^{k+1} + \gamma M^T (M^k \vec{y}^{k+1} - \vec{\psi}^k) = \vec{y}_d - \vec{y}^{k+1}, \end{cases} \tag{4.9}$$

4: define $\vec{\lambda}^{k+1} = \gamma (M^k \vec{y}^{k+1} - \vec{\psi}^k)$,

5: stop or update $k = k+1$, and go to 2.

Proposition 4.3. *There exists a unique solution $(\vec{y}^{k+1}, \vec{p}^{k+1})$ for system (4.9).*

Proof: Since for fixed γ, equation (4.9) is an elliptic system, standard argument imply existence and uniqueness of the solution.□

Proposition 4.4. *If $A_i^k = A_i^{k+1}$, for $i = 1, ..., m$, then $(\vec{y}^{k+1}, \vec{p}^{k+1}, \vec{\lambda}^{k+1})$ solves system (4.8).*

Proof: The proof is based on the key observation: if $A_i^k = A_i^{k+1}$, then $M^k \vec{y}^{k+1} - \vec{\psi}^k = (M\vec{y}^{k+1} - \vec{\psi})^+$. □

By the above proposition, it is reasonable to choose $A_i^k = A_i^{k+1}$, for $i = 1, ..., m$ as stopping criterion. Numerical results show that this stopping criterion is typically reached.

4.3. **Super-linear Convergence of the Semi-Smooth Newton Method.**

Theorem 4.5. *Let $\vec{y}^0 \in L^p(\Omega, \mathbb{R}^2)$ for some $p > 2$, and $\|\vec{y}^0 - \vec{y}_\gamma\|_{L^p}$ be sufficient small. Then $(\vec{y}^k, \vec{p}^k, \vec{\lambda}^k)$ converges to $(\vec{y}_\gamma, \vec{p}_\gamma, \vec{\lambda}_\gamma)$ super-linearly.*

Proof: We define the error between the $(k+1)^{th}$ iteration and the solution to the γ problem by

$$\delta\vec{y} = \vec{y}^{k+1} - \vec{y}_\gamma, \quad \delta\vec{p} = \vec{p}^{k+1} - \vec{p}_\gamma, \quad \delta\vec{\lambda} = \vec{\lambda}^{k+1} - \vec{\lambda}_\gamma.$$

These expressions satisfy

$$\begin{cases} \alpha\Lambda\delta\vec{y} = \delta\vec{p} \\ \Lambda^T\delta\vec{p} + M^T\delta\vec{\lambda} = -\delta\vec{y} \\ \frac{1}{\gamma}\delta\vec{\lambda} = (M^k\vec{y}^{k+1} - \vec{\psi}^k) - (M\vec{y}_\gamma - \tilde{\psi})^+. \end{cases} \tag{4.10}$$

Since

$$(M^k\vec{y}^{k+1} - \vec{\psi}^k) - (M\vec{y}_\gamma - \tilde{\psi})^+ = M^k(\vec{y}^{k+1} - \vec{y}^k) + (M\vec{y}^k - \psi)^+ - (M\vec{y}_\gamma - \psi)^+,$$

we find for $i = 1, ..., m$,

$$\begin{aligned} \frac{1}{\gamma}\delta\lambda_1 &= M_1^T(\vec{y}^{k+1} - \vec{y}^k)\chi_{A_1^k} + \max(0, M_1^T\vec{y}^k - \psi_1) - \max(0, M_1^T\vec{y}_\gamma - \psi_1) \\ &= (M_1^T\delta\vec{y} + M_1^T\vec{y}_\gamma - M_1^T\vec{y}^k)\chi_{A_1^k} + \max(0, M_1^T(\vec{y}^k - \psi_1)) - \max(0, M_1^T\vec{y}_\gamma - \psi_1). \end{aligned}$$

