

**FORMULATION AND NUMERICAL SOLUTION  
OF NASH EQUILIBRIUM MULTIOBJECTIVE  
ELLIPTIC CONTROL PROBLEMS**

Alfio Borzi and Christian Kanzow

Preprint 306

February 2012

<sup>2</sup> University of Würzburg  
Institute of Mathematics  
Emil-Fischer-Str. 30, 97074 Würzburg, Germany  
e-mail: [alfio.borzi@mathematik.uni-wuerzburg.de](mailto:alfio.borzi@mathematik.uni-wuerzburg.de)  
[kanzow@mathematik.uni-wuerzburg.de](mailto:kanzow@mathematik.uni-wuerzburg.de)

February 3, 2012

**Abstract.** The formulation and the semismooth Newton solution of Nash equilibria multiobjective elliptic optimal control problems are presented. Existence and uniqueness of a Nash equilibrium is proved. The corresponding solution is characterized by an optimality system that is approximated by second-order finite differences and solved with a semismooth Newton scheme. It is demonstrated that the numerical solution is second-order accurate and that the semismooth Newton iteration is globally and locally quadratically convergent. Results of numerical experiments confirm the theoretical estimates and show the effectiveness of the proposed computational framework.

**Key Words:** Multiobjective optimization, Nash equilibrium, optimal control problems, elliptic partial differential equations, semismooth Newton method, global convergence, quadratic convergence.

**Mathematics Subject Classification (2000):** 49J20, 65K10, 65N06, 90C30.

# 1 Introduction

Many application problems involve the simultaneous optimization of several competing objectives and constraints that define the class of multiobjective or multicriteria optimization problems. In this class of problems, we can distinguish between finite- and infinite-dimensional optimization problems. The former are characterized by finite-dimensional models and often refer to game theory problems. The latter can be associated to models in function spaces and, in particular, to differential models. A well-known application for infinite-dimensional multiobjective optimization occurs in aerodynamic shape optimization for the design of airplanes where one focus is the lift maximization in the critical phase of take-off and landing and the other one is drag minimization in the cruise regime since it directly determines kerosene consumption [9, 15]. Another application is multi-loading structural design [32]. Multiobjective PDE optimization would be also important in the treatment of models and data uncertainties where it is required to compute a robust optimal solution and minimize variance and other statistical moments [4].

We focus on multiobjective optimization problems governed by PDE models and in this case much less is known on the characterization and the solution of these problems. The starting point of our research are the works of Pareto, Nash, and Stackelberg; see [5, 8, 10] and references therein for more details on their ground breaking contributions and subsequent results.

These works usually consider finite-dimensional models, whereas part of our motivation is to consider an infinite-dimensional setting where optimization problems with partial differential equations (PDE) are involved. While the Pareto solution may be found in a number of works dealing with infinite-dimensional optimization problems, cf. [23, 24], much less is known regarding the Nash solution concept in multiobjective PDE-optimization problems. An exception are two recent papers [29, 30] by Ramos et al. devoted to a multiobjective optimization with a specific unconstrained control mechanism and solved with a conjugate gradient scheme. One of our purposes is to generalize this setting by considering a larger class of constrained controls and to solve the resulting problem with semismooth Newton methods.

Although the Nash formulation is typically viewed as a stronger solution concept than Pareto, the latter appears to be more popular in applications. On the other hand, it is only recently that powerful methods for the solution of finite-dimensional Nash equilibrium problems have been proposed. The purpose of this paper is to introduce the Nash equilibrium as a solution concept to multiobjective PDE-optimization problems and to show how to extend powerful finite-dimensional methods to solve the resulting discretized Nash equilibrium problems (NEPs, for short).

From the PDE optimization point of view, the class of problems considered in this paper are simple. Each of the two objectives are (strongly) convex as a function in the variable of the corresponding player (control), and the PDE is of the elliptic type. However, we do not consider two separate optimization problems, since we are looking for a Nash equilibrium, and even in the finite-dimensional setting with strongly convex objective functions, neither existence nor uniqueness of a Nash solution is guaranteed. This is very much in contrast

to standard optimization theory, and points out the difficulty inherent in the solution of NEPs, cf. [11, 14] for more details.

The paper is organized as follows. In Section 2, we formulate the class of multiobjective PDE-optimization problems considered throughout this paper and formally introduce the Nash solution concept. Section 3 gives an existence and uniqueness result for this class of problems in the infinite-dimensional setting. A finite difference discretization together with its approximation properties is then discussed in Section 4. In Section 5, we restate a semismooth Newton method for the solution of general NEPs which may be viewed as a standard solver and which is known to be both globally and locally quadratically convergent under certain assumptions. The two subsequent Sections 6 and 7 focus on the particular structure of the discretized multiobjective PDE-optimization problem and show that the conditions for global and local fast convergence hold. Numerical results are presented in Section 8, and we conclude with some final remarks in Section 9.

## 2 Multiobjective Elliptic Control Problems

This section defines the class of problems considered in this paper and contains the definition of the Nash solution concept. We consider multiobjective optimization problems governed by a linear elliptic partial differential equation as follows. Find  $(u_1, u_2) \in \mathcal{U}_{ad,1} \times \mathcal{U}_{ad,2} \subset L^2(\Omega) \times L^2(\Omega)$  such that

$$\begin{cases} \min_{y,u_1} & J_1(y, u_1, u_2) \\ \text{s.t.} & -\Delta y = B_1 u_1 + B_2 u_2 + f \quad \text{in } \Omega \\ & y = 0 \quad \text{on } \partial\Omega, \\ & u_1 \in \mathcal{U}_{ad,1}, \end{cases} \quad (1)$$

and

$$\begin{cases} \min_{y,u_2} & J_2(y, u_1, u_2) \\ \text{s.t.} & -\Delta y = B_1 u_1 + B_2 u_2 + f \quad \text{in } \Omega \\ & y = 0 \quad \text{on } \partial\Omega, \\ & u_2 \in \mathcal{U}_{ad,2}, \end{cases} \quad (2)$$

where  $f \in L^2(\Omega)$  and  $\Omega$  is a convex open bounded set in  $\mathbb{R}^2$ .

The cost functionals  $J_j$ ,  $j = 1, 2$ , are of the tracking type and are given by

$$J_j(y, u_1, u_2) := \frac{1}{2} \|y - z_j\|_{L^2(\Omega)}^2 + \frac{\nu_j}{2} \|B_j u_j\|_{L^2(\Omega)}^2, \quad j = 1, 2, \quad (3)$$

where  $z_1, z_2 \in L^2(\Omega)$  are given target functions and  $\nu_j > 0$ ,  $j = 1, 2$ , are the weights of the costs of the controls.

The set of admissible controls  $\mathcal{U}_{ad} = \mathcal{U}_{ad,1} \times \mathcal{U}_{ad,2}$  is the space product of two closed convex subsets of  $L^2(\Omega)$  given by

$$\mathcal{U}_{ad,j} = \{u \in L^2(\omega_j) \mid l_j(x) \leq u(x) \leq r_j(x) \text{ a.e. in } \omega_j \subset \Omega\}, \quad (4)$$

where  $l_j$  and  $r_j$ ,  $j = 1, 2$ , are elements of  $L^\infty(\omega_j)$  and  $\omega_j$  are measurable subsets of  $\Omega$ .

The linear operators  $B_j : L^2(\omega_j) \rightarrow L^2(\Omega)$ ,  $j = 1, 2$ , are chosen depending on the particular application, and suitable assumptions on them will be given later. In the case of different controls defined on different subsets of  $\Omega$ , we have the extension operator given by

$$B_j u_j = \begin{cases} u_j & \text{in } \omega_j, \\ 0 & \text{in } \Omega \setminus \omega_j. \end{cases} \quad (5)$$

With this setting, we have  $y \in H_0^1(\Omega) \cap H^2(\Omega)$ .

In order to define a suitable solution concept for the multiobjective PDE optimization problems (1) and (2), first note that  $J_1$  depends on  $y$  and  $u_1$ , but not explicitly on  $u_2$ . Similarly,  $J_2$  depends on  $y$  and  $u_2$ , but not explicitly on  $u_1$ . Since  $y$  is given by a solution of a PDE which is controlled by  $u_1$  and  $u_2$ , we are essentially in the situation that  $J_1$  is a function of  $u_1$  and  $J_2$  is a function of  $u_2$ , i.e. each of the objectives has its own variables. This is the typical situation for which the solution concept by Nash can be used.

To make this more precise, let us introduce the reduced formulation of (1) and (2). We denote with  $y(u_1, u_2)$  the unique solution of the PDE as a function of  $(u_1, u_2)$ . With our setting, the mapping  $(u_1, u_2) \mapsto y(u_1, u_2)$  is affine and continuous. Then, we define the reduced cost functionals

$$\hat{J}_j(u_1, u_2) = J_j(y(u_1, u_2), u_1, u_2), \quad j = 1, 2.$$

We have that the mappings  $(u_1, u_2) \mapsto \hat{J}_j(u_1, u_2)$ ,  $j = 1, 2$ , are twice Fréchet differentiable [19, 34].

Using these reduced cost functionals, we obtain the following reduced (equivalent) formulation of (1) and (2). We have

$$\min_{u_1} \hat{J}_1(u_1, u_2) \quad \text{s.t.} \quad u_1 \in \mathcal{U}_{ad,1} \quad \text{and} \quad (6)$$

$$\min_{u_2} \hat{J}_2(u_1, u_2) \quad \text{s.t.} \quad u_2 \in \mathcal{U}_{ad,2}. \quad (7)$$

In this formulation,  $u_1$  is the variable in optimization problem (6) and  $u_2$  is the variable in optimization problem (7). That is, we are in the situation where both objectives control different sets of variables. Moreover, since we have multiple objectives that are in conflict with one another and therefore a multiobjective optimization problem does not have a single solution that could optimize all objectives simultaneously, the solution of multiobjective optimization problems should consist of all optimization functions that can best attain the prioritized objectives as good as possible. This leads to the following multiobjective solution concept by Nash.

**Definition 2.1** *A feasible element  $\bar{u} = (\bar{u}_1, \bar{u}_2) \in \mathcal{U}_{ad}$  is called a Nash solution or a Nash equilibrium of (6)–(7), if*

$$\begin{aligned} \hat{J}_1(\bar{u}_1, \bar{u}_2) &\leq \hat{J}_1(u_1, \bar{u}_2) \quad \text{for all } u_1 \in \mathcal{U}_{ad,1} \quad \text{and} \\ \hat{J}_2(\bar{u}_1, \bar{u}_2) &\leq \hat{J}_2(\bar{u}_1, u_2) \quad \text{for all } u_2 \in \mathcal{U}_{ad,2}. \end{aligned}$$

Hence, we have a Nash equilibrium if no benefit is attained by changing one control unilaterally while the other control is kept fixed. Similarly, we call a triple  $(\bar{y}, \bar{u}_1, \bar{u}_2)$  a Nash solution of the original multiobjective optimization problem from (1), (2) if  $(\bar{u}_1, \bar{u}_2)$  is a Nash equilibrium of the reduced formulation (6), (7) and  $\bar{y} = y(\bar{u}_1, \bar{u}_2)$  holds.

### 3 Existence and Uniqueness of a Nash Equilibrium

In this section, we show that there is a unique Nash equilibrium of the problem (1)–(2). The analysis is motivated by the one in [29]. To this end, we come back to the reduced formulation (6)–(7). It is not difficult to see that the mappings  $u_1 \mapsto \hat{J}_1(u_1, \bar{u}_2)$  and  $u_2 \mapsto \hat{J}_2(\bar{u}_1, u_2)$  are convex for fixed  $\bar{u}_2$  and  $\bar{u}_1$ , respectively. Hence  $(\bar{u}_1, \bar{u}_2)$  is a Nash equilibrium if and only if this pair satisfies the following

$$\begin{aligned} \frac{\partial \hat{J}_1}{\partial u_1}(\bar{u}_1, \bar{u}_2)(v_1 - \bar{u}_1) &\geq 0 \quad \forall v_1 \in \mathcal{U}_{ad,1}, \\ \frac{\partial \hat{J}_2}{\partial u_2}(\bar{u}_1, \bar{u}_2)(v_2 - \bar{u}_2) &\geq 0 \quad \forall v_2 \in \mathcal{U}_{ad,2}. \end{aligned}$$

In order to get a convenient expression for the gradient of  $\hat{J}_j$ , we introduce the functions  $p_j \in H^2(\Omega) \cap H_0^1(\Omega)$ ,  $j = 1, 2$ , as the unique solution to the adjoint equations

$$-\Delta p_j = -(y - z_j) \quad \text{in } \Omega, \quad p_j = 0 \quad \text{on } \partial\Omega, \quad (8)$$

where  $y = y(\bar{u})$ . Then the gradient of  $\hat{J}_j$  can be represented as follows

$$\frac{\partial \hat{J}_j}{\partial u_j}(\bar{u}_1, \bar{u}_2)(v_j - \bar{u}_j) = -(p_j, B_j(v_j - \bar{u}_j))_{L^2(\Omega)} + \nu_j (B_j \bar{u}_j, B_j(v_j - \bar{u}_j))_{L^2(\Omega)}$$

for  $j = 1, 2$ . Using the adjoint operator  $B_j^*$ , the necessary and sufficient first-order optimality conditions are given by

$$\begin{aligned} (\nu_1 B_1^* B_1 \bar{u}_1 - B_1^* p_1, v_1 - \bar{u}_1)_{L^2(\omega_1)} &\geq 0 \quad \forall v_1 \in \mathcal{U}_{ad,1}, \\ (\nu_2 B_2^* B_2 \bar{u}_2 - B_2^* p_2, v_2 - \bar{u}_2)_{L^2(\omega_2)} &\geq 0 \quad \forall v_2 \in \mathcal{U}_{ad,2}. \end{aligned}$$

Altogether, it follows that the solution to (1)–(2) is characterized by the following optimality system. We have

$$\begin{aligned} -\Delta y &= B_1 u_1 + B_2 u_2 + f && \text{in } \Omega, \\ y &= 0 && \text{on } \partial\Omega, \\ -\Delta p_1 &= -(y - z_1) && \text{in } \Omega, \\ p_1 &= 0 && \text{on } \partial\Omega, \\ -\Delta p_2 &= -(y - z_2) && \text{in } \Omega, \\ p_2 &= 0 && \text{on } \partial\Omega, \\ (\nu_1 B_1^* B_1 u_1 - B_1^* p_1, v_1 - u_1) &\geq 0 && \text{for all } v_1 \in \mathcal{U}_{ad,1}, \\ (\nu_2 B_2^* B_2 u_2 - B_2^* p_2, v_2 - u_2) &\geq 0 && \text{for all } v_2 \in \mathcal{U}_{ad,2}. \end{aligned} \quad (9)$$

Next, we define the Hilbert space  $H := L^2(\omega_1) \times L^2(\omega_2)$  and the induced scalar product  $(u, v)_H := (u_1, v_1)_{L^2(\omega_1)} + (u_2, v_2)_{L^2(\omega_2)}$  for all  $u, v \in H$ , where  $u = (u_1, u_2)$  and  $v = (v_1, v_2)$ . Then, we can rewrite (9) as follows

$$(\mathcal{A}\bar{u} - b, v - \bar{u})_H \geq 0 \quad \forall v \in \mathcal{U}_{ad},$$

where we put  $b := (b_1, b_2)$ ,  $\bar{u} := (\bar{u}_1, \bar{u}_2)$ ,  $v := (v_1, v_2)$ , and the operator  $\mathcal{A} : H \rightarrow H$  is defined by

$$\mathcal{A}(u_1, u_2) := (\nu_1 B_1^* B_1 u_1 - B_1^* \tilde{p}_1, \nu_2 B_2^* B_2 u_2 - B_2^* \tilde{p}_2). \quad (10)$$

Here, for the given  $(u_1, u_2)$ , the  $\tilde{p}_1$  and  $\tilde{p}_2$  are obtained by first solving the following problem

$$-\Delta \tilde{y} = B_1 u_1 + B_2 u_2 \quad \text{in } \Omega, \quad \tilde{y} = 0 \quad \text{on } \partial\Omega, \quad (11)$$

and in the next step, we compute  $\tilde{p}_1$  and  $\tilde{p}_2$  solving the equations

$$\Delta \tilde{p}_1 = \tilde{y} \quad \text{and} \quad \Delta \tilde{p}_2 = \tilde{y}, \quad (12)$$

with homogeneous Dirichlet boundary conditions. Notice that  $\tilde{p}_1$  and  $\tilde{p}_2$  coincide in our case. However, they would be different, in general, considering different tracking functionals.

We have that  $\mathcal{A}$  results to be the linear part of the optimality condition operator which is an affine mapping. In fact, the inhomogeneous term  $b = (b_1, b_2)$  is defined in terms of  $f$  and the target functions  $z_1$  and  $z_2$ , and it is zero when these functions are zero. Specifically, the construction of  $b_1$  and  $b_2$  proceeds as follows. Define  $\hat{y}$  as a solution of  $-\Delta \hat{y} = f$  with homogeneous Dirichlet boundary conditions. Then,  $b_1$  and  $b_2$  are given by  $b_j = B_j^* \hat{p}_j$  ( $j = 1, 2$ ), where  $\hat{p}_j$  ( $j = 1, 2$ ) solve the equation  $-\Delta \hat{p}_j = -(\hat{y} - z_j)$  with homogeneous Dirichlet boundary conditions.

Now, we define the mapping  $a : H \times H \rightarrow \mathbb{R}$  by

$$a(u, v) := (\mathcal{A}u - b, v)_H \quad \forall u, v \in H. \quad (13)$$

Then  $\bar{u} = (\bar{u}_1, \bar{u}_2)$  is a Nash equilibrium if and only if it satisfies the variational inequality

$$a(\bar{u}, v - \bar{u}) \geq 0 \quad \forall v \in \mathcal{U}_{ad}. \quad (14)$$

The central properties of the mapping  $a$  are summarized in the following result.

**Proposition 3.1** *Suppose that the mappings  $B_j^* B_j$  ( $j = 1, 2$ ) are coercive and the sets  $\omega_j$  satisfy*

$$\lambda(\text{supp}(B_j v) \cap \omega_k) = 0 \quad \forall v \in L^2(\omega_j), \quad j \neq k. \quad (15)$$

where  $\lambda(\cdot)$  denotes the Lebesgue-measure of a set. Then the mapping  $a : H \times H \rightarrow \mathbb{R}$  from (10)–(13) is bilinear, continuous, and coercive.

**Proof.** Clearly the operator  $\mathcal{A}$  is linear, bounded, and continuous in  $H$ . Therefore, we immediately obtain that  $a$  is bilinear and continuous. We show that  $a$  is coercive.

For this purpose, we consider the following

$$\begin{aligned} (\mathcal{A}(u_1, u_2), (u_1, u_2))_H &= \nu_1 (B_1^* B_1 u_1, u_1)_{L^2(\omega_1)} + \nu_2 (B_2^* B_2 u_2, u_2)_{L^2(\omega_2)} \\ &\quad - \int_{\omega_1} (B_1^* \tilde{p}_1) u_1 dx - \int_{\omega_2} (B_2^* \tilde{p}_2) u_2 dx. \end{aligned} \quad (16)$$

In view of the assumed coercivity of  $B_j^* B_j$ , the statement follows if we are able to show that the last two terms in (16) are nonnegative. In fact, we have

$$\begin{aligned} - \int_{\omega_1} (B_1^* \tilde{p}_1) u_1 dx &= - \int_{\omega_1} \tilde{p}_1 (B_1 u_1) dx = - \int_{\omega_1} \tilde{p}_1 (-\Delta \tilde{y} - B_2 u_2) dx \\ &\stackrel{(15)}{=} - \int_{\omega_1} \tilde{p}_1 (-\Delta \tilde{y}) dx = \int_{\omega_1} (\Delta \tilde{p}_1) \tilde{y} dx = \int_{\omega_1} \tilde{y} \tilde{y} dx, \end{aligned} \quad (17)$$

where  $\tilde{y}$  is the solution to (11) with  $u_1$  and  $u_2$  in the right-hand side. Similarly, we prove that  $-\int_{\omega_2} (B_2^* \tilde{p}_2) u_2 dx$  is nonnegative.  $\square$

The central result of this section is a direct consequence of Proposition 3.1.

**Theorem 3.2** *There exists a unique Nash equilibrium of the reduced problem (6)–(7) (and therefore also a unique Nash solution to (1)–(2)).*

**Proof.** Recall that  $\bar{u} = (\bar{u}_1, \bar{u}_2)$  is a Nash equilibrium of (6)–(7) if and only if this pair satisfies the variational inequality (14). However, since  $a$  is bilinear, continuous and coercive in view of Proposition 3.1, it follows from the Lions-Stampacchia-Theorem (see, e.g., [1, 21]) that this variational inequality has a unique solution.  $\square$

## 4 Finite Difference Discretization

In this section, we discuss a finite difference discretization of the multiobjective linear optimality system (9). To this end, we consider the simpler case without constraints on the controls. Then we can use the resulting optimality conditions to eliminate the control functions by expressing them in terms of the state and the adjoint variables.

Specifically, consider a sequence of grids  $\{\Omega_h\}_{h>0}$  given by

$$\Omega_h = \{\mathbf{x} \in \mathbb{R}^2 \mid x_i = s_i h, s_i \in \mathbb{Z}\} \cap \Omega.$$

We assume that  $\Omega$  is a rectangular domain and that the values of the mesh size  $h$  are chosen such that the boundaries of  $\Omega$  coincide with grid lines. For grid functions  $v_h$  and  $w_h$  defined on  $\Omega_h$  we introduce the discrete  $L^2$ -scalar product

$$(v_h, w_h)_{L_h^2} = h^2 \sum_{\mathbf{x} \in \Omega_h} v_h(\mathbf{x}) w_h(\mathbf{x}),$$

with associated norm  $|v_h|_0 = (v_h, v_h)_{L_h^2}^{1/2}$ . We also use  $|v_h|_\infty = \max_{\mathbf{x} \in \Omega_h} |v_h(\mathbf{x})|$ .

First-order backward and forward partial derivatives of  $v_h$  in the  $x_i$  direction are denoted by  $\partial_i^-$  and  $\partial_i^+$ , respectively, and given by

$$\partial_i^- v_h(\mathbf{x}) = \frac{v_h(\mathbf{x}) - v_h(\mathbf{x} - \hat{i} h)}{h} \quad \text{and} \quad \partial_i^+ v_h(\mathbf{x}) = \frac{v_h(\mathbf{x} + \hat{i} h) - v_h(\mathbf{x})}{h},$$

where  $\hat{i}$  denotes the  $i$  coordinate direction vector and  $v_h$  is extended by 0 on grid points outside of  $\Omega$ ; see [16]. In this framework, the discrete  $H^1$ -norm is given by

$$|v_h|_1 = \left( |v_h|_0^2 + \sum_{i=1}^2 |\partial_i^- v_h|_0^2 \right)^{1/2}.$$

The spaces  $L_h^2$  and  $H_h^1$  consist of the sets of grid functions  $v_h$  endowed with  $|v_h|_0$ , respectively  $|v_h|_1$ , as norm. For the definition of  $H_h^2$  we refer to [16], as well.

Functions in  $L^2(\Omega)$  and  $H^1(\Omega)$  are approximated by grid functions defined through their mean values with respect to elementary cells  $[x_1 - \frac{h}{2}, x_1 + \frac{h}{2}] \times [x_2 - \frac{h}{2}, x_2 + \frac{h}{2}]$ ; see [16] for more details.

The restriction operator  $R_h : H^2(\Omega) \cap H_0^1(\Omega) \rightarrow H_h^2$  is defined by

$$(R_h v)(x, y) = \frac{1}{h^2} \int_{-h/2}^{h/2} \int_{-h/2}^{h/2} v(x + \xi, y + \eta) d\xi d\eta.$$

In the following, this operator is also used as mapping  $R_h : H^1 \rightarrow H_h^1$ . For  $L^2$  functions, we have the restriction operator  $\tilde{R}_h : L^2(\Omega) \rightarrow L_h^2$  where

$$(\tilde{R}_h v)(x, y) = \frac{1}{h^2} \int_{-h/2}^{h/2} \int_{-h/2}^{h/2} \int_{-h/2}^{h/2} \int_{-h/2}^{h/2} v(x + \xi + \xi', y + \eta + \eta') d\xi d\xi' d\eta d\eta'.$$

Alternatively, for sufficiently smooth functions  $v \in C^k(\bar{\Omega})$  (resp.  $f \in C^k(\Omega)$ ),  $k = 0, 1, \dots$ , we use the restriction operators  $(R_h v)(x) = v(x)$  (resp.  $(\tilde{R}_h f)(x) = f(x)$ ) on  $\bar{\Omega}_h$  (resp.  $\Omega_h$ ).

The second-order five-point approximation to the Laplacian with homogeneous Dirichlet boundary conditions is defined by  $\Delta_h = \partial_1^+ \partial_1^- + \partial_2^+ \partial_2^-$ . We have the following consistency result

$$|\Delta_h R_h v - \tilde{R}_h \Delta v|_\infty \leq c h^2 \|v\|_{C^4(\bar{\Omega})}; \quad (18)$$

see, e.g., [16].

Next, an a priori estimate of the accuracy of solutions to the optimality system (9) without constraints on the control is discussed. The last equation in (9) then becomes  $\nu_j B_j^* B_j u_j - B_j^* p_j = 0$ , which we can use in the first equation of the system to eliminate the control variable  $u$ . After discretization, we have the following discrete optimality system:

$$-\Delta_h y^h - B_1^h u_1^h - B_2^h u_2^h = f^h \quad \text{in } \Omega_h, \quad (19)$$

$$-\Delta_h p_1^h + y^h = z_1^h \quad \text{in } \Omega_h, \quad (20)$$

$$-\Delta_h p_2^h + y^h = z_2^h \quad \text{in } \Omega_h, \quad (21)$$

$$\nu_1 B_1^{h,*} B_1^h u_1^h - B_1^{h,*} p_1^h = 0 \quad \text{in } \omega_{h,1}, \quad (22)$$

$$\nu_2 B_2^{h,*} B_2^h u_2^h - B_2^{h,*} p_2^h = 0 \quad \text{in } \omega_{h,2}, \quad (23)$$

where  $f^h = \tilde{R}_h f$  and  $z_j^h = \tilde{R}_h z_j$ .

We assume from now on that the discrete linear operators  $B_j^h$  are injective, i.e. of full column rank. This corresponds to the coercivity condition of  $B_j^* B_j$  used in Proposition 3.1. We then define

$$Q_j^h := B_j^h (B_j^{h,*} B_j^h)^{-1} B_j^{h,*}, \quad j = 1, 2.$$

Notice that the  $Q_j^h$  are projectors onto  $L_h^2(\omega_h)$ . It follows that the  $I_h - Q_j^h$ ,  $j = 1, 2$  are also projectors. We have

$$\nu_1 B_1^h u_1^h = Q_1^h p_1^h \quad \text{and} \quad \nu_2 B_2^h u_2^h = Q_2^h p_2^h.$$

Next, we use this fact to eliminate the control variables from the optimality system. For simplicity, we take  $\nu_1 = \nu_2 = \nu$ ; we obtain

$$-\nu \Delta_h y^h - Q_1^h p_1^h - Q_2^h p_2^h = \nu f^h \quad \text{in } \Omega_h, \quad (24)$$

$$-\Delta_h p_1^h + y^h = z_1^h \quad \text{in } \Omega_h, \quad (25)$$

$$-\Delta_h p_2^h + y^h = z_2^h \quad \text{in } \Omega_h. \quad (26)$$

For the purpose of our analysis, we add and subtract  $p_1^h$  and  $p_2^h$  in (24) and consider the  $L_h^2$ -inner product of this state equation with the state variable. Further, we consider the  $L_h^2$ -inner product of the first adjoint equation with  $p_1^h$  and of the second adjoint equation with  $p_2^h$ . We then obtain

$$\begin{aligned} & \nu(-\Delta_h y^h, y^h)_{L_h^2} - (p_1^h, y^h)_{L_h^2} - (p_2^h, y^h)_{L_h^2} \\ & + ((I_h - Q_1^h) p_1^h, y^h)_{L_h^2} + ((I_h - Q_2^h) p_2^h, y^h)_{L_h^2} = \nu(f^h, y^h)_{L_h^2}, \end{aligned} \quad (27)$$

$$(-\Delta_h p_1^h, p_1^h)_{L_h^2} + (y^h, p_1^h)_{L_h^2} = (z_1^h, p_1^h)_{L_h^2}, \quad (28)$$

$$(-\Delta_h p_2^h, p_2^h)_{L_h^2} + (y^h, p_2^h)_{L_h^2} = (z_2^h, p_2^h)_{L_h^2}. \quad (29)$$

Summing up the three equations (27), (28), (29), we get

$$\begin{aligned} & \nu(-\Delta_h y^h, y^h)_{L_h^2} + (-\Delta_h p_1^h, p_1^h)_{L_h^2} + (-\Delta_h p_2^h, p_2^h)_{L_h^2} \\ & + ((I_h - Q_1^h) p_1^h, y^h)_{L_h^2} + ((I_h - Q_2^h) p_2^h, y^h)_{L_h^2} \\ & = \nu(f^h, y^h)_{L_h^2} + (z_1^h, p_1^h)_{L_h^2} + (z_2^h, p_2^h)_{L_h^2}. \end{aligned}$$

Noting that  $(-\Delta_h v_h, v_h)_{L_h^2} = (\nabla_h v_h, \nabla_h v_h)_{L_h^2} = \sum_{i=1}^2 |\partial_i^- v_h|_0^2$ , recalling that  $I_h - Q_j^h$  is a projection operator, and using the Cauchy-Schwarz inequality yields

$$\nu |\nabla_h y^h|_0^2 + |\nabla_h p_1^h|_0^2 + |\nabla_h p_2^h|_0^2 \leq \nu |f^h|_0 |y^h|_0 + |z_1^h|_0 |p_1^h|_0 + |z_2^h|_0 |p_2^h|_0 + |y^h|_0 |p_1^h|_0 + |y^h|_0 |p_2^h|_0,$$

where the operator  $\nabla_h$  is defined by  $\nabla_h f := (\partial_1^- f, \partial_2^- f)$ . Now, we need the following result from [33].

**Lemma 4.1 (Poincaré–Friedrichs inequality for finite differences)** For any grid function  $v_h$ , there exists a constant  $c_*$ , independent of  $v_h$  and  $h$ , such that

$$|v_h|_0^2 \leq c_* \sum_{i=1}^2 |\partial_i^- v_h|_0^2. \quad (30)$$

Lemma 4.1 immediately gives

$$\frac{1}{c_*} \{ \nu |y^h|_0^2 + |p_1^h|_0^2 + |p_2^h|_0^2 \} \leq \nu |f^h|_0 |y^h|_0 + |z_1^h|_0 |p_1^h|_0 + |z_2^h|_0 |p_2^h|_0 + |y^h|_0 |p_1^h|_0 + |y^h|_0 |p_2^h|_0$$

Applying the Cauchy inequality  $|ab| \leq \frac{a^2}{2} + \frac{b^2}{2}$  and assuming  $c_*$  sufficiently small, we obtain

$$\nu |y^h|_0^2 + |p_1^h|_0^2 + |p_2^h|_0^2 \leq c (|f^h|_0^2 + |z_1^h|_0^2 + |z_2^h|_0^2), \quad (31)$$

where  $c$  is a positive constant depending on  $\nu$  and  $c_*$ .

Using (31), we are now able to determine the degree of accuracy of the optimal solution. For this purpose, notice that (24)–(26) hold true with  $y^h$ ,  $p_1^h$ , and  $p_2^h$  replaced by their respective error functions, and with  $f^h$  and  $z_1^h$  and  $z_2^h$  replaced by the truncation error for  $\Delta_h$  estimated by (18). These statements are summarized in the following theorem where we explicitly consider possibly different  $\nu_1$  and  $\nu_2$ .

**Theorem 4.2** *Let  $y \in C^4(\bar{\Omega})$ , and  $p \in C^4(\bar{\Omega})$ , be solutions of (9) without constraints on the controls, assume  $B_j^h$  has full column rank, and let  $y^h$ ,  $p_1^h$ , and  $p_2^h$  be solutions to (24)–(26). Then there exists a constant  $c$ , depending on  $\Omega$ , and independent of  $h$ , such that*

$$|y^h - R_h y|_0^2 + \frac{1}{\nu_1} |p_1^h - R_h p_1|_0^2 + \frac{1}{\nu_2} |p_2^h - R_h p_2|_0^2 \leq c h^4 \left( \|y\|_{C^4(\bar{\Omega})}^2 + \|p_1\|_{C^4(\bar{\Omega})}^2 + \|p_2\|_{C^4(\bar{\Omega})}^2 \right).$$

**Remark 4.3** The second-order accuracy estimate stated in Theorem 4.2 can also be proved in the context of finite differences assuming  $y, p \in H_0^1(\Omega) \cap H^3(\Omega)$ ; see [3]. In the context of finite elements, we could prove that the same order of accuracy holds with  $y, p \in H_0^1(\Omega) \cap H^2(\Omega)$ .

Theorem 4.2 states second-order accuracy of the solution of the finite-difference approximation to (9) assuming no constraints on the control. On the other hand, in the presence of active constraints, the analysis given above does not hold. In this case, in the context of finite differences and using the analysis given in [2, 25] it is possible to prove only a first-order accuracy estimate. In the context of finite elements it would be possible to extend results given in [26, 31, 34], for the case of single-objective optimization, to prove  $\mathcal{O}(h^2)$  convergence for the state and the adjoint variables and  $\mathcal{O}(h^{3/2})$  for the constrained control variables.

## 5 The Semismooth Newton Method

In this section, we discuss the semismooth Newton method to solve our discrete multiobjective optimal control problem. Alternatively, it might be possible to develop an infinite-dimensional semismooth Newton method like in [17, 19, 35] in other contexts. Here, however, we follow the first discretize then optimize approach [18]. We assume a square domain with a uniform grid with mesh size  $h = 1/(N + 1) > 0$  where  $N$  is the number of interior grid points in one direction. For simplicity of notation, we write  $B_1, B_2, u_1, u_2, \dots$  instead of  $B_1^h, B_2^h, u_1^h, u_2^h, \dots$ . That is we omit the superscript  $h$  for the discretized matrices and vectors. Notice that the finite-difference matrix  $A$  of the negative Laplacian is nonsingular, and therefore we have the following

$$Ay = B_1u_1 + B_2u_2 + f \iff y = A^{-1}B_1u_1 + A^{-1}B_2u_2 + A^{-1}f.$$

Correspondingly, the discretized reduced objectives are given by

$$\begin{aligned} \hat{J}_1(u_1, u_2) &= \frac{1}{2} \|A^{-1}B_1u_1 + A^{-1}B_2u_2 + A^{-1}f - z_1\|_2^2 + \frac{\nu_1}{2} \|B_1u_1\|_2^2 \quad \text{and} \\ \hat{J}_2(u_1, u_2) &= \frac{1}{2} \|A^{-1}B_1u_1 + A^{-1}B_2u_2 + A^{-1}f - z_2\|_2^2 + \frac{\nu_2}{2} \|B_2u_2\|_2^2, \end{aligned}$$

respectively, whereas the (box) constraints of the controls  $u_1$  and  $u_2$  read as follows

$$g^1(u_1, u_2) := \begin{pmatrix} l_1 - u_1 \\ u_1 - r_1 \end{pmatrix} \leq 0 \quad \text{and} \quad g^2(u_1, u_2) := \begin{pmatrix} l_2 - u_2 \\ u_2 - r_2 \end{pmatrix} \leq 0.$$

Now, consider the following optimization problem

$$\min_{u_1} \hat{J}_1(u_1, u_2) \quad \text{s.t.} \quad l_1 - u_1 \leq 0, \quad u_1 - r_1 \leq 0.$$

The Lagrangian of this problem is given by

$$\frac{1}{2} \|A^{-1}B_1u_1 + A^{-1}B_2u_2 + A^{-1}f - z_1\|_2^2 + \frac{\nu_1}{2} \|B_1u_1\|_2^2 + \begin{pmatrix} \lambda_{l_1} \\ \lambda_{r_1} \end{pmatrix}^T \begin{pmatrix} l_1 - u_1 \\ u_1 - r_1 \end{pmatrix},$$

with suitable Lagrange multipliers  $\lambda_{l_1}, \lambda_{r_1}$ . Hence, the KKT conditions are as follows

$$\begin{aligned} B_1^T A^{-1} [A^{-1}B_1u_1 + A^{-1}B_2u_2 + A^{-1}f - z_1] + \nu_1 B_1^T B_1 u_1 - \lambda_{l_1} + \lambda_{r_1} &= 0, \\ l_{1,i} - u_{1,i} \leq 0, \quad \lambda_{l_{1,i}} \geq 0, \quad \lambda_{l_{1,i}} [l_{1,i} - u_{1,i}] &= 0 \quad \forall i, \\ u_{1,i} - r_{1,i} \leq 0, \quad \lambda_{r_{1,i}} \geq 0, \quad \lambda_{r_{1,i}} [u_{1,i} - r_{1,i}] &= 0 \quad \forall i. \end{aligned}$$

Similarly, for the control  $u_2$  we obtain the corresponding KKT conditions

$$\begin{aligned} B_2^T A^{-1} [A^{-1}B_1u_1 + A^{-1}B_2u_2 + A^{-1}f - z_2] + \nu_2 B_2^T B_2 u_2 - \lambda_{l_2} + \lambda_{r_2} &= 0, \\ l_{2,i} - u_{2,i} \leq 0, \quad \lambda_{l_{2,i}} \geq 0, \quad \lambda_{l_{2,i}} [l_{2,i} - u_{2,i}] &= 0 \quad \forall i, \\ u_{2,i} - r_{2,i} \leq 0, \quad \lambda_{r_{2,i}} \geq 0, \quad \lambda_{r_{2,i}} [u_{2,i} - r_{2,i}] &= 0 \quad \forall i \end{aligned}$$

with suitable multipliers  $\lambda_2, \lambda r_2$ . Introducing slack variables  $wl_1, wr_1$  for control 1 and  $wl_2, wr_2$  for control 2, we obtain from the definition of  $y$  the combined KKT conditions

$$\begin{aligned}
B_1^T A^{-1}[y - z_1] + \nu_1 B_1^T B_1 u_1 - \lambda l_1 + \lambda r_1 &= 0, \\
B_2^T A^{-1}[y - z_2] + \nu_2 B_2^T B_2 u_2 - \lambda l_2 + \lambda r_2 &= 0, \\
Ay - B_1 u_1 - B_2 u_2 - f &= 0, \\
l_1 - u_1 + wl_1 &= 0, \\
u_1 - r_1 + wr_1 &= 0, \\
l_2 - u_2 + wl_2 &= 0, \\
u_2 - r_2 + wr_2 &= 0, \\
wl_1 \circ \lambda l_1 &= 0, \\
wr_1 \circ \lambda r_1 &= 0, \\
wl_2 \circ \lambda l_2 &= 0, \\
wr_2 \circ \lambda r_2 &= 0, \\
wl_1, wr_1, wl_2, wr_2, \lambda l_1, \lambda r_1, \lambda l_2, \lambda r_2 &\geq 0
\end{aligned}$$

where  $\circ$  denotes the Hadamard (componentwise) product of two vectors. Similar to (8), we next define the adjoint variables

$$p_1 := -A^{-1}[y - z_1] \quad \text{and} \quad p_2 := -A^{-1}[y - z_2],$$

respectively. Then we can write the combined KKT-conditions as follows:

$$\begin{aligned}
\nu_1 B_1^T B_1 u_1 - B_1^T p_1 - \lambda l_1 + \lambda r_1 &= 0, \\
\nu_2 B_2^T B_2 u_2 - B_2^T p_2 - \lambda l_2 + \lambda r_2 &= 0, \\
Ay - B_1 u_1 - B_2 u_2 - f &= 0, \\
Ap_1 + y - z_1 &= 0, \\
Ap_2 + y - z_2 &= 0, \\
l_1 - u_1 + wl_1 &= 0, \\
u_1 - r_1 + wr_1 &= 0, \\
l_2 - u_2 + wl_2 &= 0, \\
u_2 - r_2 + wr_2 &= 0, \\
wl_1 \circ \lambda l_1 &= 0, \\
wr_1 \circ \lambda r_1 &= 0, \\
wl_2 \circ \lambda l_2 &= 0, \\
wr_2 \circ \lambda r_2 &= 0, \\
wl_1, wr_1, wl_2, wr_2, \lambda l_1, \lambda r_1, \lambda l_2, \lambda r_2 &\geq 0.
\end{aligned}$$

Here, the dimensions of the corresponding data and the variables are (with  $n = N^2$ ):

$$A \in \mathbb{R}^{n \times n}, \quad B_1 \in \mathbb{R}^{n \times n_1}, \quad B_2 \in \mathbb{R}^{n \times n_2},$$

$$f, p_1, p_2 \in \mathbb{R}^n, \quad u_1, w l_1, w r_1, \lambda_1, \lambda r_1 \in \mathbb{R}^{n_1}, \quad u_2, w l_2, w r_2, \lambda_2, \lambda r_2 \in \mathbb{R}^{n_2}.$$

The vector of variables is

$$z := (y, u_1, u_2, p_1, p_2, \lambda_1, \lambda r_1, \lambda_2, \lambda r_2, w l_1, w r_1, w l_2, w r_2).$$

Now, let

$$\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad \varphi(a, b) := \sqrt{a^2 + b^2} - a - b$$

denote the Fischer-Burmeister-function introduced in [13]. It has the important property that

$$\varphi(a, b) = 0 \iff a \geq 0, \quad b \geq 0, \quad ab = 0.$$

Hence the combined KKT-condition can be rewritten as an unconstrained system of non-linear equations

$$H(z) = 0,$$

where  $H : \mathbb{R}^d \longrightarrow \mathbb{R}^d$  is defined by

$$H(y, u_1, u_2, p_1, p_2, \lambda_1, \lambda r_1, \lambda_2, \lambda r_2, w l_1, w r_1, w l_2, w r_2) :=$$

$$\begin{pmatrix} \nu_1 B_1^T B_1 u_1 - B_1^T p_1 - \lambda_1 + \lambda r_1 \\ \nu_2 B_2^T B_2 u_2 - B_2^T p_2 - \lambda_2 + \lambda r_2 \\ Ay - B_1 u_1 - B_2 u_2 - f \\ Ap_1 + y - z_1 \\ Ap_2 + y - z_2 \\ l_1 - u_1 + w l_1 \\ u_1 - r_1 + w r_1 \\ l_2 - u_2 + w l_2 \\ u_2 - r_2 + w r_2 \\ \phi(\lambda_1, w l_1) \\ \phi(\lambda r_1, w r_1) \\ \phi(\lambda_2, w l_2) \\ \phi(\lambda r_2, w r_2) \end{pmatrix}$$

with

$$\phi(a, b) := (\varphi(a_1, b_1), \dots, \varphi(a_m, b_m)) \quad \text{for } m \in \{n_1, n_2\}.$$

Since  $\varphi$  is nonsmooth, the mapping  $H$  is also nonsmooth in general. However, we have the following result that can be derived in a standard way.

**Theorem 5.1** *The mapping  $H$  is strongly semismooth.*

The semismooth Newton method from [27, 28] can therefore be applied to the system of equations  $H(z) = 0$ . In order to present a globalized version of the locally convergent semismooth Newton method, we will also exploit the corresponding merit function

$$\Psi(z) := \frac{1}{2} \|H(z)\|_2^2.$$

Despite the nonsmoothness of  $H$ , it turns out that this merit function is continuously differentiable. We formulate this in the following result whose proof is again standard and therefore omitted here.

**Theorem 5.2** *The merit function  $\Psi : \mathbb{R}^d \rightarrow \mathbb{R}$  is continuously differentiable.*

Exploiting the smoothness of  $\Psi$ , we are now able to restate the globalized semismooth Newton method from [7], adapted to our context. (In this and the next two sections  $\Delta$  denotes differences and not the Laplacian operator.)

**Algorithm 5.3** *(Globalized Semismooth Newton Method)*

(S.1) Choose a starting point  $z^0 = (y^0, u_1^0, u_2^0, \dots, wr_2^0)$ , parameters  $\rho > 0, p > 2, \beta \in (0, 1), \sigma \in (0, 1/2), \varepsilon > 0$ , and set  $k := 0$ .

(S.2) If  $\|\nabla\Psi(z^k)\| \leq \varepsilon$ : STOP.

(S.3) Choose an element  $V_k \in \partial H(z^k)$ .

(S.4) Compute a solution  $\Delta z^k$  of the linear system of equations  $V_k \Delta z = -H(z^k)$ . If this system is not solvable or the solution does not satisfy the sufficient decrease condition  $\nabla\Psi(z^k)^T \Delta z^k \leq -\rho \|\Delta z^k\|^p$ , then set  $\Delta z^k := -\nabla\Psi(z^k)$ .

(S.5) Compute a stepsize  $t_k = \max\{\beta^\ell \mid \ell = 0, 1, 2, \dots\}$  satisfying the Armijo condition

$$\Psi(z^k + t_k \Delta z^k) \leq \Psi(z^k) + \sigma t_k \nabla\Psi(z^k)^T \Delta z^k.$$

(S.6) Set  $z^{k+1} := z^k + t_k \Delta z^k, k \leftarrow k + 1$ , and go to (S.1).

The corresponding local and global convergence result from [7], once again slightly adapted to our case, then reads as follows.

**Theorem 5.4** *The following statements hold:*

- (a) *Algorithm 5.3 is well-defined.*
- (b) *Every accumulation point of a sequence  $\{z^k\}$  generated by Algorithm 5.3 is a stationary point of  $\Psi$ .*
- (c) *If an accumulation point  $\bar{z}$  of the sequence  $\{z^k\}$  is such that all matrices  $V \in \partial H(\bar{z})$  are nonsingular, then  $\bar{z}$  is a solution of  $H(z) = 0$ , and the sequence  $\{z^k\}$  convergence locally quadratically to  $\bar{z}$ .*

The previous result raises the following important questions: 1) Is a stationary point of  $\Psi$  already a global minimum, hence a solution of  $H(z) = 0$  and, therefore, also a solution of our (discretized) Nash equilibrium problem? 2) Are all matrices  $V \in \partial H(\bar{z})$  at a solution  $\bar{z}$  automatically nonsingular? These two questions will be discussed in detail in the subsequent two sections.

## 6 A Stationary Point Result

The aim of this section is to show that a stationary point  $z^*$  of the merit function  $\Psi$  is already a global minimum of this function, hence a solution of the nonlinear system of equations  $H(z) = 0$  and, consequently, also a solution of the (discretized) Nash equilibrium problem. In view of Theorem 5.4, it then follows that every accumulation point of a sequence  $\{z^k\}$  generated by Algorithm 5.3 is a solution of the Nash equilibrium problem.

In order to prove the main result of this section, we need the generalized Jacobian of the mapping  $H$ . To this end, we first recall that the convex subdifferential (which is identical to the generalized gradient in this case) of the Fischer-Burmeister-function at an arbitrary point  $(a, b) \in \mathbb{R}^2$  is given by

$$\partial\varphi(a, b) = \begin{cases} \left( \frac{a}{\|(a,b)\|_2} - 1, \frac{b}{\|(a,b)\|_2} - 1 \right), & \text{if } (a, b) \neq (0, 0), \\ \{(\xi - 1, \zeta - 1) \mid \|(\xi, \zeta)\|_2 \leq 1\}, & \text{if } (a, b) = (0, 0). \end{cases} \quad (32)$$

Then we write  $D_a^{l_1}, D_b^{l_1}$  for the diagonal matrices

$$D_a^{l_1} = \text{diag}(a_1^{l_1}, \dots, a_{n_1}^{l_1}), \quad D_b^{l_1} = \text{diag}(b_1^{l_1}, \dots, b_{n_1}^{l_1})$$

with diagonal elements  $(a_i^{l_1}, b_i^{l_1}) \in \partial\varphi(\lambda_{1,i}, w_{1,i})$  for all  $i = 1, \dots, n_1$ . In a similar way, we define the corresponding diagonal matrices

$$D_a^{r_1}, D_b^{r_1}, D_a^{l_2}, D_b^{l_2}, D_a^{r_2}, D_b^{r_2}.$$

In view of the previous representation of the subdifferential of  $\varphi$ , it follows immediately that all diagonal matrices are negative semidefinite, and that the sum of each pair like  $(D_a^{l_1}, D_b^{l_1})$  is a negative definite diagonal matrix since the diagonal elements at the same position cannot be equal to zero at the same time. This simple observation will play some role in our subsequent analysis.

Using standard calculus rules for nonsmooth mappings from [6], it is now easy to see that each element  $V \in \partial H(z)$  has the following structure:

$$\begin{pmatrix} 0 & \nu_1 B_1^T B_1 & 0 & -B_1^T & 0 & -I & +I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \nu_2 B_2^T B_2 & 0 & -B_2^T & 0 & 0 & -I & +I & 0 & 0 & 0 & 0 \\ A & -B_1 & -B_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ I & 0 & 0 & A & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & A & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & +I & 0 & 0 & 0 \\ 0 & +I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & +I & 0 & 0 \\ 0 & 0 & -I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & +I & 0 \\ 0 & 0 & +I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & +I \\ 0 & 0 & 0 & 0 & 0 & D_a^{l_1} & 0 & 0 & 0 & D_b^{l_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & D_a^{r_1} & 0 & 0 & 0 & D_b^{r_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & D_a^{l_2} & 0 & 0 & 0 & D_b^{l_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & D_a^{r_2} & 0 & 0 & 0 & D_b^{r_2} \end{pmatrix}.$$

We know from the previous discussion that all the diagonal matrices  $D_a^{l_1}, D_b^{l_1}$  etc. occurring inside this matrix  $V$  are negative semidefinite. For the moment, we now assume, in addition, that

$$D_a^{l_1}, D_b^{l_1}, D_a^{r_1}, D_b^{r_1}, D_a^{l_2}, D_b^{l_2}, D_a^{r_2}, D_b^{r_2} \text{ are negative definite,} \quad (33)$$

hence nonsingular. In general, this assumption does not hold, but we will see later how to exploit this condition in the main result of this section.

**Proposition 6.1** *Given an arbitrary point*

$$z = (y, u_1, u_2, p_1, p_2, \lambda l_1, \lambda r_1, \lambda l_2, \lambda r_2, w l_1, w r_1, w l_2, w r_2)$$

*satisfying (33), each element  $V \in \partial H(z)$  of the generalized Jacobian of  $H$  at this point is nonsingular.*

**Proof.** We consider the homogeneous linear system of equations  $V \cdot \Delta z = 0$  with

$$\Delta z =: (\Delta y, \Delta u_1, \Delta u_2, \Delta p_1, \Delta p_2, \Delta \lambda l_1, \Delta \lambda r_1, \Delta \lambda l_2, \Delta \lambda r_2, \Delta w l_1, \Delta w r_1, \Delta w l_2, \Delta w r_2) \quad (34)$$

being appropriately partitioned. Taking into account the special structure of an arbitrary element  $V \in \partial H(z)$ , this can be rewritten as

$$\nu_1 B_1^T B_1 \Delta u_1 - B_1^T \Delta p_1 - \Delta \lambda l_1 + \Delta \lambda r_1 = 0, \quad (35)$$

$$\nu_2 B_2^T B_2 \Delta u_2 - B_2^T \Delta p_2 - \Delta \lambda l_2 + \Delta \lambda r_2 = 0, \quad (36)$$

$$A \Delta y - B_1 \Delta u_1 - B_2 \Delta u_2 = 0, \quad (37)$$

$$\Delta y + A \Delta p_1 = 0, \quad (38)$$

$$\Delta y + A \Delta p_2 = 0, \quad (39)$$

$$-\Delta u_1 + \Delta w l_1 = 0, \quad (40)$$

$$\Delta u_1 + \Delta w r_1 = 0, \quad (41)$$

$$-\Delta u_2 + \Delta w l_2 = 0, \quad (42)$$

$$\Delta u_2 + \Delta w r_2 = 0, \quad (43)$$

$$D_a^{l_1} \Delta \lambda l_1 + D_b^{l_1} \Delta w l_1 = 0, \quad (44)$$

$$D_a^{r_1} \Delta \lambda r_1 + D_b^{r_1} \Delta w r_1 = 0, \quad (45)$$

$$D_a^{l_2} \Delta \lambda l_2 + D_b^{l_2} \Delta w l_2 = 0, \quad (46)$$

$$D_a^{r_2} \Delta \lambda r_2 + D_b^{r_2} \Delta w r_2 = 0. \quad (47)$$

Using (40)–(43), we obtain

$$\Delta w l_1 = \Delta u_1, \quad \Delta w r_1 = -\Delta u_1, \quad \Delta w l_2 = \Delta u_2, \quad \Delta w r_2 = -\Delta u_2.$$

Substituting these expressions into the remaining equations yields

$$\nu_1 B_1^T B_1 \Delta u_1 - B_1^T \Delta p_1 - \Delta \lambda l_1 + \Delta \lambda r_1 = 0, \quad (48)$$

$$\nu_2 B_2^T B_2 \Delta u_2 - B_2^T \Delta p_2 - \Delta \lambda l_2 + \Delta \lambda r_2 = 0, \quad (49)$$

$$A \Delta y - B_1 \Delta u_1 - B_2 \Delta u_2 = 0, \quad (50)$$

$$\Delta y + A \Delta p_1 = 0, \quad (51)$$

$$\Delta y + A \Delta p_2 = 0, \quad (52)$$

$$D_a^{l_1} \Delta \lambda l_1 + D_b^{l_1} \Delta u_1 = 0, \quad (53)$$

$$D_a^{r_1} \Delta \lambda r_1 - D_b^{r_1} \Delta u_1 = 0, \quad (54)$$

$$D_a^{l_2} \Delta \lambda l_2 + D_b^{l_2} \Delta u_2 = 0, \quad (55)$$

$$D_a^{r_2} \Delta \lambda r_2 - D_b^{r_2} \Delta u_2 = 0. \quad (56)$$

Exploiting assumption (33), we may further solve equations (53)–(56) with respect to  $\Delta \lambda l_1, \Delta \lambda r_1, \Delta \lambda l_2, \Delta \lambda r_2$  and obtain

$$\begin{aligned} \Delta \lambda l_1 &= -(D_a^{l_1})^{-1} D_b^{l_1} \Delta u_1, \\ \Delta \lambda r_1 &= (D_a^{r_1})^{-1} D_b^{r_1} \Delta u_1, \\ \Delta \lambda l_2 &= -(D_a^{l_2})^{-1} D_b^{l_2} \Delta u_2, \\ \Delta \lambda r_2 &= (D_a^{r_2})^{-1} D_b^{r_2} \Delta u_2. \end{aligned}$$

Replacing these terms in (48)–(52), we get

$$\nu_1 B_1^T B_1 \Delta u_1 - B_1^T \Delta p_1 + (D_a^{l_1})^{-1} D_b^{l_1} \Delta u_1 + (D_a^{r_1})^{-1} D_b^{r_1} \Delta u_1 = 0, \quad (57)$$

$$\nu_2 B_2^T B_2 \Delta u_2 - B_2^T \Delta p_2 + (D_a^{l_2})^{-1} D_b^{l_2} \Delta u_2 + (D_a^{r_2})^{-1} D_b^{r_2} \Delta u_2 = 0, \quad (58)$$

$$A \Delta y - B_1 \Delta u_1 - B_2 \Delta u_2 = 0, \quad (59)$$

$$\Delta y + A \Delta p_1 = 0, \quad (60)$$

$$\Delta y + A \Delta p_2 = 0. \quad (61)$$

Note that (57) and (58) can be reformulated as

$$(\nu_1 B_1^T B_1 + D_1) \Delta u_1 - B_1^T \Delta p_1 = 0, \quad (62)$$

$$(\nu_2 B_2^T B_2 + D_2) \Delta u_2 - B_2^T \Delta p_2 = 0 \quad (63)$$

with suitable positive definite diagonal matrices  $D_1, D_2$ . Since  $A$  is nonsingular, it follows from (60) and (61) that

$$\Delta p_1 = -A^{-1} \Delta y, \quad \Delta p_2 = -A^{-1} \Delta y.$$

Inserting this into (62) and (63), we get

$$(\nu_1 B_1^T B_1 + D_1) \Delta u_1 + B_1^T A^{-1} \Delta y = 0, \quad (64)$$

$$(\nu_2 B_2^T B_2 + D_2) \Delta u_2 + B_2^T A^{-1} \Delta y = 0. \quad (65)$$

Furthermore, we obtain

$$\Delta y = A^{-1} B_1 \Delta u_1 + A^{-1} B_2 \Delta u_2$$

from (59). Substituting this expression into (64) and (65) yields

$$\begin{aligned}(\nu_1 B_1^T B_1 + D_1 + B_1^T A^{-1} A^{-1} B_1) \Delta u_1 + B_1^T A^{-1} A^{-1} B_2 \Delta u_2 &= 0, \\(\nu_2 B_2^T B_2 + D_2 + B_2^T A^{-1} A^{-1} B_2) \Delta u_2 + B_2^T A^{-1} A^{-1} B_1 \Delta u_1 &= 0.\end{aligned}$$

In matrix–vector notation, the previous two equations are equivalent to

$$\begin{pmatrix} \nu_1 B_1^T B_1 + D_1 + B_1^T A^{-2} B_1 & B_1^T A^{-2} B_2 \\ B_2^T A^{-2} B_1 & \nu_2 B_2^T B_2 + D_2 + B_2^T A^{-2} B_2 \end{pmatrix} \begin{pmatrix} \Delta u_1 \\ \Delta u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

which, in turn, can be rewritten as

$$\left[ \begin{pmatrix} \nu_1 B_1^T B_1 + D_1 & 0 \\ 0 & \nu_2 B_2^T B_2 + D_2 \end{pmatrix} + \begin{pmatrix} B_1^T & 0 \\ 0 & B_2^T \end{pmatrix} \begin{pmatrix} A^{-2} & A^{-2} \\ A^{-2} & A^{-2} \end{pmatrix} \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} \right] \begin{pmatrix} \Delta u_1 \\ \Delta u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Now the first matrix is obviously positive definite. We claim that the second matrix is positive semidefinite. To this end, it remains to show that

$$\begin{pmatrix} A^{-2} & A^{-2} \\ A^{-2} & A^{-2} \end{pmatrix}$$

is positive semidefinite. To see this, first recall that  $A$  is symmetric positive definite. Hence the inverse  $A^{-1}$  is also symmetric positive definite. Therefore, letting  $d = (d_1, d_2) \in \mathbb{R}^{2n}$  arbitrary and writing  $p = (p_1, p_2) := (A^{-1}d_1, A^{-1}d_2)$ , it follows that

$$\begin{aligned}(d_1^T, d_2^T) \begin{pmatrix} A^{-2} & A^{-2} \\ A^{-2} & A^{-2} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} &= d_1^T A^{-2} d_1 + 2d_1^T A^{-2} d_2 + d_2^T A^{-2} d_2 \\ &= p_1^T p_1 + 2p_1^T p_2 + p_2^T p_2 \\ &= (p_1 + p_2)^T (p_1 + p_2) \\ &= \|p_1 + p_2\|_2^2 \\ &\geq 0.\end{aligned}$$

Hence we obtain  $(\Delta u_1, \Delta u_2) = (0, 0)$ . This, in turn, successively implies  $\Delta z = 0$  so that the generalized Jacobian is indeed nonsingular.  $\square$

The previous result allows us to apply a standard trick from [12] in order to show that every stationary point of  $\Psi$  is already a global minimum of this function and, therefore, a solution of the Nash equilibrium problem.

**Theorem 6.2** *Every stationary point  $\bar{z}$  of the merit function  $\Psi(z) := \frac{1}{2} \|H(z)\|^2$  is a solution of the Nash equilibrium problem.*

**Proof.** Using Clarke’s generalized chain rule, it follows that  $0 = \nabla \Psi(\bar{z}) = V^T H(\bar{z})$  for some matrix  $V \in \partial H(\bar{z})$ . Now, we can assume without loss of generality that assumption (33) holds since zero entries in one of the diagonal matrices from (33) can only occur if the

corresponding entry of the vector  $H(z)$  is zero, hence these entries of the diagonal matrices can be modified without changing the gradient. Hence, Proposition 6.1 implies that  $V$  is nonsingular. However, the nonsingularity of  $V$  together with  $0 = V^T H(\bar{z})$  immediately gives  $H(\bar{z}) = 0$ . This, in turn, implies that the corresponding components of  $z$  are a solution of the underlying Nash equilibrium problem.  $\square$

## 7 A Nonsingularity Result

Here we want to show that all matrices  $V \in \partial H(\bar{z})$  are nonsingular at a solution  $\bar{z}$ . In view of Theorem 5.4, this implies that the semismooth Newton method is locally quadratically convergent.

To this end, we state the following assumption.

**Assumption 7.1** *The matrices  $B_1$  and  $B_2$  have orthonormal columns, i.e. it holds that  $B_1^T B_1 = I_{n_1}$  and  $B_2^T B_2 = I_{n_2}$ .*

This assumption holds, for example, for

$$B_1 := \begin{pmatrix} I_{n_1} \\ 0 \end{pmatrix}, \quad B_2 := \begin{pmatrix} 0 \\ I_{n_2} \end{pmatrix}. \quad (66)$$

This is precisely the situation that we will consider in our numerical experiments, and it is the only case discussed in [29] (where, however, the controls  $u_1$  and  $u_2$  are unconstrained).

Now, let  $\bar{z} = (y, u_1, u_2, p_1, p_2, \lambda_1, \lambda r_1, \lambda_2, \lambda r_2, w l_1, w r_1, w l_2, w r_2)$  be a solution of  $H(z) = 0$ , and let  $V \in \partial H(\bar{z})$  an arbitrary element from the generalized Jacobian of  $H$  at  $\bar{z}$ . Then  $V$  has precisely the structure as indicated before Proposition 6.1 except that, in addition, we have  $B_1^T B_1 = I_{n_1}$  and  $B_2^T B_2 = I_{n_2}$  in our particular situation, cf. Assumption 7.1. We need to show that this matrix  $V$  is nonsingular, and we want to verify this statement without the additional condition from (33). Unfortunately, the proof is quite involved and needs some technical notation.

To this end, we first recall that  $\bar{z}$  being a solution of  $H(z) = 0$ , we, in particular, have that  $u_1$  is feasible, hence, for any index  $i \in \{1, \dots, n_1\}$ , we either have  $u_{1,i} = l_{1,i}$  (the variable is equal to the lower bound) or  $u_{1,i} \in (l_{1,i}, r_{1,i})$  (the variable is inactive) or  $u_{1,i} = r_{1,i}$  (the variable is equal to the upper bound).

In the first case, where  $u_{1,i} = l_{1,i}$ , we automatically have  $w l_{1,i} = 0$  and  $w r_{1,i} > 0$ , which in turn implies  $\lambda r_{1,i} = 0$ . Since  $\lambda_1 \geq 0$ , we therefore have the two subcases  $\lambda l_{1,i} > 0$  and  $\lambda l_{1,i} = 0$ . On the other hand, the second case  $u_{1,i} \in (l_{1,i}, r_{1,i})$  automatically gives  $w l_{1,i} > 0$  and  $w r_{1,i} > 0$  which, in turn, yields  $\lambda l_{1,i} = 0$  and  $\lambda r_{1,i} = 0$ . Finally, in the third case  $u_{1,i} = r_{1,i}$ , we have  $w r_{1,i} = 0$  and  $w l_{1,i} > 0$ , which then gives  $\lambda l_{1,i} = 0$ . Using  $\lambda r_1 \geq 0$ , there are the two remaining subcases  $\lambda r_{1,i} > 0$  and  $\lambda r_{1,i} = 0$ . Summarizing this discussion, we see that the following five index sets form a partition of the set  $\{1, \dots, n_1\}$ :

$$\alpha_1 := \{i \mid \lambda l_{1,i} > 0, w l_{1,i} = 0 \text{ and } \lambda r_{1,i} = 0, w r_{1,i} > 0\},$$

$$\begin{aligned}
\beta_1 &:= \{i \mid \lambda_{1,i} = 0, \, wl_{1,i} = 0 \text{ and } \lambda r_{1,i} = 0, \, wr_{1,i} > 0\}, \\
\gamma_1 &:= \{i \mid \lambda_{1,i} = 0, \, wl_{1,i} > 0 \text{ and } \lambda r_{1,i} = 0, \, wr_{1,i} > 0\} \\
\delta_1 &:= \{i \mid \lambda_{1,i} = 0, \, wl_{1,i} > 0 \text{ and } \lambda r_{1,i} > 0, \, wr_{1,i} = 0\}, \\
\varepsilon_1 &:= \{i \mid \lambda_{1,i} = 0, \, wl_{1,i} > 0 \text{ and } \lambda r_{1,i} = 0, \, wr_{1,i} = 0\}.
\end{aligned}$$

A similar reasoning applied to the control  $u_2$  shows that the subsequent index sets form a partitioning of the set  $\{1, \dots, n_2\}$ :

$$\begin{aligned}
\alpha_2 &:= \{i \mid \lambda_{2,i} > 0, \, wl_{2,i} = 0 \text{ and } \lambda r_{2,i} = 0, \, wr_{2,i} > 0\}, \\
\beta_2 &:= \{i \mid \lambda_{2,i} = 0, \, wl_{2,i} = 0 \text{ and } \lambda r_{2,i} = 0, \, wr_{2,i} > 0\}, \\
\gamma_2 &:= \{i \mid \lambda_{2,i} = 0, \, wl_{2,i} > 0 \text{ and } \lambda r_{2,i} = 0, \, wr_{2,i} > 0\} \\
\delta_2 &:= \{i \mid \lambda_{2,i} = 0, \, wl_{2,i} > 0 \text{ and } \lambda r_{2,i} > 0, \, wr_{2,i} = 0\}, \\
\varepsilon_2 &:= \{i \mid \lambda_{2,i} = 0, \, wl_{2,i} > 0 \text{ and } \lambda r_{2,i} = 0, \, wr_{2,i} = 0\}.
\end{aligned}$$

Next, we recall that the generalized gradient  $\partial\varphi(a, b)$  of the Fischer-Burmeister-function is given by (32), and that the diagonal matrices  $D_a^{l_1}, \dots, D_b^{r_2}$  are defined by

$$\begin{aligned}
D_a^{l_1} &= \text{diag}(a_1^{l_1}, \dots, a_{n_1}^{l_1}), & D_b^{l_1} &= \text{diag}(b_1^{l_1}, \dots, b_{n_1}^{l_1}) && \text{with } (a_i^{l_1}, b_i^{l_1}) \in \partial\varphi(\lambda_{1,i}, wl_{1,i}), \\
D_a^{r_1} &= \text{diag}(a_1^{r_1}, \dots, a_{n_1}^{r_1}), & D_b^{r_1} &= \text{diag}(b_1^{r_1}, \dots, b_{n_1}^{r_1}) && \text{with } (a_i^{r_1}, b_i^{r_1}) \in \partial\varphi(\lambda r_{1,i}, wr_{1,i}), \\
D_a^{l_2} &= \text{diag}(a_1^{l_2}, \dots, a_{n_2}^{l_2}), & D_b^{l_2} &= \text{diag}(b_1^{l_2}, \dots, b_{n_2}^{l_2}) && \text{with } (a_i^{l_2}, b_i^{l_2}) \in \partial\varphi(\lambda_{2,i}, wl_{2,i}), \\
D_a^{r_2} &= \text{diag}(a_1^{r_2}, \dots, a_{n_2}^{r_2}), & D_b^{r_2} &= \text{diag}(b_1^{r_2}, \dots, b_{n_2}^{r_2}) && \text{with } (a_i^{r_2}, b_i^{r_2}) \in \partial\varphi(\lambda r_{2,i}, wr_{2,i}).
\end{aligned}$$

Taking into account the definition of these index sets as well as the particular structure of the subdifferential  $\partial\varphi$ , we immediately obtain the following representation of the above diagonal matrices.

**Lemma 7.2** *The diagonal matrices  $D_a^{l_1}, D_b^{l_1}, D_a^{r_1}, D_b^{r_1}$  have the following diagonal entries:*

	$\alpha_1$	$\beta_1$	$\gamma_1$	$\delta_1$	$\varepsilon_1$
$D_a^{l_1}$	$0_{\alpha_1\alpha_1}$	*	$-I_{\gamma_1\gamma_1}$	$-I_{\delta_1\delta_1}$	$-I_{\varepsilon_1\varepsilon_1}$
$D_b^{l_1}$	$-I_{\alpha_1\alpha_1}$	*	$0_{\gamma_1\gamma_1}$	$0_{\delta_1\delta_1}$	$0_{\varepsilon_1\varepsilon_1}$
$D_a^{r_1}$	$-I_{\alpha_1\alpha_1}$	$-I_{\beta_1\beta_1}$	$-I_{\gamma_1\gamma_1}$	$0_{\delta_1\delta_1}$	*
$D_b^{r_1}$	$0_{\alpha_1\alpha_1}$	$0_{\beta_1\beta_1}$	$0_{\gamma_1\gamma_1}$	$-I_{\delta_1\delta_1}$	*

Similarly, the diagonal matrices  $D_a^{l_2}, D_b^{l_2}, D_a^{r_2}, D_b^{r_2}$  consist of the following entries:

	$\alpha_2$	$\beta_2$	$\gamma_2$	$\delta_2$	$\varepsilon_2$
$D_a^{l_2}$	$0_{\alpha_2\alpha_2}$	*	$-I_{\gamma_2\gamma_2}$	$-I_{\delta_2\delta_2}$	$-I_{\varepsilon_2\varepsilon_2}$
$D_b^{l_2}$	$-I_{\alpha_2\alpha_2}$	*	$0_{\gamma_2\gamma_2}$	$0_{\delta_2\delta_2}$	$0_{\varepsilon_2\varepsilon_2}$
$D_a^{r_2}$	$-I_{\alpha_2\alpha_2}$	$-I_{\beta_2\beta_2}$	$-I_{\gamma_2\gamma_2}$	$0_{\delta_2\delta_2}$	*
$D_b^{r_2}$	$0_{\alpha_2\alpha_2}$	$0_{\beta_2\beta_2}$	$0_{\gamma_2\gamma_2}$	$-I_{\delta_2\delta_2}$	*

Here, \* stands for a negative semidefinite matrix whose precise entries do not matter.

In order to understand the meaning of the two tables from Lemma 7.2, let us take a closer look at the diagonal matrix  $D_a^{l_1}$ , for example: Then the first line of the first table means that the subblocks of this diagonal matrix are given by

$$[D_a^{l_1}]_{\alpha_1\alpha_1} = 0_{\alpha_1\alpha_1}, \quad [D_a^{l_1}]_{\gamma_1\gamma_1} = -I_{\gamma_1\gamma_1}, \quad [D_a^{l_1}]_{\delta_1\delta_1} = -I_{\delta_1\delta_1}, \quad [D_a^{l_1}]_{\varepsilon_1\varepsilon_1} = -I_{\varepsilon_1\varepsilon_1},$$

whereas the block entry  $[D_a^{l_1}]_{\beta_1\beta_1}$  has no special structure (it is negative semidefinite, but nothing else can be said). The other lines in the two tables have to be interpreted in a corresponding way. This particular structure of the diagonal matrices will play a central role in our subsequent analysis.

We now consider the homogeneous linear system of equations  $V \cdot \Delta z = 0$  and partition the vector  $\Delta z$  in exactly the same way as in (34). Following the proof of Proposition 6.1, we arrive at (48)–(56). From that point on, we have to change the proof since condition (33) does not hold any longer. To this end, we first observe that (51) and (52) immediately give

$$\Delta p_1 = \Delta p_2 =: \Delta p \quad \text{and, therefore,} \quad \Delta y = -A\Delta p. \quad (67)$$

Substituting this into (50) yields

$$-A^2\Delta p - B_1\Delta u_1 - B_2\Delta u_2 = 0.$$

Taking into account Assumption 7.1, system (48)–(56) thus reduces to

$$\nu_1\Delta u_1 - B_1^T\Delta p - \Delta\lambda l_1 + \Delta\lambda r_1 = 0, \quad (68)$$

$$\nu_2\Delta u_2 - B_2^T\Delta p - \Delta\lambda l_2 + \Delta\lambda r_2 = 0, \quad (69)$$

$$A^2\Delta p + B_1\Delta u_1 + B_2\Delta u_2 = 0, \quad (70)$$

$$D_a^{l_1}\Delta\lambda l_1 + D_b^{l_1}\Delta u_1 = 0, \quad (71)$$

$$D_a^{r_1}\Delta\lambda r_1 - D_b^{r_1}\Delta u_1 = 0, \quad (72)$$

$$D_a^{l_2}\Delta\lambda l_2 + D_b^{l_2}\Delta u_2 = 0, \quad (73)$$

$$D_a^{r_2}\Delta\lambda r_2 - D_b^{r_2}\Delta u_2 = 0. \quad (74)$$

Solving (68), (69) for  $\Delta u_1$  and  $\Delta u_2$ , respectively, we obtain

$$\Delta u_1 = \frac{1}{\nu_1} (B_1^T\Delta p + \Delta\lambda l_1 - \Delta\lambda r_1), \quad (75)$$

$$\Delta u_2 = \frac{1}{\nu_2} (B_2^T\Delta p + \Delta\lambda l_2 - \Delta\lambda r_2). \quad (76)$$

Inserting these expressions into (70) and rearranging terms gives

$$M\Delta p + \frac{1}{\nu_1}B_1(\Delta\lambda l_1 - \Delta\lambda r_1) + \frac{1}{\nu_2}B_2(\Delta\lambda l_2 - \Delta\lambda r_2) = 0, \quad (77)$$

where, for simplicity of notation, we put

$$M := A^2 + \frac{1}{\nu_1}B_1B_1^T + \frac{1}{\nu_2}B_2B_2^T. \quad (78)$$

We next replace  $\Delta u_1, \Delta u_2$  from (75), (76) also in equations (71)–(74) to obtain

$$\begin{aligned} D_a^{l_1} \Delta \lambda_1 + \frac{1}{\nu_1} D_b^{l_1} (B_1^T \Delta p + \Delta \lambda_1 - \Delta \lambda r_1) &= 0, \\ D_a^{r_1} \Delta \lambda r_1 - \frac{1}{\nu_1} D_b^{r_1} (B_1^T \Delta p + \Delta \lambda_1 - \Delta \lambda r_1) &= 0, \\ D_a^{l_2} \Delta \lambda_2 + \frac{1}{\nu_2} D_b^{l_2} (B_2^T \Delta p + \Delta \lambda_2 - \Delta \lambda r_2) &= 0, \\ D_a^{r_2} \Delta \lambda r_2 - \frac{1}{\nu_2} D_b^{r_2} (B_2^T \Delta p + \Delta \lambda_2 - \Delta \lambda r_2) &= 0. \end{aligned}$$

Reordering terms, system (68)–(74) therefore reduces to

$$\begin{aligned} M \Delta p + \frac{1}{\nu_1} B_1 (\Delta \lambda_1 - \Delta \lambda r_1) + \frac{1}{\nu_2} B_2 (\Delta \lambda_2 - \Delta \lambda r_2) &= 0, \\ \frac{1}{\nu_1} D_b^{l_1} B_1^T \Delta p + (D_a^{l_1} + \frac{1}{\nu_1} D_b^{l_1}) \Delta \lambda_1 - \frac{1}{\nu_1} D_b^{l_1} \Delta \lambda r_1 &= 0, \\ -\frac{1}{\nu_1} D_b^{r_1} B_1^T \Delta p + (D_a^{r_1} + \frac{1}{\nu_1} D_b^{r_1}) \Delta \lambda r_1 - \frac{1}{\nu_1} D_b^{r_1} \Delta \lambda_1 &= 0, \\ \frac{1}{\nu_2} D_b^{l_2} B_2^T \Delta p + (D_a^{l_2} + \frac{1}{\nu_2} D_b^{l_2}) \Delta \lambda_2 - \frac{1}{\nu_2} D_b^{l_2} \Delta \lambda r_2 &= 0, \\ -\frac{1}{\nu_2} D_b^{r_2} B_2^T \Delta p + (D_a^{r_2} + \frac{1}{\nu_2} D_b^{r_2}) \Delta \lambda r_2 - \frac{1}{\nu_2} D_b^{r_2} \Delta \lambda_2 &= 0. \end{aligned}$$

Let us define

$$\Delta q_1 := B_1^T \Delta p, \quad \Delta q_2 := B_2^T \Delta p \quad (79)$$

as well as

$$D^{l_1} := D_a^{l_1} + \frac{1}{\nu_1} D_b^{l_1}, \quad D^{r_1} := D_a^{r_1} + \frac{1}{\nu_1} D_b^{r_1}, \quad D^{l_2} := D_a^{l_2} + \frac{1}{\nu_2} D_b^{l_2}, \quad D^{r_2} := D_a^{r_2} + \frac{1}{\nu_2} D_b^{r_2}.$$

Noting that these diagonal matrices are nonsingular (in fact, negative definite) and denoting by

$$D^{-l_1}, D^{-r_1}, D^{-l_2}, D^{-r_2} \quad \text{the inverses of } D^{l_1}, D^{r_1}, D^{l_2}, D^{r_2},$$

respectively, our linear system can be rewritten as

$$0 = M \Delta p + \frac{1}{\nu_1} B_1 (\Delta \lambda_1 - \Delta \lambda r_1) + \frac{1}{\nu_2} B_2 (\Delta \lambda_2 - \Delta \lambda r_2), \quad (80)$$

$$\Delta \lambda_1 = \frac{1}{\nu_1} D^{-l_1} D_b^{l_1} (\Delta \lambda r_1 - \Delta q_1), \quad (81)$$

$$\Delta \lambda r_1 = \frac{1}{\nu_1} D^{-r_1} D_b^{r_1} (\Delta \lambda_1 + \Delta q_1), \quad (82)$$

$$\Delta \lambda_2 = \frac{1}{\nu_2} D^{-l_2} D_b^{l_2} (\Delta \lambda r_2 - \Delta q_2), \quad (83)$$

$$\Delta\lambda r_2 = \frac{1}{\nu_2} D^{-r_2} D_b^{r_2} (\Delta\lambda l_2 + \Delta q_2). \quad (84)$$

We now exploit the fact that the vector  $\bar{z}$  under consideration is a solution of  $H(z) = 0$ . Hence, the index sets  $\alpha_i, \beta_i, \gamma_i, \delta_i, \varepsilon_i$  form a partition of the set  $\{1, \dots, n_i\}$  for  $i = 1, 2$ . Therefore, considering equations (81), (82) for each of the blocks  $\alpha_1, \beta_1, \gamma_1, \delta_1, \varepsilon_1$  as well as equations (83), (84) for each of the block components  $\alpha_2, \beta_2, \gamma_2, \delta_2, \varepsilon_2$  separately and exploiting the special structure of the diagonal matrices from Lemma 7.2, we obtain

$$\begin{aligned} [\Delta\lambda l_1]_{\alpha_1} &= -\frac{1}{\nu_1} [D^{-l_1}]_{\alpha_1\alpha_1} ([\Delta\lambda r_1]_{\alpha_1} - [\Delta q_1]_{\alpha_1}), \\ [\Delta\lambda l_1]_{\beta_1} &= \frac{1}{\nu_1} [D^{-l_1}]_{\beta_1\beta_1} [D_b^{l_1}]_{\beta_1\beta_1} ([\Delta\lambda r_1]_{\beta_1} - [\Delta q_1]_{\beta_1}), \\ [\Delta\lambda l_1]_{\gamma_1} &= [\Delta\lambda l_1]_{\delta_1} = [\Delta\lambda l_1]_{\varepsilon_1} = 0, \end{aligned}$$


---

$$\begin{aligned} [\Delta\lambda r_1]_{\alpha_1} &= [\Delta\lambda r_1]_{\beta_1} = [\Delta\lambda r_1]_{\gamma_1} = 0, \\ [\Delta\lambda r_1]_{\delta_1} &= -\frac{1}{\nu_1} [D^{-r_1}]_{\delta_1\delta_1} ([\Delta\lambda l_1]_{\delta_1} + [\Delta q_1]_{\delta_1}), \\ [\Delta\lambda r_1]_{\varepsilon_1} &= \frac{1}{\nu_1} [D^{-r_1}]_{\varepsilon_1\varepsilon_1} [D_b^{r_1}]_{\varepsilon_1\varepsilon_1} ([\Delta\lambda l_1]_{\varepsilon_1} + [\Delta q_1]_{\varepsilon_1}), \end{aligned}$$


---

$$\begin{aligned} [\Delta\lambda l_2]_{\alpha_2} &= -\frac{1}{\nu_2} [D^{-l_2}]_{\alpha_2\alpha_2} ([\Delta\lambda r_2]_{\alpha_2} - [\Delta q_2]_{\alpha_2}), \\ [\Delta\lambda l_2]_{\beta_2} &= \frac{1}{\nu_2} [D^{-l_2}]_{\beta_2\beta_2} [D_b^{l_2}]_{\beta_2\beta_2} ([\Delta\lambda r_2]_{\beta_2} - [\Delta q_2]_{\beta_2}), \\ [\Delta\lambda l_2]_{\gamma_2} &= [\Delta\lambda l_2]_{\delta_2} = [\Delta\lambda l_2]_{\varepsilon_2} = 0, \end{aligned}$$


---

$$\begin{aligned} [\Delta\lambda r_2]_{\alpha_2} &= [\Delta\lambda r_2]_{\beta_2} = [\Delta\lambda r_2]_{\gamma_2} = 0, \\ [\Delta\lambda r_2]_{\delta_2} &= -\frac{1}{\nu_2} [D^{-r_2}]_{\delta_2\delta_2} ([\Delta\lambda l_2]_{\delta_2} + [\Delta q_2]_{\delta_2}), \\ [\Delta\lambda r_2]_{\varepsilon_2} &= \frac{1}{\nu_2} [D^{-r_2}]_{\varepsilon_2\varepsilon_2} [D_b^{r_2}]_{\varepsilon_2\varepsilon_2} ([\Delta\lambda l_2]_{\varepsilon_2} + [\Delta q_2]_{\varepsilon_2}). \end{aligned}$$

Taking into account that several block components of  $\Delta\lambda l_1, \Delta\lambda r_1, \Delta\lambda l_2, \Delta\lambda r_2$  are equal to zero, we may further reduce the nontrivial terms to

$$\begin{aligned} [\Delta\lambda l_1]_{\alpha_1} &= \frac{1}{\nu_1} [D^{-l_1}]_{\alpha_1\alpha_1} [\Delta q_1]_{\alpha_1}, \\ [\Delta\lambda l_1]_{\beta_1} &= -\frac{1}{\nu_1} [D^{-l_1}]_{\beta_1\beta_1} [D_b^{l_1}]_{\beta_1\beta_1} [\Delta q_1]_{\beta_1}, \end{aligned}$$


---

$$\begin{aligned}
[\Delta\lambda r_1]_{\delta_1} &= -\frac{1}{\nu_1} [D^{-r_1}]_{\delta_1\delta_1} [\Delta q_1]_{\delta_1}, \\
[\Delta\lambda r_1]_{\varepsilon_1} &= \frac{1}{\nu_1} [D^{-r_1}]_{\varepsilon_1\varepsilon_1} [D_b^{r_1}]_{\varepsilon_1\varepsilon_1} [\Delta q_1]_{\varepsilon_1}, \\
\hline
[\Delta\lambda l_2]_{\alpha_2} &= \frac{1}{\nu_2} [D^{-l_2}]_{\alpha_2\alpha_2} [\Delta q_2]_{\alpha_2}, \\
[\Delta\lambda l_2]_{\beta_2} &= -\frac{1}{\nu_2} [D^{-l_2}]_{\beta_2\beta_2} [D_b^{l_2}]_{\beta_2\beta_2} [\Delta q_2]_{\beta_2}, \\
\hline
[\Delta\lambda r_2]_{\delta_2} &= -\frac{1}{\nu_2} [D^{-r_2}]_{\delta_2\delta_2} [\Delta q_2]_{\delta_2}, \\
[\Delta\lambda r_2]_{\varepsilon_2} &= \frac{1}{\nu_2} [D^{-r_2}]_{\varepsilon_2\varepsilon_2} [D_b^{r_2}]_{\varepsilon_2\varepsilon_2} [\Delta q_2]_{\varepsilon_2}.
\end{aligned}$$

Hence we obtain

$$\Delta\lambda l_1 - \Delta\lambda r_1 = \Lambda_1\Delta q_1 \quad \text{and} \quad \Delta\lambda l_2 - \Delta\lambda r_2 = \Lambda_2\Delta q_2 \quad (85)$$

with diagonal matrices  $\Lambda_1$  and  $\Lambda_2$  given by

$$\begin{aligned}
\Lambda_1 &:= \frac{1}{\nu_1} \text{diag} \left( [D^{-l_1}]_{\alpha_1\alpha_1}, -[D^{-l_1}]_{\beta_1\beta_1} [D_b^{l_1}]_{\beta_1\beta_1}, 0_{\gamma_1\gamma_1}, [D^{-r_1}]_{\delta_1\delta_1}, -[D^{-r_1}]_{\varepsilon_1\varepsilon_1} [D_b^{r_1}]_{\varepsilon_1\varepsilon_1} \right), \\
\Lambda_2 &:= \frac{1}{\nu_2} \text{diag} \left( [D^{-l_2}]_{\alpha_2\alpha_2}, -[D^{-l_2}]_{\beta_2\beta_2} [D_b^{l_2}]_{\beta_2\beta_2}, 0_{\gamma_2\gamma_2}, [D^{-r_2}]_{\delta_2\delta_2}, -[D^{-r_2}]_{\varepsilon_2\varepsilon_2} [D_b^{r_2}]_{\varepsilon_2\varepsilon_2} \right).
\end{aligned}$$

The central property of these two diagonal matrices are summarized in the following result.

**Lemma 7.3** *The matrices  $I + \Lambda_1$  and  $I + \Lambda_2$  are both positive semidefinite.*

**Proof.** We verify the statement only for the matrix  $I + \Lambda_1$  since the proof is similar for the second matrix  $I + \Lambda_2$ . To this end, we take a closer look at the block components of the matrix  $\Lambda_1$  corresponding to the blocks defined by the index sets  $\alpha_1, \beta_1, \gamma_1, \delta_1$ , and  $\varepsilon_1$ , respectively. We have to show that all entries of the diagonal matrix  $\Lambda_1$  are greater or equal to  $-1$ .

First, consider the elements from the index set  $\alpha_1$ . The definition of the matrix  $D^{l_1}$  together with Lemma 7.2 then shows that  $\frac{1}{\nu_1} [D^{-l_1}]_{\alpha_1\alpha_1} = -I_{\alpha_1\alpha_1}$ . Next, for the index set  $\beta_1$ , we can argue as for the index set  $\varepsilon_1$ , see below. The diagonal elements of  $\Lambda_1$  belonging to the index set  $\gamma_1$  are equal to zero by definition. Using Lemma 7.2 once again, we immediately see that the  $\delta_1$ -block is given by  $\frac{1}{\nu_1} [D^{-r_1}]_{\delta_1\delta_1} = -I_{\delta_1\delta_1}$ .

Finally, let us consider the block corresponding to the index set  $\varepsilon_1$ , and take an arbitrary element  $i \in \varepsilon_1$ . The definition of the matrix  $\Lambda_1$  shows that the diagonal element  $d_i$  belonging to this index is given by

$$d_i = \frac{1}{\nu_1} (\zeta_i - 1) \frac{1}{(1 - \xi_i) + \frac{1}{\nu_1}(1 - \zeta_i)} = \frac{\zeta_i - 1}{\nu_1(1 - \xi_i) + (1 - \zeta_i)}$$

for some vector  $(\xi_i, \zeta_i)$  satisfying  $\|(\xi_i, \zeta_i)\| \leq 1$ . An elementary calculation then shows that  $d_i \geq -1$  holds. Altogether, this completes the proof.  $\square$

We are now in a position to verify the nonsingularity of the given matrix  $V \in \partial H(\bar{z})$ . To this end, let us come back to equation (77). Exploiting (85), we obtain

$$M\Delta p + \frac{1}{\nu_1}B_1\Lambda_1\Delta q_1 + \frac{1}{\nu_2}B_2\Lambda_2\Delta q_2 = 0.$$

In view of the definition (79) of the two vectors  $\Delta q_1$  and  $\Delta q_2$ , this may be rewritten as

$$\left(M + \frac{1}{\nu_1}B_1\Lambda_1B_1^T + \frac{1}{\nu_2}B_2\Lambda_2B_2^T\right)\Delta p = 0.$$

Recalling the definition (78) of  $M$ , this is equivalent to

$$\left(A^2 + \frac{1}{\nu_1}B_1(I + \Lambda_1)B_1^T + \frac{1}{\nu_2}B_2(I + \Lambda_2)B_2^T\right)\Delta p = 0.$$

Since  $A$  and, therefore, also  $A^2$  is positive definite, it now follows from Lemma 7.3 that  $\Delta p = 0$ . By definition, this means that  $\Delta p_1 = 0$  and  $\Delta p_2 = 0$ , which in turn also gives  $\Delta y = 0$ , cf. (67). Furthermore, it also follows from (79) that  $\Delta q_1 = 0$  and  $\Delta q_2 = 0$ . This immediately gives  $\Delta\lambda_1 = 0, \Delta\lambda r_1 = 0, \Delta\lambda_2 = 0$ , and  $\Delta\lambda r_2 = 0$ . Hence (75), (76) yield  $\Delta u_1 = 0$  and  $\Delta u_2 = 0$ . Finally, this also gives  $\Delta w l_1 = 0, \Delta w r_1 = 0, \Delta w l_2 = 0$ , and  $\Delta w r_2 = 0$ . We have thus proven the following result.

**Theorem 7.4** *Let  $\bar{z}$  be a solution of  $H(z) = 0$  such that Assumption 7.1 holds. Then all elements  $V \in \partial H(\bar{z})$  are nonsingular.*

Recall that Theorem 7.4 is highly important since it guarantees the local quadratic convergence of the semismooth Newton method applied to  $H(z) = 0$ . Hence, together with Theorem 6.2, it follows that Algorithm 5.3 has very nice global and local convergence properties.

## 8 Numerical Results

We validate our globalized semismooth Newton scheme (Algorithm 5.3 in MATLAB) using the stepsize parameters  $\sigma = 10^{-4}$  and  $\beta = 0.5$  as well as the termination parameter  $\varepsilon = 10^{-8}$ . Since we never observed singularity problems, we always took the Newton-type direction and remove the switching to the antigradient direction from Algorithm 5.3 in our implementation. We use  $\Omega = (0, 1) \times (0, 1)$ , the matrices  $B_1$  and  $B_2$  from (66), and the starting point  $z^0 = 0$ . We tested several examples with different choices of the stepsize  $h = 1/(N + 1)$ , the lower and upper bounds  $l_1, r_1, l_2, r_2$ , the weights  $\nu_1, \nu_2$ , the target values  $z_1, z_2$ , and the right-hand side  $f$ . Our method was able to solve all cases

with typical Newton efficiency and showing robustness with respect to the optimization parameters.

For the purpose of validating the accuracy of the computed optimal solution, we define the following test case with known analytical solution. We have

$$\begin{aligned} y(x) &:= \sin(\pi x_1) \sin(\pi x_2), \\ p_1(x) &:= \sin(2\pi x_1) \sin(2\pi x_2), \\ p_2(x) &:= \sin(3\pi x_1) \sin(3\pi x_2). \end{aligned}$$

Then set

$$z_1(x) := y(x) - \Delta p_1(x) \quad \text{and} \quad z_2(x) := y(x) - \Delta p_2(x).$$

Now, using

$$u_1(x) := \max \left\{ l_1, \min \left\{ r_1, \frac{1}{\nu_1} B_1^* p_1 \right\} \right\}, \quad (86)$$

$$u_2(x) := \max \left\{ l_2, \min \left\{ r_2, \frac{1}{\nu_2} B_2^* p_2 \right\} \right\} \quad (87)$$

and, finally,

$$f(x) := -\Delta y(x) - B_1 u_1(x) - B_2 u_2(x),$$

it is not difficult to see that  $u_1, u_2$  satisfy the optimality conditions from (9), hence  $u_1, u_2$  are the optimal controls and  $y$  the optimal state, i.e., the example is constructed in such a way that we know, a priori, the solution of the underlying Nash equilibrium problem.

We illustrate different properties of our semismooth Newton method by looking at this test case from different perspectives. For all test runs, we take the lower bounds  $l_j$  equal to  $-0.5$  and the upper bounds  $r_j$  equal to  $+0.5$ . As can be seen from the corresponding figures, this choice guarantees that the constraints on the controls are active in certain regions. Furthermore, we use  $\nu_1 = \nu_2 = 1$  for our first set of test runs where we want to investigate the behaviour of the semismooth Newton method when the dimension increases. Table 1 summarizes the results that we obtained using different step sizes  $h = 1/(N + 1)$ ,  $N \in \mathbb{N}$ . More precisely, we report the function values  $\Psi(z^k)$  at each iteration for different discretizations. The resulting optimal controls  $u_1$  and  $u_2$  are shown in Figure 1 for the case  $N = 64$ . As can be seen from Table 1, the number of iterations remains essentially constant and therefore mesh-size independent.

Next, we take the same data and investigate the behaviour of the errors

$$|u_1^* - u_1^f|_0, \quad |u_2^* - u_2^f|_0, \quad |y^* - y^f|_0,$$

where  $u_j^*, y^*$  denote the (known) continuous optimal solution, evaluated at the discrete points, and  $u_j^f, y^f$  are the approximate solutions at the final iterate. Table 2 presents these errors for different discretizations. The results show that, doubling the dimension  $N$ ,

$k$	$N = 16$	$N = 32$	$N = 64$	$N = 128$
0	1379566.3890632149	5198447.9808923909	20168480.847509615	79437620.7950397581
1	144.6654260953	573.1040391275	2283.3415965325	9117.1902512238
2	6.1186669271	22.4315506473	87.0916543274	344.1033825083
3	1.1705448216	3.0182299762	15.3555186035	62.5034453387
4	0.0494094196	0.1134067727	0.6388759404	2.6232961841
5	0.0010851899	0.0019904354	0.0137002072	0.0571646369
6	0.0000026786	0.0000025625	0.0000207808	0.0000826226
7	0.0000000018	0.0000000003	0.0000000060	0.0000000199
8				0.0000000000

Table 1: Function values  $\Psi(z^k)$  for different discretizations

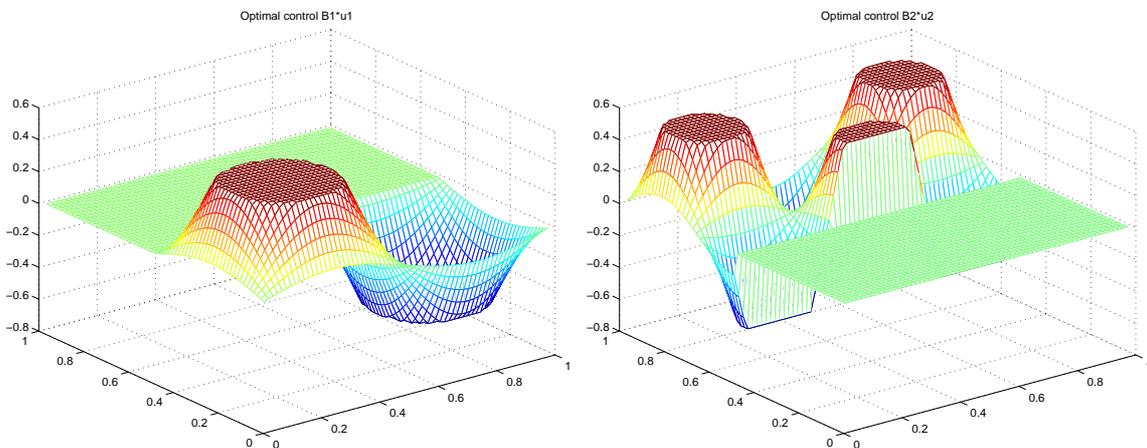


Figure 1: Computed optimal controls  $B_1u_1, B_2u_2$  for  $N = 64$

i.e. halving the stepsize  $h$ , results in an accuracy that is about four times better. This is consistent with the second-order error estimate given in Theorem 4.2, although the controls are active and a lower degree of convergence of the controls could be expected.

Finally, we take a look at the behaviour of our method for decreasing values of the regularization parameters  $\nu_1, \nu_2$ . To this end, we take the same test problem data as before, using the fixed discretization  $N = 64$ , and let  $\nu_1 = \nu_2 =: \nu$  go down from  $\nu = 1$  to  $\nu = 10^{-8}$ . The number of iterations needed by our semismooth Newton method are reported in Table 3. The number of iterations increases slightly, but eventually stays constant, using nine iterations for all sufficiently small values of  $\nu_j$ .

In fact, the method is able to solve the resulting problem also in the limiting case  $\nu_1 = \nu_2 = 0$  in nine iterations, though this case is not covered by our theory. More precisely, in this case we cannot use (86), (87) to construct the functions  $u_1, u_2$ . Therefore, we take small positive values  $\nu_1 = \nu_2 := 10^{-16}$  to define these functions, just to get a suitable test

error	$N = 16$	$N = 32$	$N = 64$	$N = 128$
$ u_1^* - u_1^f _0$	0.0015673656	0.0004297785	0.0001122424	0.0000285216
$ u_2^* - u_2^f _0$	0.0035529235	0.0010132386	0.0002515317	0.0000649400
$ y^* - y^f _0$	0.0014315618	0.0003772036	0.0000971426	0.0000246742

Table 2: Behaviour of the errors  $|u_1^* - u_1^f|_0$ ,  $|u_2^* - u_2^f|_0$  and  $|y^* - y^f|_0$

$\nu$	1	$10^{-1}$	$10^{-2}$	$10^{-3}$	$10^{-4}$	$10^{-5}$	$10^{-6}$	$10^{-7}$	$10^{-8}$
iterations	7	8	9	10	9	9	9	9	9

Table 3: Number of iterations for different values of  $\nu := \nu_1 = \nu_2$

problem, and afterwards we do all the calculations with the semismooth Newton method using  $\nu_1 = \nu_2 = 0$ , i.e. without a regularization term. This means that the functions  $u_1, u_2$  generated via (86), (87) are no longer the analytical solutions of this example. However, the algorithm still works very well, and the (bang-bang) solution is depicted in Figure 2.

## 9 Conclusion

The formulation and the semismooth Newton solution of Nash equilibria multiobjective elliptic optimal control problems was presented. The convergence of the semismooth Newton method resulted to be quadratic as predicted by the theoretical investigation. Second-order accuracy of the finite difference approximation to the unique Nash equilibrium solution was demonstrated. Future research will focus on the analysis of the case of more than two objectives and of different control mechanisms.

## References

- [1] C. BAIOCCHI AND A. CAPELO, *Variational and Quasivariational Inequalities. Applications to Free-Boundary Problems*, John Wiley & Sons, Chichester, 1984.
- [2] A. BORZÌ AND K. KUNISCH, *A multigrid scheme for elliptic constrained optimal control problems*, Computational Optimization and Applications, 31 (2005), pp. 309–333.
- [3] A. BORZÌ, K. KUNISCH, AND D.Y. KWAK, *Accuracy and convergence properties of the finite difference multigrid solution of an optimal control optimality system*, SIAM J. