IMA Journal of Numerical Analysis (2013) **33**, 875–897
doi:10.1093/imanum/drs029
Advance Access publication on November 23, 2012

Optimal control for an elliptic system with convex polygonal control constraints

KARL KUNISCH
Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria
karl.kunisch@uni-graz.at

AND

XILIANG LU*
School of Mathematics and Statistics, Wuhan University, Wuhan 430072, P.R. China
*Corresponding author: xllv.math@whu.edu.cn

[Received on 30 March 2012; revised on 29 July 2012]

The semismooth Newton method for optimal control problems for systems of partial differential equations with polygonal constraints on the controls is developed. The Newton derivatives are characterized for the case of systems of dimension two, superlinear convergence is verified and a simple proof-of-concept numerical example is provided.

Keywords: optimal control problem; control constraints; semismooth Newton method; superlinear convergence.

1. Introduction

In recent years, optimal control problems governed by partial differential equations with pointwise constraints on the controls have received a considerable amount of attention. The challenge consists in finding efficient numerical methods in spite of the nonsmoothness introduced by pointwise inequality constraints.

Most investigations so far have considered the case of scalar-valued control with simple constraints (e.g. unilateral or bilateral constraints). For vector-valued controls, which appear, for example, in the context of the Navier–Stokes equations, the constraints that were investigated were of a box type. For scalar-valued controls with unilateral or bilateral constraints, the numerical realization of optimal control problems can advantageously be performed with the semismooth Newton method or, equivalently, by the primal–dual active set method. We refer to Hintermüller *et al.* (2003), Ito & Kunisch (2004), De Los Reyes & Kunisch (2005), Hintermüller & Hinze (2006), Ito & Kunisch (2008) and Ulbrich (2011) and the references cited therein. An alternative approach is based on interior point methods. They were analysed, for example, in Schiela & Weiser (2008) and Weiser *et al.* (2008). The semismooth Newton method is different from penalty—or barrier—function methods which rely on penalty—or barrier—parameters, respectively. It was proven to be locally superlinearly convergent, and for unilaterally constrained problems to be globally convergent; cf. Hintermüller *et al.* (2003).

For vector-valued controls, there are only a few papers that deal with the case of general convex control constraints. In Wachsmuth (2006a,b) and Wachsmuth (2007), the first-order necessary and the second-order sufficient optimality conditions for general convex constraints are obtained. In Wachsmuth (2006b), a numerical procedure is proposed which, however, does not concretely exploit the structure of the constraint set and for which no convergence analysis is given. In Ulbrich (2003a), an abstract semismooth Newton approach is investigated that is applicable for constraints on the vector-valued controls. The detailed realization in that paper is carried out for pointwise constraints on the components of the control vector and does not cover the cases considered in the present paper. In Kunisch & Lu (submitted), the semismooth Newton method was investigated for a pointwise Euclidean norm constraint and in De Los Reyes & Kunisch (2009), for a class of affine control constraints, where the number of equations characterizing the constraints is not larger than the number of constraints.

In the present paper, we analyse the semismooth Newton method for constraints of polygonal type, a topic that has not been treated before. We suppose that Ω is a bounded domain in $\mathbb{R}^d$, $d=2,3$, which is either convex or has a $C^{1,1}$ boundary, and that state and control vectors map from Ω to $\mathbb{R}^l$. We assume $l=2$. A generalization of our approach to the case of l-dimensional controls is possible.

The control variables $\vec{u}$ are supposed to satisfy pointwise polygonal constraints, i.e.

$$\vec{u}(x) \in K \quad \text{a.e. } x \in \tilde{\Omega} \subset \Omega,$$

where $K \subset \mathbb{R}^2$ is a convex, closed polygon, and a.e. $x \in \tilde{\Omega}$ stands for 'almost every $x \in \tilde{\Omega}$'. At the end of this section, the constraints will be expressed by means of inequalities. The cost functional is chosen to be quadratic:

$$J_\alpha(\vec{y},\vec{u}) = \frac{1}{2}\|\vec{y}-\vec{y}_d\|^2_{L^2(\Omega,\mathbb{R}^2)} + \frac{\alpha}{2}\|\vec{u}\|^2_{L^2(\tilde{\Omega},\mathbb{R}^2)},$$

where $\alpha > 0$ and $\vec{y}_d \in L^2(\Omega,\mathbb{R}^2)$ are given, and $\tilde{\Omega}$ is a subdomain of Ω where the controls are localized. The equation constraint is chosen as an elliptic system in Ω with homogeneous, Dirichlet boundary conditions

$$\Lambda\vec{y} = \mathcal{B}\vec{u} \quad \text{in } \Omega. \tag{1.1}$$

The operator Λ is supposed to be a bounded linear second-order elliptic operator $H_0^1(\Omega,\mathbb{R}^2) \to H^{-1}(\Omega,\mathbb{R}^2)$ with L^∞ coefficients, and moreover, it is assumed to be strongly elliptic, i.e. there exists $\eta > 0$ such that $\langle \Lambda\vec{y},\vec{y}\rangle \geqslant \eta\|\vec{y}\|^2_{H^1}$ for every $\vec{y} \in H_0^1(\Omega,\mathbb{R}^2)$. The operator $\mathcal{B}$ is the extension-by-zero operator from the subdomain $\tilde{\Omega} \subset \Omega$ to Ω. Hence $\mathcal{B}$ is a bounded linear operator from $L^2(\tilde{\Omega},\mathbb{R}^2)$ to $L^2(\Omega,\mathbb{R}^2)$. Then by the Lax–Milgram theorem, for every $\vec{u} \in L^2(\tilde{\Omega},\mathbb{R}^2)$, the underlying system (1.1) admits a unique variational solution $y \in H_0^1(\Omega,\mathbb{R}^2)$. Moreover, we assume that the coefficients of the operator Λ are sufficiently regular such that for any $u \in L^2(\tilde{\Omega},\mathbb{R}^2)$ the solution satisfies $\vec{y} \in H^2(\Omega,\mathbb{R}^2)$ and the following *a priori* estimates hold:

$$\|\vec{y}\|_{H^2(\Omega,\mathbb{R}^2)} \leqslant C\|\Lambda\vec{y}\|_{L^2(\Omega,\mathbb{R}^2)}, \quad \|\vec{y}\|_{H^2(\Omega,\mathbb{R}^2)} \leqslant C\|\Lambda^*\vec{y}\|_{L^2(\Omega,\mathbb{R}^2)}, \tag{1.2}$$

for any $\vec{y} \in H_0^1(\Omega,\mathbb{R}^2) \cap H^2(\Omega,\mathbb{R}^2)$.

The optimal control problem under consideration is then given by the following problem.

PROBLEM 1.1

$$\min J_\alpha(\vec{y},\vec{u}) \quad \text{subject to (1.1) and } \vec{u}(x) \in K, \ \text{a.e. } x \in \tilde{\Omega}.$$

The main aim of this work is to verify that Problem 1.1 can be solved by a locally superlinearly convergent semismooth Newton method. In passing let us recall that linear quadratic problems also arise as auxiliary problems in the sequential quadratic programming approach to genuinely nonlinear optimal control problems.

For the optimal control problem with unilateral or bilateral control constraints, the semismooth Newton method was proved to be an efficient, superlinearly convergent technique; see Hintermüller *et al.* (2003) and Ito & Kunisch (2004). We briefly recall the notion of differentiability which will be used in this paper. Let X and Y be two Banach spaces, with D an open set in X and a map F from D to Y.

DEFINITION 1.2 $F: D \mapsto Y$ is called Newton (slantly) differentiable at $x \in D$ if there exists an open neighbourhood $N_x \subset D$, and mappings $G: N_x \mapsto \mathcal{L}(X, Y)$ (bounded linear operator from X to Y) such that

$$\lim_{\|h\|_X \to 0} \frac{\|F(x+h) - F(x) - G(x+h)(h)\|_Y}{\|h\|_X} = 0. \tag{1.3}$$

The mapping G is called the Newton derivative of F at x.

The key step to utilizing the semismooth Newton method for the optimal control problem, Problem 1.1, is to reformulate its optimality condition as a nonlinear, Newton-differentiable, operator equation $F(x) = 0$ in an appropriately chosen function-space setting.

The paper is organized as follows: in the remaining part of this section, we give the notation for the function spaces and the description of the constraints. Section 2 contains the existence and uniqueness of the optimal solution, the first-order necessary condition and its equivalent formulations. Moreover, Newton differentiability for the nonlinear equation representing the optimality system is proved. A semismooth Newton algorithm, together with its convergence analysis, is given in Section 3. In Section 4, a numerical example is presented to depict the superlinear convergence property of the algorithm.

Throughout this paper, standard notation for Sobolev spaces is used. $L^p(\Omega, \mathbb{R}^2)$ denotes the space of vector-valued L^p-integrable functions defined in Ω. The norm in L^2 is expressed as $\|\cdot\|$. Furthermore, $|\cdot|$ expresses the Euclidean vector norm in $\mathbb{R}^2$ or $\mathbb{R}^m$. The inner product $(\cdot,\cdot)$ is taken as the integration of two L^2 functions over Ω, and $(\cdot,\cdot)_{\tilde{\Omega}}$ stands for integration over the subdomain $\tilde{\Omega} \subset \Omega$.

We turn to the description of the polygonal set K. It is taken as the intersection of m half spaces. For $i = 1, 2, \ldots, m$, each half space is represented by the affine inequality $\vec{n}_i \cdot \vec{u} \leqslant \psi_i$, where $\vec{\mathrm{n}}_i$ is the unit outer normal vector to the half space. We denote $\vec{n}_i = (\cos\theta_i, \sin\theta_i)'$, with $\theta_i \in [0, 2\pi)$.

We can check that the angles must satisfy $\theta_i \neq \theta_j$ for all $i \neq j$. In fact, if $\theta_i = \theta_j$ and $\psi_i \leqslant \psi_j$, then $\vec{n}_i \cdot \vec{p} \leqslant \psi_i \Rightarrow \vec{n}_j \cdot \vec{p} \leqslant \psi_j$, and hence the constraint $\vec{h}_j \cdot \vec{p} \leqslant \psi_j$ is redundant. Without loss of generality, we utilize the ordering $2\pi > \theta_1 > \theta_2 > \cdots > \theta_m \geqslant 0$. To guarantee that K is convex, we assume that $\theta_i - \theta_{i+1} < \pi$ for $i = 1, \ldots, m$, where $\theta_{m+1} = \theta_1 - 2\pi$. We denote the sides of the polygonal domain by l_i (the two end points are not included) and the vertices by $a_{i,i+1}$ (at times we use $a_{m,1} = a_{m,m+1} = a_{0,1}$). Thus the sides and vertices are ordered in a clockwise manner; see Fig. 1.

The polygonal constraints can be put into matrix form by introducing

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

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

$$M = \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ \vdots & \vdots \\ \cos\theta_m & \sin\theta_m \end{pmatrix}, \quad \vec{\psi} = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_m \end{pmatrix}. \tag{1.4}$$

Then the convex polygonal domain K can be expressed as

$$K = \{\vec{u} \in \mathbb{R}^2 : M\vec{u} \leqslant \vec{\psi}\}.$$

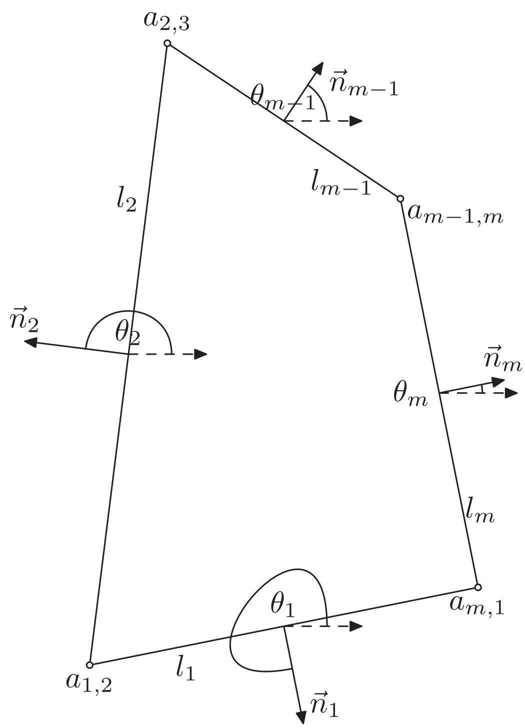

FIG. 1. Notation for the constraints.

2. Optimality system

We define the subset of admissible controls $\vec{u} \in L^2(\tilde{\Omega}, \mathbb{R}^2)$ as

$$D = \{\vec{u} \in L^2(\tilde{\Omega}, \mathbb{R}^2) : \vec{u}(x) \in K, \text{ a.e. } x \in \tilde{\Omega}\}. \tag{2.1}$$

Clearly, D is a closed convex set in $L^2(\tilde{\Omega}, \mathbb{R}^2)$. Denote the control-to-state mapping $\vec{u} \to \vec{y}(\vec{u})$ by T, where $T : L^2(\tilde{\Omega}, \mathbb{R}^2) \to L^2(\Omega, \mathbb{R}^2)$ is given by $T = \Lambda^{-1}\mathcal{B}$ and Λ^{-1} is the solution operator of the underlying elliptic system (1.1). Hence Problem 1.1 can equivalently be expressed as

$$\inf_{\vec{u} \in D} \hat{J}(\vec{u}), \tag{2.2}$$

where $\hat{J}(\vec{u}) = J_\alpha(T\vec{u}, \vec{u})$ denotes the reduced cost functional. Since $\hat{J}(\vec{u})$ is strictly convex and coercive, and the admissible set D is a closed convex set in the Hilbert space $L^2(\tilde{\Omega}, \mathbb{R}^2)$, existence and uniqueness of a solution to the minimization problem (2.2), and hence to Problem 1.1, can be proved with well-known techniques; see e.g. Ekeland & Temam (1999, p. 35, Proposition 1.2). The following result then follows from the equivalence between problem (2.2) and Problem 1.1.

THEOREM 2.1 There exists a unique solution to Problem 1.1.

Next we present the optimality system for the minimizer $(\vec{y}_*, \vec{u}_*)$ of Problem 1.1. Since $\hat{J}$ is differentiable and D is a closed convex set, we have

$$(\hat{J}'(\vec{u}_*), \vec{u}_* - \vec{v}) \leqslant 0 \quad \text{for all } \vec{v} \in D.$$

From the definition,

$$\hat{J}'(\vec{u}_*) = T^*(T\vec{u}_* - \vec{y}_d) + \alpha\vec{u}_* = \mathcal{B}^* \Lambda^{-*}(\vec{y}_* - \vec{y}_d) + \alpha\vec{u}_*,$$

where Λ^{-*} is the adjoint operator of Λ^{-1} and $\mathcal{B}^* : L^2(\Omega, \mathbb{R}^2) \to L^2(\tilde{\Omega}, \mathbb{R}^2)$ denotes the adjoint operator of $\mathcal{B}$, which is the restriction operator from Ω to $\tilde{\Omega}$. If we define the dual variable $\vec{p}_* = -\Lambda^{-*}(\vec{y}_* - \vec{y}_d)$, then the optimality system can be put into the following form:

$$\begin{aligned} \text{primal equation} \quad & \Lambda\vec{y}_* = \mathcal{B}\vec{u}_*, \\ \text{adjoint equation} \quad & \Lambda^*\vec{p}_* = \vec{y}_d - \vec{y}_*, \\ \text{optimality condition} \quad & (\mathcal{B}^*\vec{p}_* - \alpha\vec{u}_*, \vec{u}_* - \vec{v}) \geqslant 0 \quad \text{for all } \vec{v} \in D. \end{aligned} \tag{2.3}$$

LEMMA 2.2 The optimality condition in (2.3) is equivalent to

$$\vec{u}_* = \mathrm{Proj}_D\left(\frac{1}{\alpha}\mathcal{B}^*\vec{p}_*\right), \tag{2.4}$$

which is equivalent to the pointwise projection

$$\vec{u}_*(x) = \mathrm{Proj}_K\left(\frac{1}{\alpha}\vec{p}_*(x)\right), \quad \text{a.e. } x \in \tilde{\Omega}. \tag{2.5}$$

Here Proj_D and Proj_K denote the projections onto D in $L^2(\tilde{\Omega}, \mathbb{R}^2)$ and K in $\mathbb{R}^2$, respectively.

Proof. The first equivalence stated above follows from the projection onto a closed convex subset in a Hilbert space. Equivalence between the pointwise and the function-space projection follows, for instance, by observing that for $x \in \tilde{\Omega}$,

$$\mathrm{Proj}_K\left(\frac{1}{\alpha}\vec{p}_*(x)\right) = \underset{\vec{v}\in K}{\mathrm{argmin}} \left|\frac{1}{\alpha}\vec{p}_*(x) - \vec{v}\right|^2,$$

and uniqueness of the projection onto convex sets. □

By introducing Lagrange multipliers we will obtain a complementarity system that is equivalent to the pointwise projection.

LEMMA 2.3 For every vector $\vec{q} \in \mathbb{R}^2$, the projection $\vec{w} = \mathrm{Proj}_K((1/\alpha)\vec{q})$ is equivalent to the existence of a unique vector $\vec{\lambda} \in \mathbb{R}^m$ such that

$$\vec{q} = \alpha\vec{w} + M^{\mathrm{T}}\vec{\lambda}, \quad \vec{\lambda} \geqslant \vec{0}, \quad M\vec{w} - \vec{\psi} \leqslant \vec{0}, \quad \vec{\lambda} \cdot (M\vec{w} - \vec{\psi}) = 0. \tag{2.6}$$

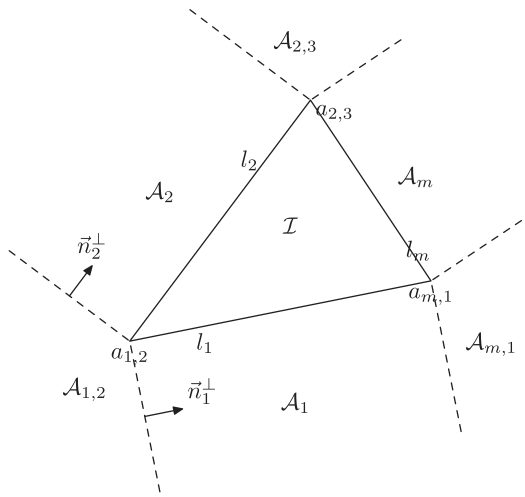

FIG. 2. The decomposition of the $\vec{q}$ plane.

Proof. Note at first that

$$\vec{w} = \underset{v \in K}{\operatorname{argmin}} \left| \frac{1}{\alpha}\vec{q} - \vec{v} \right|^2. \tag{2.7}$$

Due to the assumption $\theta_i - \theta_{i+1} \in (0, \pi)$, the solution $\vec{w}$ satisfies the regular point condition (in the sense of Zowe; cf. Zowe & Kurcyusz, 1979, p. 50) for the constraint $M\vec{w} \leqslant \vec{\psi}$ for each $\vec{q} \in \mathbb{R}^2$. Hence, by standard Lagrange multiplier theory for inequality constraints, there exists $\vec{\lambda} \in \mathbb{R}^m$ such that (2.6) holds (Luenberger, 1984; Ito & Kunisch, 2008). □

For the proof of semismoothness in Lemma 2.2, an explicit expression for $\vec{\lambda}$ in terms of $\vec{q}$ is derived next. For this purpose the $\vec{q}$ space $\mathbb{R}^2$ is decomposed into disjoint subsets, $\mathcal{I}$, $\mathcal{A}_i$, $\mathcal{A}_{i,i+1}$ for $i = 1, \ldots, m$ ($\mathcal{A}_{m,m+1} = \mathcal{A}_{m,1}$); see Fig. 2. The formula for the Lagrange multiplier $\vec{\lambda}$ depends on which subdomain the vector $(1/\alpha)\vec{q}$ belongs to. Here and below, for $i = m$, the index $m+1$ is set equal to the index 1. Three types of subdomain are defined as follows:

$$\begin{cases} \mathcal{I} = \{\vec{h} : \operatorname{Proj}_K(\vec{h}) \in \text{ interior of } K\}, \\ \mathcal{A}_i = \{\vec{h} : \operatorname{Proj}_K(\vec{h}) \in l_i\}, \\ \mathcal{A}_{i,i+1} = \{\vec{h} : \operatorname{Proj}_K(\vec{h}) = a_{i,i+1}\}. \end{cases} \tag{2.8}$$

The dotted lines in Fig. 2 are perpendicular to adjacent sides of the polygon.
Case 1: if $(1/\alpha)\vec{q} \in \mathcal{I}$, then all the constraints are inactive, and hence $\vec{w} = (1/\alpha)\vec{q}$ and $\vec{\lambda} = \vec{0}$.
Case 2: if $(1/\alpha)\vec{q} \in \mathcal{A}_i$, then the constraints $(1/\alpha)\vec{n}_j \cdot \vec{q} \leqslant \psi_j$ are not active for all $j \neq i$, and hence $\lambda_j = 0$ for all $j \neq i$. By definition of the projection operator, we have for some $\lambda_i \geqslant 0$,

$$\frac{1}{\alpha}\vec{q} - \vec{w} = \frac{1}{\alpha}\lambda_i \vec{n}_i,$$
$$\vec{w} \cdot \vec{n}_i = \psi_i.$$

Solving this linear system, we obtain

$$\lambda_i = \vec{q}\cdot\vec{n}_i - \alpha\psi_i, \quad \lambda_j = 0 \ \text{ for } j \neq i, \quad \vec{w} = \frac{1}{\alpha}\vec{q} - \frac{1}{\alpha}(\vec{q}\cdot\vec{n}_i)\vec{n}_i + \psi_i\vec{n}_i.$$

This provides the unique representation of $\vec{\lambda}$ and $\vec{w}$ on $\cup_{i=1}^m \mathcal{A}_i$.

Case 3: if $(1/\alpha)\vec{q} \in \mathcal{A}_{i,i+1}$, the constraints $(1/\alpha)\vec{n}_j\cdot\vec{q} \leqslant \psi_i$ for $j \neq i$, $j \neq i+1$ are not active, and hence necessarily $\lambda_j = 0$ for all $j = 1,\ldots,m$ with $j \neq i$, $j \neq i+1$. The vertex $a_{i,i+1}$ is presented by the intersection of two lines and we have $\vec{w} = M_{i,i+1}^{-1}\vec{\psi}_{i,i+1}$, where $M_{i,i+1}$ is the square matrix arising from the ith and $(i+1)$st rows of M. Furthermore, $\vec{\lambda}_{i,i+1}$ (or $\vec{\psi}_{i,i+1}$) denotes the ith and $(i+1)$st coordinates of $\vec{\lambda}$ (or $\vec{\psi}$). Therefore,

$$\vec{\lambda}_{i,i+1} = \alpha M_{i,i+1}^{-\mathrm{T}}\left(\frac{1}{\alpha}\vec{q} - \vec{w}\right) = M_{i,i+1}^{-\mathrm{T}}\vec{q} - \alpha(M_{i,i+1}M_{i,i+1}^{\mathrm{T}})^{-1}\vec{\psi}_{i,i+1},$$

where

$$M_{i,i+1}^{-\mathrm{T}} = -\frac{1}{\sin(\theta_i - \theta_{i+1})}\begin{pmatrix}\sin\theta_{i+1} & -\cos\theta_{i+1}\\ -\sin\theta_i & \cos\theta_i\end{pmatrix},$$

$$(M_{i,i+1}M_{i,i+1}^{\mathrm{T}})^{-1} = \frac{1}{\sin^2(\theta_i - \theta_{i+1})}\begin{pmatrix}1 & -\cos(\theta_i - \theta_{i+1})\\ -\cos(\theta_i - \theta_{i+1}) & 1\end{pmatrix}.$$

Then for $(1/\alpha)\vec{q} \in \mathcal{A}_{i,i+1}$ we find $\lambda_j = 0$ for $j \neq i$, $j \neq i+1$, and

$$\begin{aligned}\lambda_i &= \frac{1}{\sin^2(\theta_i - \theta_{i+1})}\left[-\sin(\theta_i - \theta_{i+1})(q_1\sin\theta_{i+1} - q_2\cos\theta_{i+1}) - \alpha\psi_i + \alpha\cos(\theta_i - \theta_{i+1})\psi_{i+1}\right],\\ \lambda_{i+1} &= \frac{1}{\sin^2(\theta_i - \theta_{i+1})}\left[-\sin(\theta_i - \theta_{i+1})(q_2\cos\theta_i - q_1\sin\theta_i) - \alpha\psi_{i+1} + \alpha\cos(\theta_i - \theta_{i+1})\psi_i\right].\end{aligned} \tag{2.9}$$

We have thus obtained the unique representation of $\vec{\lambda}$ as a function of $\vec{q}$.

It is convenient to summarize the representation of $\vec{\lambda}(\vec{q})$ that was obtained:

$$\lambda_i(\vec{q}) = \begin{cases} \vec{q}\cdot\vec{n}_i - \alpha\psi_i & \text{if } \frac{1}{\alpha}\vec{q} \in \mathcal{A}_i,\\ \frac{1}{\sin^2(\theta_i - \theta_{i+1})}\begin{pmatrix}-\sin(\theta_i - \theta_{i+1})(q_1\sin\theta_{i+1} - q_2\cos\theta_{i+1})\\ -\alpha\psi_i + \alpha\cos(\theta_i - \theta_{i+1})\psi_{i+1}\end{pmatrix} & \text{if } \frac{1}{\alpha}\vec{q} \in \mathcal{A}_{i,i+1},\\ \frac{1}{\sin^2(\theta_i - \theta_{i-1})}\begin{pmatrix}\sin(\theta_i - \theta_{i-1})(q_2\cos\theta_{i-1} - q_1\sin\theta_{i-1})\\ -\alpha\psi_i + \alpha\cos(\theta_i - \theta_{i-1})\psi_{i-1}\end{pmatrix} & \text{if } \frac{1}{\alpha}\vec{q} \in \mathcal{A}_{i-1,i},\\ 0 & \text{otherwise,}\end{cases} \tag{2.10}$$

where $\theta_0 = \theta_m, q_0 = q_m, \mathcal{A}_{0,1} = \mathcal{A}_{m,1}, \theta_{m+1} = \theta_1, q_{m+1} = q_1, \mathcal{A}_{m,m+1} = \mathcal{A}_{m,1}$.

LEMMA 2.4 The mapping $H : \vec{q} \to \vec{\lambda}(\vec{q})$ from $\mathbb{R}^2 \mapsto \mathbb{R}^m$ is piecewise affine and Lipschitz continuous.

Proof. The fact that H is piecewise affine follows directly from well-known nonlinear programming theory; see Scholtes. Alternatively, it is a direct consequence of (2.10). Local Lipschitz continuity follows from an abstract result on the sensitivity of solutions and Lagrange multipliers with respect to problem data, in abstract optimization problems; see e.g. Ito & Kunisch (2008, p. 46) and the references therein. It is applied to (2.7), with $\vec{q}$ denoting the perturbation parameter. The active set structure (see (2.10)), then implies that H is globally Lipschitz continuous as a mapping from $\mathbb{R}^2 \to \mathbb{R}^m$. □

By collecting the pointwise information in Lemma 2.3, we find a Lagrange multiplier associated to $\vec{u} = \mathrm{Proj}_D((1/\alpha)\mathcal{B}^*\vec{p})$ in function space.

LEMMA 2.5 For any $\vec{p} \in L^2(\Omega, \mathbb{R}^2)$, the projection $\vec{u} = \mathrm{Proj}_D((1/\alpha)\mathcal{B}^*\vec{p})$ is equivalent to the existence of a unique $\vec{\lambda} \in L^2(\tilde{\Omega}, \mathbb{R}^m)$ such that

$$\mathcal{B}^*\vec{p} = \alpha\vec{u} + M^{\mathrm{T}}\vec{\lambda}, \quad \vec{\lambda} \geqslant \vec{0}, \quad M\vec{u} - \vec{\psi} \leqslant \vec{0}, \quad (\vec{\lambda}, M\vec{u} - \vec{\psi})_{\tilde{\Omega}} = 0. \tag{2.11}$$

Proof. For $x \in \tilde{\Omega}$, we define $\vec{\lambda}(x)$ according to Lemma 2.3. From the explicit representation (2.10), we conclude that $\vec{\lambda} \in L^2(\tilde{\Omega}, \mathbb{R}^m)$ for $\vec{p} \in L^2(\Omega, \mathbb{R}^2)$. The pointwise information (2.6) implies that (2.11) holds. □

REMARK 2.6 For $\vec{p} \in C(\Omega, \mathbb{R}^2)$, by Lemma 2.4, the associated Lagrange multiplier $\vec{\lambda} = H(\mathcal{B}^*\vec{p})$ is also continuous on $\tilde{\Omega}$.

LEMMA 2.7 The mapping $\vec{p} \to \vec{\lambda} := H(\mathcal{B}^*\vec{p})$ is Newton differentiable from $L^s(\Omega, \mathbb{R}^2)$ to $L^t(\tilde{\Omega}, \mathbb{R}^m)$ for any $1 \leqslant t < s \leqslant \infty$.

Proof. It clearly suffices to consider the case when $\tilde{\Omega} = \Omega$, i.e. when $\mathcal{B}^*$ equals the identity. In the first part of the proof, we verify Newton differentiability for a special piecewise linear function. In the second step we reduce the general case to the special one. We shall use the fact that by Hölder's inequality, for any $1 \leqslant t < s \leqslant \infty$,

$$\|w\|_{L^t(\tilde{\Omega})} \leqslant |\tilde{\Omega}|^r \|w\|_{L^s(\tilde{\Omega})}, \quad \text{where } r = \begin{cases} \dfrac{s-t}{st} & \text{for } s < \infty, \\ \dfrac{1}{t} & \text{for } s = \infty. \end{cases}$$

Subsequently, we consider only the case $s < \infty$, and leave modifications for the case $s = \infty$ to the reader. *Step* 1: we first consider the specific piecewise linear function $f(\vec{p}) : \mathbb{R}^2 \to \mathbb{R}$ defined as (see Fig. 3):

$$f(\vec{p}) = \begin{cases} p_1, & \vec{p} \in I_1, \\ p_1 + p_2, & \vec{p} \in I_2, \\ 0, & \vec{p} \in I_3. \end{cases}$$

Accordingly, let $\mathbb{R}^2$ be decomposed into three mutually disjoint subsets:

$$I_1 = \{\vec{p} : p_1 > 0, p_2 > 0\}, \quad I_2 = \{\vec{p} : p_1 + p_2 > 0, p_2 \leqslant 0\}, \quad I_3 = \mathbb{R}^2 \backslash (I_1 \cup I_2).$$

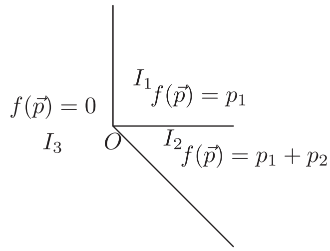

FIG. 3. Piecewise linear function.

Moreover, the vector function G is defined piecewise:

$$G(\vec{p}) = \begin{cases} (1,0), & \vec{p} \in I_1, \\ (1,1), & \vec{p} \in I_2, \\ (0,0), & \vec{p} \in I_3. \end{cases}$$

Associated to f defined above, there is a canonical substitution operator which we denote by the same symbol $f : L^s(\Omega, \mathbb{R}^2)$ to $L^t(\Omega, \mathbb{R})$. We shall show that f is Newton differentiable and that G is a Newton derivative of f. To this end, for any increment $\vec{h} \in L^s(\Omega, \mathbb{R}^2)$, we define the remainder term

$$R = f(\vec{p} + \vec{h}) - f(\vec{p}) - G(\vec{p} + \vec{h}) \cdot \vec{h}.$$

By the definition of a Newton derivative, it is sufficient to show that

$$\lim_{\|\vec{h}\|_{L^s(\Omega)} \to 0} \frac{\|R\|_{L^t(\Omega)}}{\|\vec{h}\|_{L^s(\Omega)}} = 0.$$

For this purpose, a partition of Ω is given as follows:

$$\Omega_0 = \{x \in \Omega : \vec{p}\,(x) \text{ and } (\vec{p} + \vec{h})(x) \text{ lie in same subset } I_i, \text{ for } i = 1,2,3\},$$
$$\Omega_{1,2} = \{x \in \Omega : \vec{p}\,(x) \in I_1,\ (\vec{p} + \vec{h})(x) \in I_2\},$$

and $\Omega_{2,1}$, $\Omega_{1,3}$, $\Omega_{3,1}$, $\Omega_{2,3}$, $\Omega_{3,2}$ are defined analogously. Let the index set $I = \{0, (1,2), (2,1), (1,3), (3,1), (2,3), (3,2)\}$. It can be checked that Ω_i, $i \in I$ are disjoint subsets of Ω and that their union is Ω. A computation shows that for $x \in \Omega_i$, $i \in I$, we have

$$R = \begin{cases} 0, & i = 0, \\ p_2, & i = (1,2), \\ -p_2, & i = (2,1), \\ -p_1, & i = (1,3), \\ p_1, & i = (3,1), \\ -p_1 - p_2, & i = (2,3), \\ p_1 + p_2, & i = (3,2). \end{cases}$$

Since Ω_i forms a disjoint partition of Ω, we have

$$\|R\|_{L^t(\Omega)} \leqslant \sum_{i\in I} \|R\|_{L^t(\Omega_i)}.$$

By definition, we need to show that for any small $\epsilon > 0$, there exists δ such that, for any $\|\vec{h}\|_{L^s(\tilde{\Omega})} \leqslant \delta$, we have

$$\frac{\sum_{i\in I} \|R\|_{L^t(\Omega_i)}}{\|\vec{h}\|_{L^s(\Omega)}} \leqslant \epsilon.$$

It is sufficient to check that this is true for each $i \in I$ separately. First consider the domain $\Omega_{1,2}$. For any $x \in \Omega_{1,2}$, by definition we have

$$\vec{p}(x) \in I_1, \quad (\vec{p}+\vec{h})(x) \in I_2,$$

which implies that $p_2(x) > 0$, $(p_2 + h_2)(x) \leqslant 0$, and in particular that

$$|p_2(x)| \leqslant |h_2(x)| \quad \text{for } x \in \Omega_{1,2}. \tag{2.12}$$

For the subdomain Ω_η given by

$$\Omega_\eta = \{x \in \Omega : 0 < p_2(x) < \eta\},$$

we have $\lim_{\eta\to 0^+} |\Omega_\eta| = 0$, where $|\cdot|$ denotes the measure of the given domain. Hence there exists η, independent of h, such that

$$|\Omega_\eta| \leqslant \left(\frac{\epsilon}{2}\right)^{st/(s-t)}.$$

For any $x \in \Omega_{1,2}\backslash\Omega_\eta$, we have $|h_2(x)| \geqslant |p_2(x)| \geqslant \eta$, which implies that

$$\|\vec{h}\|_{L^s(\Omega)} \geqslant \|h_2\|_{L^s(\Omega)} \geqslant \|h_2\|_{L^s(\Omega_{1,2}\backslash\Omega_\eta)} \geqslant \eta|\Omega_{1,2}\backslash\Omega_\eta|^{1/s}.$$

The choice $\delta = \eta(\epsilon/2)^{t/(s-t)}$ guarantees that for every $\|\vec{h}\|_{L^s(\Omega)} \leqslant \delta$, we have

$$\begin{aligned}
\frac{\|R\|_{L^t(\Omega_{1,2})}}{\|\vec{h}\|_{L^s(\Omega)}} &\leqslant \frac{\|R\|_{L^t(\Omega_{1,2}\cap\Omega_\eta)}}{\|\vec{h}\|_{L^s(\Omega)}} + \frac{\|R\|_{L^t(\Omega_{1,2}\backslash\Omega_\eta)}}{\|\vec{h}\|_{L^s(\Omega)}} \\
&\leqslant |\Omega_{1,2}\cap\Omega_\eta|^{(s-t)/st}\frac{\|p_2\|_{L^s(\Omega_{1,2}\cap\Omega_\eta)}}{\|h_2\|_{L^s(\Omega)}} + |\Omega_{1,2}\backslash\Omega_\eta|^{(s-t)/st}\frac{\|p_2\|_{L^s(\Omega_{1,2}\backslash\Omega_\eta)}}{\|h_2\|_{L^s(\Omega)}} \\
&\leqslant |\Omega_\eta|^{(s-t)/st} + |\Omega_{1,2}\backslash\Omega_\eta|^{(s-t)/st} \leqslant \frac{\epsilon}{2} + \left(\frac{\|\vec{h}\|_{L^s(\Omega)}}{\eta}\right)^{(s-t)/t} \leqslant \epsilon.
\end{aligned}$$

The remaining subsets Ω_i can be treated analogously. We note that the observation (2.12) plays an essential role in the proof. For the other subdomains Ω_i, $i \in I$, we have

- $|p_2(x)| \leqslant |h_2(x)|$ if $x \in \Omega_{2,1}$,
- $|p_1(x)| \leqslant \sqrt{2}|\vec{h}(x)|$ if $x \in \Omega_{1,3}$ or $x \in \Omega_{3,1}$,
- $|p_1(x) + p_2(x)| \leqslant \sqrt{2}|\vec{h}(x)|$ if $x \in \Omega_{2,3}$ or $x \in \Omega_{3,2}$.

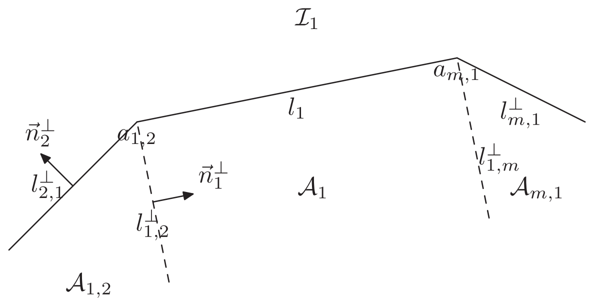

FIG. 4. Lagrange multiplier λ_1.

The case $x \in \Omega_{1,3}$, for example, can be checked as follows (other cases are very similar). Either $p_1(x) + h_1(x) \leqslant 0$, then $0 \leqslant p_1(x) \leqslant |h_1(x)|$ or $p_1(x) + h_1(x) > 0$, then $0 \leqslant p_1(x) \leqslant -h_1(x) - h_2(x) - p_2(x) \leqslant |h_1(x)| + |h_2(x)| \leqslant \sqrt{2}|\vec{h}(x)|$. In either case, $|p_1(x)| \leqslant \sqrt{2}|\vec{h}(x)|_{\mathbb{R}^2}$.

So far we have proved that $f(\vec{p})$ is Newton differentiable from $L^s(\Omega, \mathbb{R}^2)$ to $L^t(\Omega)$ for the specific choice of f. Next we show that $\vec{\lambda} = H(\vec{p})$ is Newton differentiable as a map from $L^s(\Omega, \mathbb{R}^2)$ to $L^t(\Omega, \mathbb{R}^m)$.

Step 2: with reference to (2.10) it suffices to consider one coordinate of $\vec{\lambda}(\vec{p})$. Without loss of generality, we show that λ_1 is Newton differentiable as a function of $\vec{p}$ from $L^s(\Omega, \mathbb{R}^2)$ to $L^t(\Omega)$. It will be convenient to refer back to Fig. 2 (see also Fig. 4) and to recall the formula for λ_1:

$$\lambda_1(\vec{p}) = \begin{cases} \vec{p} \cdot \vec{n}_1 - \alpha \psi_1 & \text{if } \frac{1}{\alpha}\vec{p} \in \mathcal{A}_1, \\ \dfrac{1}{\sin^2(\theta_1 - \theta_2)} \begin{pmatrix} -\sin(\theta_1 - \theta_2)(p_1 \sin\theta_2 - p_2 \cos\theta_2) \\ -\alpha\psi_1 + \alpha\cos(\theta_1 - \theta_2)\psi_2 \end{pmatrix} & \text{if } \frac{1}{\alpha}\vec{p} \in \mathcal{A}_{1,2}, \\ \dfrac{1}{\sin^2(\theta_1 - \theta_m)} \begin{pmatrix} \sin(\theta_1 - \theta_m)(p_2 \cos\theta_m - p_1 \sin\theta_m) \\ -\alpha\psi_1 + \alpha\cos(\theta_1 - \theta_m)\psi_m \end{pmatrix} & \text{if } \frac{1}{\alpha}\vec{p} \in \mathcal{A}_{m,1}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.13}$$

We next decompose λ_1 into two additive parts by the following construction. Choose two lines $\hat{l}_-$ and $\hat{l}_+$ which are perpendicular to l_1 (see Fig. 5) and denote the half space to the right of $\hat{l}_-$ by $\hat{A}_+$, and the half space to the left of $\hat{l}_+$ by $\hat{A}_-$. One can find two positive smooth functions ξ_+ and ξ_- which have support in $\hat{A}_+$ and $\hat{A}_-$, and satisfy $\xi_+ + \xi_- = 1$. We further introduce $\lambda_{1,-}$ (see Fig. 6) by

$$\lambda_{1,-}(\vec{p}) = \begin{cases} \vec{p} \cdot \vec{n}_1 - \alpha\psi_1, & \frac{1}{\alpha}\vec{p} \in \tilde{\mathcal{A}}_1, \\ \lambda_1, & \frac{1}{\alpha}\vec{p} \in \mathcal{A}_{1,2}, \\ 0, & \frac{1}{\alpha}\vec{p} \in \tilde{\mathcal{I}} = \Omega \setminus (\tilde{\mathcal{A}}_1 \cup \mathcal{A}_{1,2}), \end{cases}$$

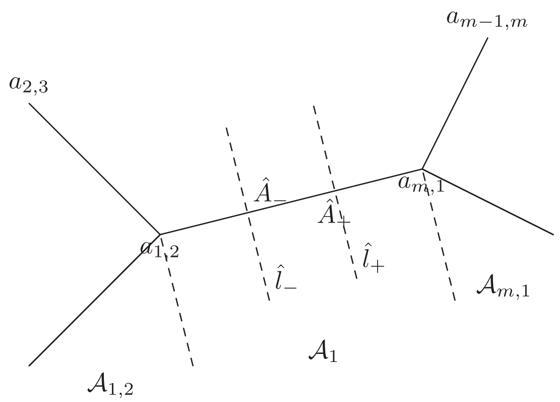

FIG. 5. Partition for λ_1.

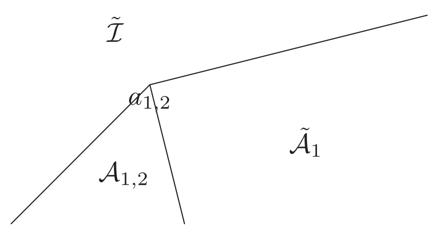

FIG. 6. Subdomains for $\lambda_{1,-}$.

where $\tilde{\mathcal{A}}_1$ denotes the sector depicted in Fig. 6. Furthermore, $\lambda_{1,+}$ is defined analogously to $\lambda_{1,-}$ with $\mathcal{A}_{1,2}$ replaced by $\mathcal{A}_{m,1}$.

With these preliminaries, the Lagrange multiplier λ_1 can be decomposed as

$$\lambda_1 = \lambda_1 \xi_- + \lambda_1 \xi_+ = \lambda_{1,-}\xi_- + \lambda_{1,+}\xi_+. \tag{2.14}$$

We next argue that $\lambda_{1,-}$ and $\lambda_{1,+}$ are affine transformations of f which was defined in Step 1. Without loss of generality, we focus on Newton differentiability of the function $\lambda_{1,-}$.

We first consider the case $\theta_1 - \theta_2 \in (0, \pi/2)$; the case $\theta_1 - \theta_2 \in [\pi/2, \pi)$ will be treated later. We define a new variable $\vec{q}$ according to

$$\vec{q} = \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ -\cot(\theta_1 - \theta_2)\sin\theta_1 & \cot(\theta_1 - \theta_2)\cos\theta_1 \end{pmatrix} (\vec{p} - \alpha M_{1,2}^{-1}\vec{\psi}_{1,2}). \tag{2.15}$$

A computation shows that

$$
\begin{aligned}
q_1 &= \cos\theta_1 \left(p_1 + \frac{\alpha}{\sin(\theta_1-\theta_2)}(\sin\theta_2\psi_1 - \sin\theta_1\psi_2)\right) \\
&\quad + \sin\theta_1 \left(p_2 + \frac{\alpha}{\sin(\theta_1-\theta_2)}(-\cos\theta_2\psi_1 + \cos\theta_1\psi_2)\right) \\
&= \vec{n}_1 \cdot \vec{p} - \alpha\psi_1, \\
q_2 &= \cot(\theta_1-\theta_2)\vec{n}_1^{\perp} \cdot \vec{p}_{1,2},
\end{aligned}
$$

where

$$
\vec{n}_1^{\perp} = \left(\cos\left(\theta_1 + \frac{\pi}{2}\right), \sin\left(\theta_1 + \frac{\pi}{2}\right)\right)' = (-\sin\theta_1, \cos\theta_1)',
$$

and

$$
\vec{p}_{1,2} = \vec{p}_{1,2}(\vec{p}) = \vec{p} - \alpha M_{1,2}^{-1}\vec{\psi}_{1,2} = \begin{pmatrix} p_1 + \dfrac{\alpha}{\sin(\theta_1-\theta_2)}(\sin\theta_2\psi_1 - \sin\theta_1\psi_2) \\ p_2 + \dfrac{\alpha}{\sin(\theta_1-\theta_2)}(-\cos\theta_2\psi_1 + \cos\theta_1\psi_2) \end{pmatrix}. \tag{2.16}
$$

We also define

$$
\vec{n}_2^{\perp} = \left(\cos\left(\theta_2 - \frac{\pi}{2}\right), \sin\left(\theta_2 - \frac{\pi}{2}\right)\right)' = (\sin\theta_2, -\cos\theta_2)'.
$$

Then from Fig. 4 it is noted that

$$
l_{1,2}^{\perp} = \{\vec{p} : \vec{n}_1^{\perp} \cdot \vec{p}_{1,2}(\vec{p}) = 0\}, \quad l_{2,1}^{\perp} = \{\vec{p} : \vec{n}_2^{\perp} \cdot \vec{p}_{1,2}(\vec{p}) = 0\},
$$

and

$$
\begin{cases} \vec{n}_1^{\perp} \cdot \vec{p}_{1,2}(\vec{p}) > 0 & \text{for } \dfrac{1}{\alpha}\vec{p} \in \mathcal{A}_1, \\ \vec{n}_1^{\perp} \cdot \vec{p}_{1,2}(\vec{p}) \leqslant 0 \text{ and } \vec{n}_2^{\perp} \cdot \vec{p}_{1,2}(\vec{p}) < 0 & \text{for } \dfrac{1}{\alpha}\vec{p} \in \mathcal{A}_{1,2}. \end{cases} \tag{2.17}
$$

One can check as follows.

(1) For $(1/\alpha)\vec{p} \in \tilde{\mathcal{A}}_1$, we have $q_1 > 0$ and $q_2 = \cot(\theta_1-\theta_2)\vec{n}_1^{\perp} \cdot \vec{p}_{1,2} > 0$. Hence $\vec{q} \in I_1$ (see Fig. 3) and $\lambda_{1,-}(\vec{p}) = q_1 = f(\vec{q})$.

(2) For $(1/\alpha)\vec{p} \in \mathcal{A}_{1,2}$, we have $q_2 \leqslant 0$ and

$$
\begin{aligned}
q_1 + q_2 &= (\vec{n}_1 + \cot(\theta_1-\theta_2)\vec{n}_1^{\perp})\vec{p}_{1,2}(\vec{p}) + \alpha(\vec{n}_1 \cdot M_{1,2}^{-1}\vec{\psi}_{1,2} - \psi_1) \\
&= -\frac{1}{\sin(\theta_1-\theta_2)}\vec{n}_2^{\perp} \cdot \vec{p}_{1,2}(\vec{p}) > 0;
\end{aligned} \tag{2.18}
$$

hence $\vec{q} \in I_2$. By (2.13) and (2.16), we find

$$
\lambda_{1,2}(\vec{p}) = q_1 + q_2 = f(\vec{q}).
$$

(3) For $(1/\alpha)\vec{p} \in \tilde{\mathcal{I}}$, similarly $\vec{q} \in I_3$ and $\lambda_{1,2}(\vec{p}) = 0 = f(\vec{q})$.

Therefore, for $0 < \theta_1 - \theta_2 < \pi/2$, we have established that

$$\lambda_{1,-}(\vec{p}) = f(\vec{q}(\vec{p})),$$

and by the chain rule (Ito & Kunisch, 2008, p. 238), $\vec{p} \to \lambda_{1,-}(\vec{p})$ is Newton differentiable.

The case $\theta_1 - \theta_2 \in [\pi/2, \pi)$ can be treated similarly. We define the variable $\vec{s}$ as

$$\vec{s} = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \begin{pmatrix} \vec{n}_1 \cdot \vec{p}_{1,2} + \cot(\theta_1 - \theta_2)\vec{n}_1^{\perp} \cdot \vec{p}_{1,2} \\ -\cot(\theta_1 - \theta_2)\vec{n}_1^{\perp} \cdot \vec{p}_{1,2} \end{pmatrix}.$$

Using (2.17) we find the following.

(1) For $(1/\alpha)\vec{p} \in \tilde{\mathcal{A}}_1$, we have $s_2 \leqslant 0$ and $s_1 + s_2 > 0$, and hence $\vec{s} \in I_2$ (see Fig. 3) and $\lambda_{1,2}(\vec{p}) = s_1 + s_2 = f(\vec{s})$.

(2) For $(1/\alpha)\vec{p} \in \mathcal{A}_{1,2}$, $s_2 > 0$ and (2.18) imply that $s_1 = -(1/\sin(\theta_1 - \theta_2))\vec{n}_2^{\perp} \cdot \vec{p}_{1,2}(\vec{p}) > 0$; hence $\vec{q} \in I_1$ and $\lambda_{1,2}(\vec{p}) = s_1 = f(\vec{s})$.

(3) For $(1/\alpha)\vec{p} \in \tilde{\mathcal{I}}$, similarly $\vec{s} \in I_3$ and $\lambda_{1,2}(\vec{p}) = 0 = f(\vec{s})$.

Again by the chain rule, $\vec{p} \to f(\vec{s}(\vec{p})) = \lambda_{1,-}(\vec{p})$ is Newton differentiable.

Combining the above results, Newton differentiability of $\vec{p} \to \lambda_1(\vec{p})$ from $L^t(\Omega, \mathbb{R}^2)$ to $L^s(\Omega)$ now follows from (2.14). □

REMARK 2.8 Alternatively, the proof of the previous lemma could be based on a superposition operator argument considering the chain of operations: $\vec{p} \to \mathcal{B}^*\vec{p} \to H(\mathcal{B}^*\vec{p})$ from $L^s(\Omega, \mathbb{R}^2)$ to $L^s(\tilde{\Omega}, \mathbb{R}^2)$ and further to $L^t(\tilde{\Omega}, \mathbb{R}^m)$. The first of these two mappings is clearly regular. The second one, considered as a pointwise mapping from $\mathbb{R}^2 \to \mathbb{R}^m$, is globally Lipschitz continuous by Lemma 2.4. Moreover, it can be argued that it is semismooth as a mapping from $\mathbb{R}^2$ to $\mathbb{R}^m$; see e.g. Ito & Kunisch (2008, p. 218) or Ulbrich (2011, p. 27). Then, due to the topology gap between $L^s(\tilde{\Omega}, \mathbb{R}^2)$ and $L^t(\tilde{\Omega}, \mathbb{R}^m)$, the composite mapping is semismooth; see e.g. Ito & Kunisch (2008, p. 235) or Ulbrich (2011, p. 61). For the definitions of semismoothness in finite- and infinite-dimensional spaces, we also refer, respectively, to these references. For the proof of the Lemma we preferred to give a self-contained argument.

3. Semismooth Newton algorithm

From (2.3) and Lemma 2.5, the first-order optimality condition has the form

$$\begin{aligned} \alpha \Lambda \vec{y}_* &= \mathcal{B}(\mathcal{B}^*\vec{p}_* - M^{\mathrm{T}}\vec{\lambda}_*), \\ \Lambda^*\vec{p}_* &= \vec{y}_d - \vec{y}_*, \\ \vec{\lambda}_* &= H(\mathcal{B}^*\vec{p}_*). \end{aligned} \tag{3.1}$$

Since Problem 1.1 is strictly convex, the optimality system is also a sufficient condition for Problem 1.1. Moreover, just like $\vec{y}_*$, the adjoint state $\vec{p}_*$ and the multiplier λ_* are unique.

We next aim at solving (3.1) by a Newton-type method. For this purpose, we introduce a nonlinear operator $F : H^2(\Omega, \mathbb{R}^2) \cap H_0^1(\Omega, \mathbb{R}^2) \times H^2(\Omega, \mathbb{R}^2) \cap H_0^1(\Omega, \mathbb{R}^2) \times L^2(\tilde{\Omega}, \mathbb{R}^m) \mapsto$

$L^2(\Omega,\mathbb{R}^2)\times L^2(\Omega,\mathbb{R}^2)\times L^2(\tilde{\Omega},\mathbb{R}^m)$ by

$$F(x)=\begin{pmatrix}\alpha\Lambda\vec{y}-\mathcal{B}(\mathcal{B}^*\vec{p}-M^{\mathrm{T}}\vec{\lambda})\\ \Lambda^*\vec{p}-\vec{y}_d+\vec{y}\\ \vec{\lambda}-H(\mathcal{B}^*\vec{p})\end{pmatrix}, \tag{3.2}$$

where $x=(\vec{y},\vec{p},\vec{\lambda})$. The Newton derivative of the nonlinear operator $F(x)$ can be expressed as

$$G_NF(x)=\begin{pmatrix}\alpha\Lambda & -\mathcal{B}\mathcal{B}^* & \mathcal{B}M^{\mathrm{T}}\\ I & \Lambda^* & 0\\ 0 & -G_NH(\mathcal{B}^*\vec{p})\mathcal{B}^* & I\end{pmatrix}, \tag{3.3}$$

where the specific form of the Newton derivative G_NH will be given below in Lemma 3.1.

Then the semismooth Newton algorithm for (3.1) consists in an iteration solving

$$G_NF(x^k)(x^{k+1}-x^k)=-F(x^k) \tag{3.4}$$

for x^{k+1}. It is equivalent to solving

$$\begin{cases}\alpha\Lambda\vec{y}^{k+1}=\mathcal{B}\mathcal{B}^*\vec{p}^{k+1}-\mathcal{B}M^{\mathrm{T}}\vec{\lambda}^{k+1},\\ \Lambda^*\vec{p}^{k+1}=\vec{y}_d-\vec{y}^{k+1},\\ \vec{\lambda}^{k+1}=G_NH(\mathcal{B}^*\vec{p}^k)[\mathcal{B}^*(\vec{p}^{k+1}-\vec{p}^k)]+H(\mathcal{B}^*\vec{p}^k).\end{cases} \tag{3.5}$$

The complete algorithm is given in Algorithm 1. One can observe that if $\vec{p}^{k+1}=\vec{p}^k$, then the Newton step (3.5) coincides with the optimality system (3.1), and uniqueness implies that $\vec{p}^k=\vec{p}_*$. Moreover, as we will prove in Theorem 3.3, the semismooth Newton method is locally superlinearly convergent. Therefore, it is reasonable to choose $\|\vec{p}^{k+1}-\vec{p}^k\|\leqslant\epsilon$, for a given small ϵ, as a stopping rule. In practice, we observe that the algorithm often converges in finitely many steps to the exact (discretized) solution.

Algorithm 1 Semismooth Newton algorithm

1: set k=0, initialize $\vec{p}^0\in H^2(\Omega,\mathbb{R}^2)\cap H^1_0(\Omega,\mathbb{R}^2)$.
2: solve (3.5) for $(\vec{y}^{k+1},\vec{p}^{k+1},\vec{\lambda}^{k+1})$.
3: stop or update $k=k+1$, and go to 2.

For the error functions defined by $\vec{z}^k=\vec{y}^k-\vec{y}_*$, $\vec{q}^k=\vec{p}^k-\vec{p}_*$ and $\vec{\xi}^k=\vec{\lambda}^k-\vec{\lambda}_*$, we have

$$\begin{cases}\alpha\Lambda\vec{z}^{k+1}=\mathcal{B}\mathcal{B}^*\vec{q}^{k+1}-\mathcal{B}M^{\mathrm{T}}\vec{\xi}^{k+1},\\ \Lambda^*\vec{q}^{k+1}=-\vec{z}^{k+1},\\ \vec{\xi}^{k+1}=G_NH(\mathcal{B}^*\vec{p}^k)\mathcal{B}^*\vec{q}^{k+1}+\vec{r}^k,\end{cases} \tag{3.6}$$

where $\vec{r}^k=H(\mathcal{B}^*\vec{p}^k)-H(\mathcal{B}^*\vec{p}^*)-G_N(\mathcal{B}^*\vec{p}^k)\mathcal{B}^*\vec{q}^k$. By Lemma 2.2, we find that if $\|\vec{q}^k\|_{L^q(\Omega)}\to 0$ for some $q>2$, then

$$\|\vec{r}^k\|_{L^2(\tilde{\Omega})}=\mathcal{O}(\|\vec{q}^k\|_{L^q(\Omega)}). \tag{3.7}$$

Substituting $\vec{\xi}^{k+1}$ into the first error equation, we obtain

$$\begin{cases} \alpha \Lambda \vec{z}^{k+1} - \mathcal{B}\mathcal{B}^* \vec{q}^{k+1} + \mathcal{B} M^{\mathrm{T}} G_N H(\mathcal{B}^* \vec{p}^k) \mathcal{B}^* \vec{q}^{k+1} = -\mathcal{B} M^{\mathrm{T}} \vec{r}^k, \\ \Lambda^* \vec{q}^{k+1} = -\vec{z}^{k+1}. \end{cases} \tag{3.8}$$

To obtain the well posedness of the Newton iteration and to derive error estimates, we need the following lemma.

LEMMA 3.1 For any given function $\vec{p} \in L^2(\Omega, \mathbb{R}^2)$, the operator $\mathcal{C} = \mathcal{B}\mathcal{B}^* - \mathcal{B} M^{\mathrm{T}} G_N H(\mathcal{B}^* \vec{p}) \mathcal{B}^*$ is non-negative in the sense that

$$(\vec{v}, [\mathcal{B}\mathcal{B}^* - \mathcal{B} M^{\mathrm{T}} G_N H(\mathcal{B}^* \vec{p}) \mathcal{B}^*] \vec{v}) \geqslant 0, \qquad \text{for all } \vec{v} \in L^2(\Omega, \mathbb{R}^2).$$

Proof. From Lemma 2.2, we can obtain the explicit form of $G_N H(\vec{q})$ for any vector $\vec{q}$. We will show that for any $\vec{q}$, the matrix $G = I - M^{\mathrm{T}} G_N H(\vec{q})$ is symmetric semidefinite. Recalling the notation in (2.8), we consider three characteristic locations for the vector $\vec{q}$, which are one inactive and two active components.
Case 1: if $(1/\alpha)\vec{q} \in \mathcal{I}$, $G_N H(\vec{q}) = 0$, which implies that $G = I$.
Case 2: if $(1/\alpha)\vec{q} \in \mathcal{A}_1$, then

$$\vec{\lambda} = (\vec{q} \cdot \vec{n}_1 - \alpha \psi_1, 0, \ldots, 0)^{\mathrm{T}},$$

which implies that

$$G_N H(\vec{q}) = \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix}, \quad M^{\mathrm{T}} G_N H(\vec{q}) = \begin{pmatrix} \cos^2\theta_1 & \cos\theta_1 \sin\theta_1 \\ \cos\theta_1 \sin\theta_1 & \sin^2\theta_1 \end{pmatrix}.$$

Hence the matrix

$$I - M^{\mathrm{T}} G_N H(\vec{q}) = \begin{pmatrix} \sin^2\theta_1 & -\cos\theta_1 \sin\theta_1 \\ -\cos\theta_1 \sin\theta_1 & \cos^2\theta_1 \end{pmatrix}$$

is symmetric semidefinite.
Case 3: If $(1/\alpha)\vec{q} \in \mathcal{A}_{1,2}$, then $\vec{\lambda} = (\lambda_1, \lambda_2, 0, \ldots, 0)^{\mathrm{T}}$, where

$$\lambda_1 = \frac{1}{\sin^2(\theta_1 - \theta_2)} [-\sin(\theta_1 - \theta_2)(q_1 \sin\theta_2 - q_2 \cos\theta_2) - \alpha\psi_1 + \alpha\cos(\theta_1 - \theta_2)\psi_2],$$

$$\lambda_2 = \frac{1}{\sin^2(\theta_1 - \theta_2)} [-\sin(\theta_1 - \theta_2)(q_2 \cos\theta_1 - q_1 \sin\theta_1) - \alpha\psi_2 + \alpha\cos(\theta_1 - \theta_2)\psi_1].$$

This implies that

$$G_N H(\vec{q}) = \frac{1}{\sin(\theta_2 - \theta_1)} \begin{pmatrix} \sin\theta_2 & -\cos\theta_2 \\ -\sin\theta_1 & \cos\theta_1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix}, \quad M^{\mathrm{T}} G_N H(\vec{q}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In this case, the matrix $I - M^{\mathrm{T}} G_N H(\vec{q})$ is zero and hence symmetric semidefinite.

Using the above fact and the definition of $\mathcal{B}^*$, for any function $\vec{v} \in L^2(\Omega, \mathbb{R}^2)$, we have

$$(\vec{v}, [\mathcal{B}\mathcal{B}^* - \mathcal{B}M^{\mathrm{T}} G_N H(\mathcal{B}^*\vec{p})\mathcal{B}^*]\vec{v}) = \int_{\Omega} \mathcal{B}^*\vec{p} \cdot (I - M^{\mathrm{T}} G_N H(\mathcal{B}^*\vec{p}))(\mathcal{B}^*\vec{v})\, \mathrm{d}x$$
$$= \int_{\tilde{\Omega}} \vec{v} \cdot (I - M^{\mathrm{T}} G_N H(\vec{p}))\vec{v}\, \mathrm{d}x \geqslant 0. \qquad \square$$

COROLLARY 3.2 For any initialization, the Newton iterates of (3.5) are well defined and satisfy $\vec{x}^k = (\vec{y}^k, \vec{u}^k, \vec{\lambda}^k) \in H^2(\Omega, \mathbb{R}^2) \cap H_0^1(\Omega, \mathbb{R}^2) \times H^2(\Omega, \mathbb{R}^2) \cap H_0^1(\Omega, \mathbb{R}^2) \times L^2(\tilde{\Omega}, \mathbb{R}^m)$ for $k = 1, \ldots$.

Proof. Let $\mathcal{C}$ be the operator defined in Lemma 3.1 and set $\vec{g} = \mathcal{B}M^{\mathrm{T}}(G_N H(\mathcal{B}^*\vec{p}^k)B^*\vec{p}^k - H(\mathcal{B}^*\vec{p}^k))$, $(\vec{y}, \vec{p}) = (\vec{y}^{k+1}, \vec{p}^{k+1})$. Then, inserting the last equation of (3.5) into the first, we find the following equations:

$$\begin{cases} \alpha \Lambda \vec{y} = \mathcal{C}\vec{p} + \vec{g}, \\ \Lambda^* \vec{p} = \vec{y}_d - \vec{y}, \end{cases} \tag{3.9}$$

and $\vec{\lambda} = G_N H(\mathcal{B}^*\vec{p}^k)[\mathcal{B}^*(\vec{p} - \vec{p}^k)] + H(\mathcal{B}^*\vec{p}^k)$. It is not difficult to check that $\vec{g} \in L^2(\Omega, \mathbb{R}^2)$.

For any $\mu > 0$, let us introduce the operator $\mathcal{A}_\mu : (H_0^1(\Omega, \mathbb{R}^2))^2 \to (H^{-1}(\Omega, \mathbb{R}^2))^2$ by

$$\mathcal{A}_\mu(\vec{y}, \vec{p}) = \begin{pmatrix} \alpha \Lambda \vec{y} - \mathcal{C}\vec{p} \\ \Lambda^* \vec{p} - \vec{y} \end{pmatrix} + \mu \begin{pmatrix} \vec{y} \\ \vec{p} \end{pmatrix}.$$

For $\mu > 0$, sufficiently large, the Lax–Milgram theorem implies that $\mathcal{A}_\mu$ is continuously invertible. In particular, we can consider the restriction of $\mathcal{A}_\mu^{-1}$ to $(L^2(\Omega, \mathbb{R}^2))^2$, and $\mathcal{A}_\mu^{-1} : (L^2(\Omega, \mathbb{R}^2))^2 \to (L^2(\Omega, \mathbb{R}^2))^2$ is a compact operator. Therefore, by a standard Fredholm alternative argument,

$$\mathcal{A}_0(\vec{y}, \vec{p}) = \begin{pmatrix} \vec{g} \\ \vec{y}_d \end{pmatrix}$$

has a unique solution $(\vec{y}, \vec{p})$ in $(H_0^1(\Omega, \mathbb{R}^2))^2$ if

$$\begin{cases} \alpha \Lambda \vec{y} - \mathcal{C}\vec{p} = 0, \\ \Lambda^* \vec{p} + \vec{y} = 0 \end{cases} \tag{3.10}$$

has only the trivial solution; see e.g. Evans (1998, p. 303). Taking inner products with $-\vec{p}$ and $\alpha\vec{y}$ in the first and second equations of (3.10) and adding them, we obtain $(\mathcal{C}\vec{p}, \vec{p}) + \alpha\|\vec{y}\|^2 = 0$. Using the positive semidefiniteness of $\mathcal{C}$, it follows that (3.10) has only the zero solution. Hence (3.9) admits a unique solution $(\vec{y}, \vec{p}) \in (H_0^1(\Omega, \mathbb{R}^2))^2$.

Finally, $(\vec{y}, \vec{p}) \in (H^2(\Omega, \mathbb{R}^2) \cap H_0^1(\Omega, \mathbb{R}^2))^2$ follows from (1.2), and $\vec{\lambda} \in L^2(\tilde{\Omega}, \mathbb{R}^m)$ can be verified from its formula. This concludes the proof. $\square$

Finally, we address convergence of the semismooth Newton method.

THEOREM 3.3 Algorithm 1 is locally superlinearly convergent; more precisely, if $\|\vec{p}^0 - \vec{p}_*\|_{H^1(\Omega)}$ is sufficiently small, then $(\vec{y}^k, \vec{p}^k, \vec{\lambda}^k)$ converges superlinearly in $H^2(\Omega, \mathbb{R}^2) \times H^2(\Omega, \mathbb{R}^2) \times L^2(\tilde{\Omega}, \mathbb{R}^m)$ to $(\vec{y}_*, \vec{p}_*, \vec{\lambda}_*)$.

Proof. It is sufficient to show

$$\|\vec{p}^{k+1} - \vec{p}_*\|_{H^2(\Omega)} + \|\vec{y}^{k+1} - \vec{y}_*\|_{H^2(\Omega)} + \|\vec{\lambda}^{k+1} - \vec{\lambda}_*\|_{L^2(\tilde{\Omega})} = \mathcal{O}(\|\vec{p}^k - \vec{p}_*\|_{H^2(\Omega)}).$$

We utilize the error equation (3.6). Taking the inner product with $\vec{q}^{k+1}$ in the first equation of (3.8) results in

$$\begin{aligned}(\vec{q}^{k+1}, [\mathcal{B}\mathcal{B}^* - \mathcal{B}M^{\mathrm{T}}G_N H(\mathcal{B}^*\vec{p})\mathcal{B}^*]\vec{q}^{k+1}) - (\mathcal{B}M^{\mathrm{T}}\vec{r}^k, \vec{q}^{k+1}) &= \alpha(\Lambda\vec{z}^{k+1}, \vec{q}^{k+1})\\ &= \alpha(\vec{z}^{k+1}, \Lambda^*\vec{q}^{k+1})\\ &= -\alpha\|\Lambda^*\vec{q}^{k+1}\|^2.\end{aligned}$$

By (1.2), using Lemma 3.1 and Young's inequality, we obtain

$$\|\vec{q}^{k+1}\|_{H^2(\Omega)} \leqslant C\|\Lambda^*\vec{q}^{k+1}\| \leqslant C\|\mathcal{B}M^{\mathrm{T}}\vec{r}^k\| \leqslant C\|\vec{r}^k\|_{L^2(\tilde{\Omega})},$$

for a constant C independent of k. For $d = 2, 3$, we have $H^1(\Omega) \hookrightarrow L^p$ for all $2 < p \leqslant 6$. Hence by observing (3.7), we conclude that

$$\|\vec{q}^{k+1}\|_{H^2(\Omega)} = \mathcal{O}(\|\vec{p}^k - \vec{p}_*\|_{H^2(\Omega)}).$$

For the term $\|\vec{p}^{k+1} - \vec{p}_*\|_{H^2(\Omega)}$, we can use the first equation of (3.6). Since

$$\alpha\|\Lambda\vec{z}^{k+1}\| \leqslant \|\mathcal{B}\mathcal{B}^*\vec{q}^{k+1}\| + \|\mathcal{B}M^{\mathrm{T}}G_N H(\mathcal{B}^*\vec{p}^k)\mathcal{B}^*\vec{q}^{k+1}\| + \|\mathcal{B}M^{\mathrm{T}}\vec{r}^k\|,$$

using the pointwise uniform boundedness of $H(\mathcal{B}^*\vec{p})$ for $\vec{p} \in \mathbb{R}^2$ and (1.2), we find that

$$\|\vec{p}^{k+1} - \vec{p}_*\|_{H^2(\Omega)} \leqslant C\|\Lambda\vec{z}^{k+1}\| = \mathcal{O}(\|\vec{p}^k - \vec{p}_*\|_{H^2(\Omega)}).$$

The last assertion, $\|\vec{\xi}^{k+1}\|_{L^2(\tilde{\Omega})} = \mathcal{O}(\|\vec{p}^k - \vec{p}_*\|_{H^2(\Omega)})$, follows by the representation $\vec{\xi}^{k+1} = G_N H(\mathcal{B}^*\vec{p}^k)\mathcal{B}^*\vec{q}^{k+1} + \vec{r}^k$, and the same argument as in the case of $\|\vec{p}^{k+1} - \vec{p}_*\|_{H^2(\Omega)}$. □

Remark 3.4 The optimal control for $\vec{u}$ also enjoys the superlinear convergence property in L^2. This immediately follows from

$$\vec{u}_* = \frac{1}{\alpha}(B^*\vec{p}_* - M^{\mathrm{T}}\vec{\lambda}_*) \quad \text{and} \quad \vec{u}^k = \frac{1}{\alpha}(B^*\vec{p}^k - M^{\mathrm{T}}\vec{\lambda}^k)$$

and Theorem 3.3.

We close this section by remarking that the optimality system (2.3) allows extra regularity of the optimal control and the multiplier to be asserted.

Proposition 3.5 The optimal control and associated Lagrange multiplier satisfy $(\vec{u}_*, \vec{\lambda}_*) \in W^{1,p}(\tilde{\Omega}, \mathbb{R}^2) \times W^{1,p}(\tilde{\Omega}, \mathbb{R}^m)$ for any $2 \leqslant p \leqslant 6$.

Proof. From (2.3) we have that $\vec{p}_* \in H^2(\Omega, \mathbb{R}^2)$ and hence $\vec{p}_* \in W^{1,p}(\Omega, \mathbb{R}^2)$. By Lemma 2.4 the mapping H is Lipschitz continuous and hence in $W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^m)$. Combined, this implies that $\vec{\lambda}_* \in W^{1,p}(\Omega, \mathbb{R}^m)$, and by (2.11) we obtain $\vec{u}_* \in W^{1,p}(\Omega, \mathbb{R}^2)$. □

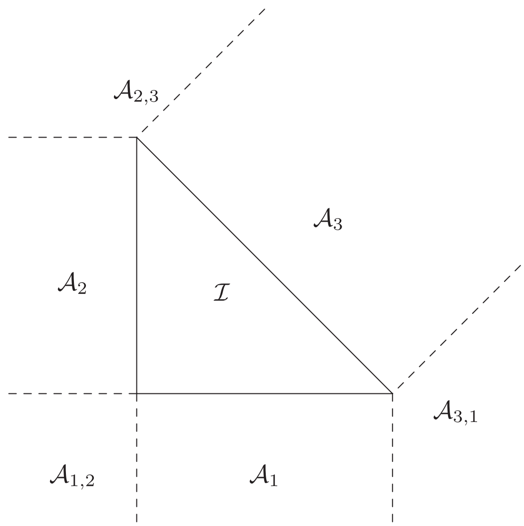

FIG. 7. The decomposition of the $\vec{p}$ plane.

4. Numerical example

We illustrate the efficiency of the semismooth Newton algorithm in the case of polygonal constraints for a numerical example. The underlying control system consists of an elliptic system in the unit square with a homogeneous Dirichlet boundary condition, and the controls act on the whole domain, i.e. we consider

$$-\triangle y_1 + y_2 = u_1, \quad -\triangle y_2 - y_1 = u_2, \text{ in } \Omega, \quad \vec{y}|_{\partial\Omega} = \vec{0},$$

where $\Omega = (0, 1)^2$. It can be verified that this system is strongly elliptic and satisfies inequality (1.2). Let the polygonal domain K be given by

$$u_1 \geqslant 0, \quad u_2 \geqslant 0, \quad u_1 + u_2 \leqslant 1.$$

Using the notation introduced in Section 2, we have

$$\theta_1 = \tfrac{3}{2}\pi, \quad \theta_2 = \pi, \quad \theta_3 = \tfrac{1}{4}\pi,$$

and the constraints can be put into matrix form $M\vec{u} \leqslant \vec{\psi}$, where

$$M = \begin{pmatrix} 0 & -1 \\ -1 & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}, \quad \psi = \begin{pmatrix} 0 \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix}.$$

The space $\mathbb{R}^2$ for $\vec{p}$ is decomposed into seven subdomains: $\mathcal{I}$, $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$, $\mathcal{A}_{1,2}$, $\mathcal{A}_{2,3}$, $\mathcal{A}_{3,1}$, as shown in Fig. 7. The Newton derivative $G(\vec{p}) = G_N H(\vec{p})$ is given by

(1)

$$G(\vec{p}) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{for } \frac{1}{\alpha}\vec{p} \in \mathcal{I},$$

(2)

$$G(\vec{p}) = \begin{pmatrix} 0 & -1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{for } \frac{1}{\alpha}\vec{p} \in \mathcal{A}_1,$$

(3)

$$G(\vec{p}) = \begin{pmatrix} 0 & 0 \\ -1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{for } \frac{1}{\alpha}\vec{p} \in \mathcal{A}_2,$$

(4)

$$G(\vec{p}) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \quad \text{for } \frac{1}{\alpha}\vec{p} \in \mathcal{A}_3,$$

(5)

$$G(\vec{p}) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{for } \frac{1}{\alpha}\vec{p} \in \mathcal{A}_{1,2},$$

(6)

$$G(\vec{p}) = \sqrt{2}\begin{pmatrix} 0 & 0 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & 1 \end{pmatrix} \quad \text{for } \frac{1}{\alpha}\vec{p} \in \mathcal{A}_{2,3},$$

(7)

$$G(\vec{p}) = \sqrt{2}\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{for } \frac{1}{\alpha}\vec{p} \in \mathcal{A}_{3,1}.$$

Recall that the cost functional is given by

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

Here we choose $\vec{y}_d = (0.2\sin(3\pi xy), 0.2(\sin(5\pi x) + \cos(10\pi y)))^{\mathrm{T}}$, $\alpha = 0.0005$. The choice of $\vec{y}_d$ and α guarantees that all three constraints are active on some parts of the domain. Since the exact solution is in general not known, the question arises of how to demonstrate convergence rates. For this example, however, the algorithm finds the exact (discretized) solution in finitely many (l) steps. Therefore, we replace $\|\vec{y}_* - \vec{y}^k\|_{H^1(\Omega)}$ by $\|\vec{y}^l - \vec{y}^k\|$ and similarly for $\vec{u}$ and $\vec{p}$. In fact, for this example, $l = 5$. The control variables do not appear in the Newton iteration (3.5), rather they are recovered by $\vec{u}^k = (1/\alpha)(\vec{p}^k - M^{\mathrm{T}}\vec{\lambda}^k)$. In Table 1, superlinear convergence can be observed numerically. It should be noted that the initial guess is the zero function which is quite far from the final solution. The control variables are depicted in Fig. 8. It can be noted that all three constraints $u_1 \geqslant 0$, $u_2 \geqslant 0$ and $u_1 + u_2 \leqslant 1$ are active in some parts of the domain.

TABLE 1 *Superlinear convergence*

	Iteration number				
	1	2	3	4	5
$\|y^k - y^5\|_{H^2(\Omega)}$	0.0685	0.0044	3.9068e−005	3.4286e−008	0
$\frac{\|y^k - y^5\|_{H^2(\Omega)}}{\|y^{k-1} - y^5\|_{H^2(\Omega)}}$		0.0642	0.008879	0.000878	0
$\|p^k - p^5\|_{H^2(\Omega)}$	0.0034	9.8179e−005	5.6036e−007	3.6116e−010	0
$\frac{\|p^k - p^5\|_{H^2(\Omega)}}{\|p^{k-1} - p^5\|_{H^2(\Omega)}}$		0.02887	0.0057	0.000645	0
$\|u^k - u^5\|_{L^2(\Omega)}$	6.3597	5.8744	0.1769	1.9255e−004	0
$\frac{\|u^k - u^5\|_{L^2(\Omega)}}{\|u^{k-1} - u^5\|_{L^2(\Omega)}}$		0.9237	0.03	0.001088	0

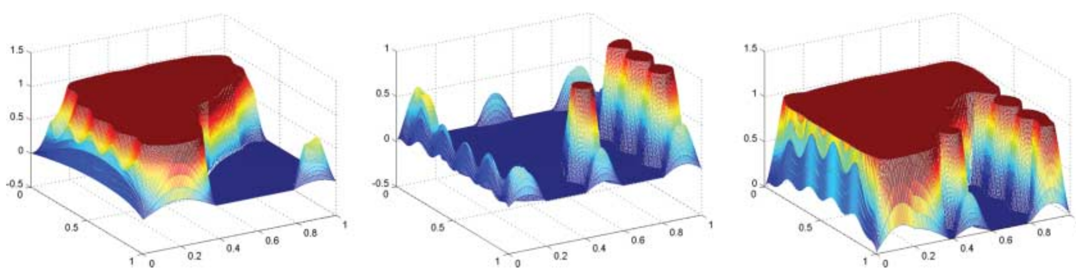

FIG. 8. Plot for the control variables. From left to right: u_1, u_2 and $u_1 + u_2$.

TABLE 2 *Number of iterations*

Grid number	8	16	32	64	128	256
Number of iterations	5	5	5	5	5	5

The numerical realization is based on a finite difference discretization with respect to a uniform axis-parallel grid. The mesh size for the results reported in Table 1 is $h = 1/256$. To demonstrate mesh independence of the algorithm, we compute the same example on a series of grids for mesh sizes refined by a factor of $1/2$. The number of iterations to obtain the finite-dimensional solution is found to remain constant for these grid levels; see Table 2.

REMARK 4.1 We also considered the case where the differential operator was changed to be $\Lambda\vec{y} = \begin{pmatrix} -\mathrm{div}(\beta_1 \nabla y_1) + y_2 \\ -\mathrm{div}(\beta_2 \nabla y_2) - y_1 \end{pmatrix}$, where β_1 and β_2 are piecewise constant coefficients defined by $\beta_1 = 1 + \chi_{x>0.5}$ and $\beta_2 = 1 + \chi_{y>0.5}$, and $\chi_{y>0.5}$ is the characteristic function of the set $\{(x, y) \in \Omega : y > 0.5\}$. The remaining specifications are as in the example above. For this numerical example, the operator Λ does not satisfy the *a priori* estimate (1.2). Still, the convergence behaviour is similar to that of the previous example; in particular, the iteration number to find the exact discretized solution is less than 6 and the convergence rate is superlinear.

Acknowledgement

The authors would like to thank the anonymous referees for their constructive feedback and useful comments.

Funding

The first author was supported in part by the Fonds zur Förderung der wissenschaftlichen Forschung under SFB 32, Mathematical Optimization and Applications in the Biomedical Sciences. The second author was supported by the National Natural Science Foundation of China (11101316).

References

De Los Reyes, J. C. & Kunisch, K. (2005) A semi-smooth Newton method for control constrained optimal control of the Navier Stokes equations. *Nonlinear Anal.*, **62**, 1289–1316.

De Los Reyes, J. C. & Kunisch, K. (2009) Optimal control of partial differential equations with affine control constraints. *Control Cybernet.*, **38**, 1217–1250.

Ekeland, I. & Temam, R. (1999) *Convex Analysis and Variational Problems*. Philadelphia: SIAM.

Evans, L. (1998) *Partial Differential Equations*. Providence: American Mathematical Society.

Hintermüller, M. & Hinze, M. (2006) A SQP-semismooth Newton-type algorithm applied to control of the instationary Navier-Stokes system subject to control constraints. *SIAM J. Optim.*, **16**, 1177–1200.

Hintermüller, M., Ito, K. & Kunisch, K. (2003) The primal-dual active set strategy as a semi-smooth Newton method. *SIAM J. Optim.*, **13**, 865–888.

Ito, K. & Kunisch, K. (2004) The primal-dual active set method for nonlinear optimal control problems with bilateral constraints. *SIAM J. Control Optim.*, **43**, 357–376.

Ito, K. & Kunisch, K. (2008) *Lagrange Multiplier Approach to Variational Problems and Applications*. Philadelphia: SIAM.

Kunisch, K. & Lu, X. Optimal control for an elliptic system with pointwise Euclidean norm constraints on the controls. *Math. Prog.* (Accepted for publication).

Luenberger, D. G. (1984) *Linear and Nonlinear Programming*. Menlo Park: Addison-Wesley.

Schiela, A. & Weiser, M. (2008) Superlinear convergence of the control reduced interior point method for PDE constrained optimization. *Comput. Optim. Appl.*, **39**, 369–393.

Scholtes, S. Introduction to piecewise differentiable equations. *Habilitation Thesis*, Preprint No. 53/1994, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe.

Ulbrich, M. (2003a) Semismooth Newton methods for operator equations in function spaces. *SIAM J. Optim.*, **13**, 805–842.

Ulbrich, M. (2011) *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*. Philadelphia: SIAM.

Wachsmuth, D. (2006a) Sufficient second-order optimality conditions for convex control constraints. *J. Math. Anal. Appl.*, **319**, 228–247.

Wachsmuth, D. (2006b) Numerical solution of optimal control problems with convex control constraints. *Systems, Control, Modeling and Optimization* (F. Ceragioli, A. Dontchev, H. Furuta, K. Marti & L. Pandolfi eds). New York: Springer, pp. 319–327.

WACHSMUTH, D. (2007) Optimal control problems with convex constraints. *Control of Coupled Partial Differential Equations* (K. Kunisch, G. Leugering, J. Sprekels & F. Tröltzsch eds). International Series of Numerical Mathematics, vol. 155. Birkhäuser: Verlag AG, pp. 311–328.

WEISER, M., GÄNZLER, T. & SCHIELA, A. (2008) A control reduced primal interior point method for PDE constrained optimization. *Comput. Optim. Appl.*, **41**, 127–145.

ZOWE, J. & KURCYUSZ, S. (1979) Regularity and stability for the mathematical programming problem in Banach spaces. *Appl. Math. Optim.*, **5**, 49–62.