For $g = M_1^T\vec{y}_\gamma - \psi_1$, $h = M_1^T\vec{y}^k - M_1^T\vec{y}_\gamma$, we have $g + h = M_1^T\vec{y}^k - \psi_1$. It is well known that the characteristic function $\chi_{A_1^k}$ is a Newton-derivative of $\max(0, g)$, considered as mapping from L^p to L^2, see e.g. [7]. Hence

$$\|\max(0, g + h) - \max(0, g) - h\chi_{A_1^k}\|_2 = o(\|h\|_{L^p}),$$

and

$$\delta\vec{\lambda} = \gamma M^k\delta\vec{y} + o(\|M(\vec{y}^k - \vec{y}_\gamma)\|_{L^p}).$$

Then we multiply $\Lambda\delta\vec{y}$ on the first equation of (4.10),

$$0 = \alpha\|\Lambda\delta\vec{y}\|^2 - (\delta\vec{p}, \Lambda\delta\vec{y}) = \alpha\|\Lambda\delta\vec{y}\|^2 + \|\delta\vec{y}\|^2 + (M\delta\vec{y}, \delta\vec{\lambda}).$$

Since $(M\delta\vec{y}, M^k\delta\vec{y}) = \|M^k\delta\vec{y}\|^2$, we have

$$\|\delta\vec{y}\|_2 = o(\|\vec{y}^k - \vec{y}_\gamma\|_{L^p}),$$

and hence the desired estimate follows:

$$\|\delta\vec{y}\|_2 + \|\delta\vec{\lambda}\| + \|\delta\vec{p}\|_2 = o(\|\vec{y}^k - \vec{y}_\gamma\|_{L^p}).$$

□

5. Numerical Results

Here we present a numerical example by utilizing a finite difference discretization to the following elliptic system in the unit square with the homogeneous Dirichlet boundary conditions:

$$-\triangle y_1 + y_2 = u_1, \tag{5.1}$$

$$-\triangle y_2 - y_1 = u_2. \tag{5.2}$$

Define the elliptic operator $\Lambda = -\triangle I + \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. It can be shown that Λ satisfies $\|\vec{y}\|_2 \le C\|\vec{u}\|$.

The state variable $\vec{y}$ has to satisfy the constraint

$$y_1 \ge 0, \quad y_2 \ge 0, \quad y_1 + y_2 \le 1. \tag{5.3}$$

It can be equivalently represented in matrix form as $M\vec{y} \le \vec{\psi}$, where

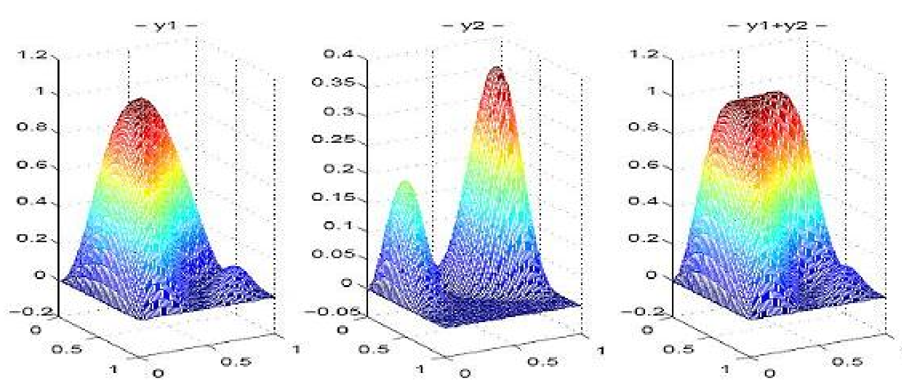
$$M = \begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 1 \end{pmatrix}, \quad \vec{\psi} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Recall that the cost functional J is

$$J(\vec{y}, \vec{u}) = \frac{1}{2}\|\vec{y} - \vec{y}_d\|^2 + \frac{\alpha}{2}\|\vec{u}\|^2.$$

Let $\vec{y}_d = (4\sin(4\pi xy), 2\sin(2\pi x) + 2\cos(2\pi y))^T$, $\gamma = 10^4$, $\alpha = 0.001$, $h = \frac{1}{64}$. For this choice of y_d all three constraints are active on some part of the domain. We refer to Figure 5.3 and 5.4 for the optimal state and control. The active sets associated with the three different constraints are depicted in Figure 5.5. We note that the active set can be a "slim set" (as for the first constraint) or a set which clearly has on open interior (as for the second constraint). The corresponding Lagrange multipliers are depicted in Figure 5.6, and as expected they are oscillatory near the boundary of active set.

Figure 5.3. Optimal State



Super-linear convergence can be observed numerically. In fact, denoting by $\vec{y}_h^*$ the solution to the discretized problem, we compute the ratios

$$r_k = \frac{\|\vec{y}_h^{k+1} - \vec{y}_h^*\|}{\|\vec{y}_h^k - \vec{y}_h^*\|},$$

FIGURE 5.4. Optimal Control

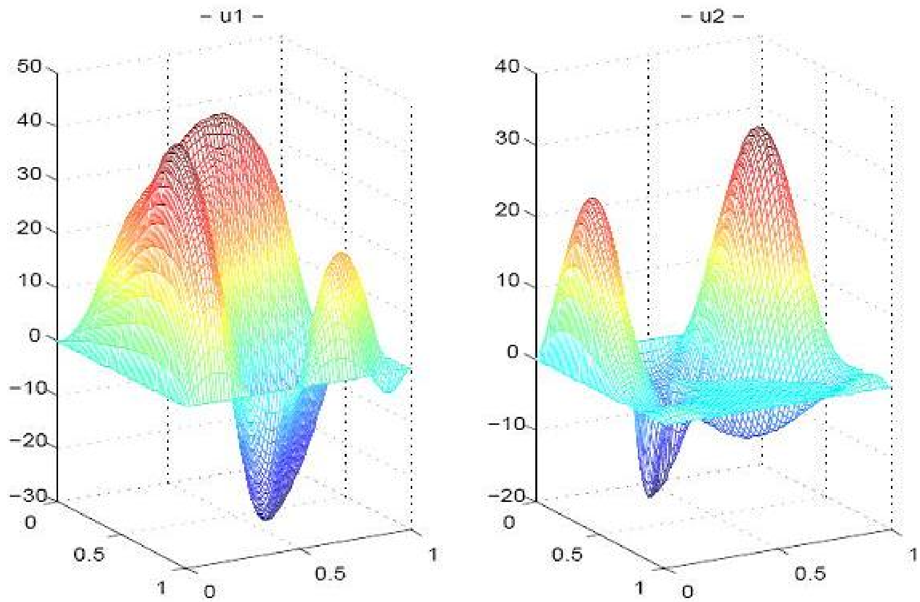


FIGURE 5.5. Active set

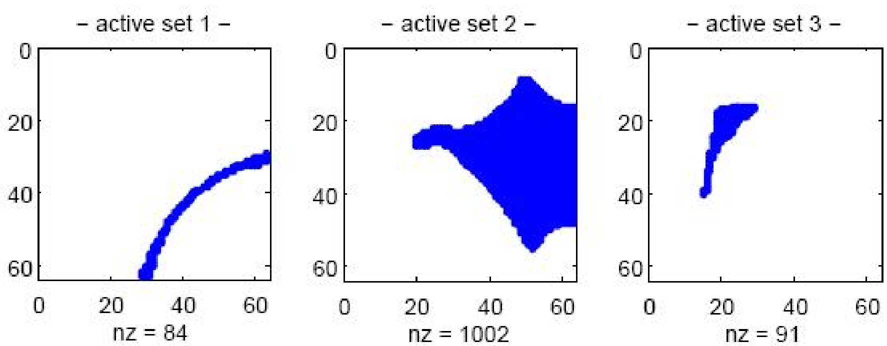


TABLE 5.1. super-linear convergence

k	11	12	13	14	15	16	17
r_k	0.9449	0.8910	0.7685	0.6405	0.4670	0.0118	0.0

for a fixed $\gamma = 1000$ and $\alpha = 0.01$. The result for iterations $k = 11, \ldots, 17$ is given in Table 5.1.

The above test problem has the same diffusion parameter in the underlying equations and the constraint set is a triangular with two equally long sides. We

FIGURE 5.6. Lagrange Multiplier

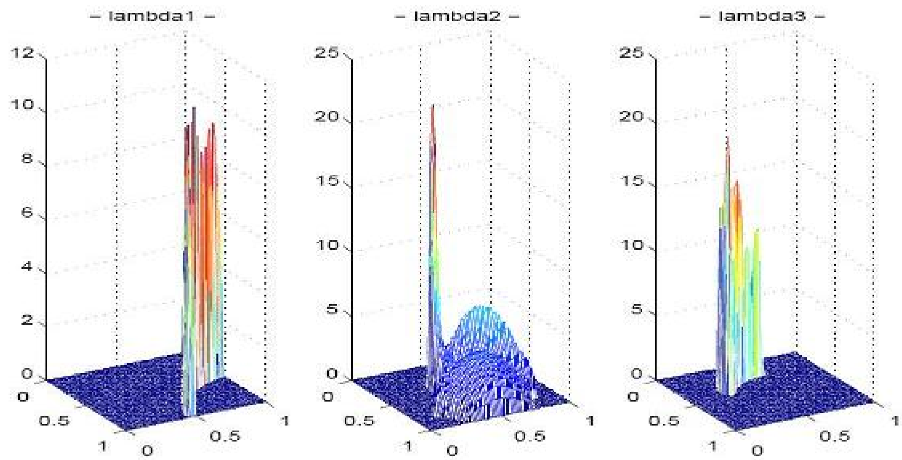


TABLE 5.2. $\gamma = 10^6$, $\alpha = 0.001$, $\beta = 1$

θ	0.1	0.5	1	10
iteration number	36	23	15	18
active components	1	2	3	3

TABLE 5.3. $\gamma = 10^6$, $\alpha = 0.001$, $\theta = 1$

β	0.01	1	100
iteration number	23	15	17
active components	3	3	3

then change the governing equation and the constraints to be

$$-\triangle y_1 + y_2 = u_1, \tag{5.4}$$

$$-\beta \triangle y_2 - y_1 = u_2. \tag{5.5}$$

and

$$y_1 \geq 0, \quad y_2 \geq 0, \quad \theta y_1 + y_2 \leq 1. \tag{5.6}$$

As before we set $\vec{y}_d = (4\sin(4\pi xy), 2\sin(2\pi x) + 2\cos(2\pi y))^T$, and fix $\gamma = 10^6$, $\alpha = 0.001$, $h = \frac{1}{64}$. We test the performance of the algorithm under changes of θ and β. It can be noted from Tables 5.2 and 5.3 that it converges for a wide range of values for θ and β, and that the number of iterations before convergence is achieved does not depend on these parameters significantly. As expect, taking into consideration the shape of y_d the number of components as well as the associated active sets increase with θ.

Appendix

Proof to the Proposition 3.4. Let $(\vec{y}^*, \vec{u}^*) \in \overrightarrow{W} \times L^2(\Omega, \mathbb{R}^2)$ denote the unique optimal solution. Since the Salter condition is not assumed, convex analysis is not directly applicable. To prove the existence of a Lagrange multiplier, we consider the convergence of the solutions to the γ problems. Let $(\vec{y}_\gamma, \vec{u}_\gamma, \vec{p}_\gamma, \vec{\lambda}_\gamma)$ satisfy

$$\begin{cases} \Lambda \vec{y}_\gamma = \vec{u}_\gamma, \\ (\vec{p}_\gamma, \Lambda \vec{z}) + (\vec{\lambda}_\gamma, M\vec{z}) = (\vec{y}_d - \vec{y}_\gamma, \vec{z}), \quad \text{for all } \vec{z} \in \overrightarrow{W}, \\ \alpha \vec{u}_\gamma = \vec{p}_\gamma, \\ \vec{\lambda}_\gamma = \gamma (M\vec{y}_\gamma - \psi)^+. \end{cases} \tag{5.7}$$

By similar arguments as in Section 4.1, we obtain

$$\begin{aligned} \vec{y}_\gamma &\rightharpoonup \hat{y} \quad \text{in } \overrightarrow{W}, \\ \vec{u}_\gamma &\rightharpoonup \hat{u} \quad \text{in } L^2(\Omega, \mathbb{R}^2), \end{aligned}$$

and $M\hat{y} \le \vec{\psi}$. For $J_\gamma(\vec{y}, \vec{u}) = J(\vec{y}, \vec{u}) + \frac{\gamma}{2}\|(M\vec{y} - \psi)^+\|^2$, we have

$$J_\gamma(\vec{y}_\gamma, \vec{u}_\gamma) \le J_\gamma(\vec{y}^*, \vec{u}^*) = J(\vec{y}^*, \vec{u}^*).$$

The weak lower semi-continuous of J implies that

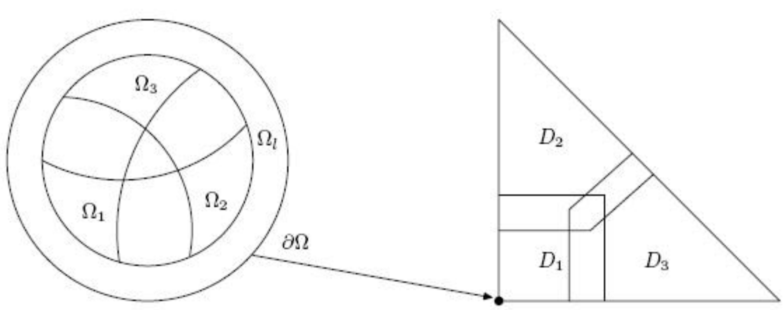
$$J(\hat{y}, \hat{u}) \le \liminf J(\vec{y}_\gamma, \vec{u}_\gamma) \le \liminf J_\gamma(\vec{y}_\gamma, \vec{u}_\gamma) \le J(\vec{y}^*, \vec{u}^*).$$

Since $\hat{y}$ satisfies the state constraints and $(\hat{y}, \hat{u})$ satisfies $\Lambda \hat{y} = \hat{u}$, uniqueness of the optimal solution $(\vec{y}^*, \vec{u}^*)$ implies that $(\vec{y}^*, \vec{u}^*) = (\hat{y}, \hat{u})$. From the second equation in system (5.7), we deduce that $\{\|M\vec{\lambda}_\gamma\|_{\overrightarrow{W}^*}\}_{\gamma \ge 1}$ is bounded. To prove that $\{\vec{\lambda}_\gamma\}_{\gamma \ge 1}$ is bounded in space $(W^*)^m$, we need to guarantee that for some constant $C > 0$ independent of $\gamma \ge 1$,

$$\langle \vec{\lambda}_\gamma, \vec{\phi} \rangle_{W,W^*} \le C\|\vec{\phi}\|_2, \qquad \text{for all } \vec{\phi} \in W(\Omega, \mathbb{R}^m).$$

We will follow the notation and the arguments in Section 4.1, and note that the construction of Ω_i for $i = 1, ..., m$ does not use assumption 3.2). Since we use homogenous Dirichlet boundary condition, $\vec{y}^*$ maps from $\partial\Omega$ to $\vec{0}$. Without loss of

Figure 5.7. Partition with Homogenous Dirichlet Boundary Condition



generality, we assume that $\vec{0}$ has strictly positive distance to D_i, for $i = 2, ..., m$, (see the figure 5.7). Hence a neighborhood of $\partial\Omega$ is in Ω_1. Therefore there exists a compact subdomain $\Omega_c \Subset \Omega$ such that $\overline{\Omega}\backslash\Omega_c \subset \Omega_1$. Since Ω_c is covered by open sets Ω_i, by partition of unity (c.f. [1]), there are smooth functions $\xi_i \in C_0^\infty(\Omega)$ with support in Ω_i, and

$$1 = \sum_{i=1}^{m} \xi_i(\mathbf{x}), \qquad \text{for all } \mathbf{x} \in \Omega_c.$$

Define $\xi_0 = 1 - \sum_{i=1}^m \xi_i$, which has support in Ω_1. Then any fixed function $\phi \in H^2 \cap H_0^1(\Omega)$, can be expressed as

$$\phi = \phi(\xi_0 + \sum_{i=1}^{m} \xi_i).$$

Let $\phi_i = \phi\xi_i$, for $i = 0, 1, ..., m$. Clear ϕ_i has support in Ω_i, for $i = 1, ..., m$, and ϕ_0 has support in Ω_1. Moreover, the domain Ω_i depends on the optimal solution $\vec{y}^*$, but not on γ. Therefore there exists a constant $\bar{C}$ which does not dependent on γ, such that

$$\|\phi_i\|_2 \le \bar{C}\|\phi\|_2, \quad i = 0, 1, ..., m. \tag{5.8}$$

Using the same technique as in Lemma 4.2, and noticing the uniform boundedness of $\|M\vec{\lambda}_\gamma\|_{\overrightarrow{W}^*}$, we obtain for $\vec{\varphi} \in W^m$ and with support in Ω_i,

$$\langle\vec{\lambda}_\gamma, \vec{\varphi}\rangle_{W,W^*} \le C_i\|\vec{\varphi}\|_2,$$

where C_i only depends on domain Ω_i and the bound on $\{\|M\vec{\lambda}_\gamma\|_{\overrightarrow{W}^*}\}_{\gamma\ge 1}$. Combining this with estimate (5.8), we have

$$\langle\vec{\lambda}_\gamma, \vec{\phi}\rangle_{W,W^*} = \sum_{i=0}^{m}\langle\vec{\lambda}_\gamma, \vec{\phi}_i\rangle_{W,W^*} \le \sum_i C_i\|\phi_i\|_2 \le (m+1)\bar{C}\max_i\{C_i\}\|\phi\|_2.$$

Hence $\{\vec{\lambda}_\gamma\}_{\gamma\ge 1}$ is uniformly bounded in $(W^*)^m$. Then there exists $\vec{\lambda}^* \in W^*(\Omega, \mathbb{R}^m)$ and a subsequence still denoted by $\vec{\lambda}_\gamma$, such that

$$\vec{\lambda}_\gamma \rightharpoonup \vec{\lambda}^* \quad \text{in } W^*(\Omega, \mathbb{R}^m).$$

This implies the existence of Lagrange multiplier of the Problem 1.1. In fact, to obtain the optimality system, we need to define the dual variable $\vec{p}^* = \alpha\vec{u}^*$. By passing to the limit in the first three equations of system (5.7), $(\vec{y}^*, \vec{u}^*, \vec{p}^*, \vec{\lambda}^*)$ satisfies

$$\begin{aligned}
&\Lambda\vec{y}^* = \vec{u}^*,\\
&(\vec{p}^*, \Lambda\vec{z}) + (\vec{\lambda}^*, M\vec{z}) = (\vec{y}_d - \vec{y}^*, \vec{z}), \quad \text{for all } \vec{z} \in \overrightarrow{W},\\
&\alpha\vec{u}^* = \vec{p}^*.
\end{aligned}$$

We need to check the linear complimentary condition to close the system. Positivity of $\vec{\lambda}^*$ follows from weak convergence of λ_γ. Utilizing the estimate (4.4), we find

$$\alpha\|\Lambda(\vec{y}_\gamma - \vec{y}^*)\|^2 \le -\|\vec{y}_\gamma - \vec{y}^*\|^2 - \frac{1}{\gamma}\|\vec{\lambda}_\gamma\|^2 + \langle\vec{\lambda}^*, M(\vec{y}_\gamma - \vec{y}^*)\rangle_{W^*,W}.$$

Since $\vec{y}_\gamma \rightharpoonup \vec{y}^*$ in $\overrightarrow{W}$, and $M^T\vec{\lambda}^* \in \overrightarrow{W}^*$, we obtain

$$\frac{1}{\gamma}\|\vec{\lambda}_\gamma\|^2 \to 0, \quad \vec{y}_\gamma \to \vec{y}^* \quad \text{in } \overrightarrow{W}.$$

Using that $\vec{\lambda}_\gamma = \gamma(M\vec{y}_\gamma - \vec{\psi})$, we have

$$(\vec{\lambda}_\gamma, M\vec{y}_\gamma - \vec{\psi}) = \frac{1}{\gamma}\|\vec{\lambda}_\gamma\|^2.$$

Strong convergence $\vec{y}_\gamma \to \vec{y}^*$ in $\overrightarrow{W}$ and weak convergence $M^T\vec{\lambda}_\gamma \rightharpoonup M^T\vec{\lambda}^*$ in $\overrightarrow{W}^*$ imply that

$$(\vec{\lambda}_\gamma, M\vec{y}_\gamma - \vec{\psi}) \to \langle \vec{\lambda}^*, M\vec{y}^* - \vec{\psi}\rangle_{W^*,W}.$$

Combining the above estimates leads to linear complimentary condition

$$\vec{\lambda}^* \geq 0, \quad M\vec{y}^* \leq \vec{\psi}, \quad \langle \vec{\lambda}^*, M\vec{y}^* - \vec{\psi}\rangle_{W^*,W} = 0.$$

This completes the proof. □

References

[1] R. Adams and J. Fournier, *Sobolev Spaces*, Academic Press, 2003.

[2] V. Barbu and T. Precupanu, *Convexity and Optimization in Banach Spaces*, Springer, 1986.

[3] M. Bergounioux and K. Kunisch, *On the Structure of Lagrange Multipliers for State-Constrained Optimal Control Problems*, Systems and Control Letters, 48 (2002), 169-176.

[4] E. Casas, *Control of an Elliptic Problem with Pointwise State Constraints*, SIAM J. Control Optim., 24 (1986), 1309-1322.

[5] D. Cohn, *Measure Theory*, Birkhäuser, 1980.

[6] K. Deckelnick, A. Günther and M. Hinze, *Finite element approximation of elliptic control problems with constraints on the gradient*, Numer. Math. 111 (2009), 335-350.

[7] M. Hintermüller, K. Ito, and K. Kunisch, *The Primal-Dual Active Set Strategy as a Semi-Smooth Newton Method*, SIAM J. Optim. 13 (2003), 865-888.

[8] K. Ito and K. Kunisch, *Semi-smooth Newton Methods for State-Constrained Optimal Control Problems*, Systems and Control Letters, 50 (2003), 221-228.

[9] P. Neittaanmaki, J. Sprekels and D. Tiba, *Optimization of Elliptic Systems: Theory and Applications*, Springer, 2005.

[10] U. Prüfert, F. Tröltzsch and M. Weiser, *The convergence of an interior point method for an elliptic control problem with mixed control-state constraints*, Comput. Optim. Appl. 39 (2008), 183-218.

[11] J.C. De Los Reyes, C. Meyer and B. Vexler, *Finite element error analysis for state-constrained optimal control of the Stokes equations*, Control Cybernet, 37 (2008), 251-284.

[12] A. Rösch and F. Tröltzsch, *On regularity of solutions and Lagrange multipliers of optimal control problems for semilinear equations with mixed pointwise control-state constraints*, SIAM J. Control Optim. 46 (2007), no. 3, 1098–1115.

[13] G. Stampacchia, *Le problème de Dirichlet pour les èquations elliptiques du second ordre à coefficients discontinus*, Ann. Inst. Fourier 15 (1965), 189-258.

Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36 A-8010 Graz, Austria

E-mail address: karl.kunisch@uni-graz.at

Department of Mathematics, Zhejiang University, P.R.China

E-mail address: matlkw@zju.edu.cn

Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstrasse 69, A-4040 Linz, Austria

E-mail address: xiliang.lu@ricam.oeaw.ac.at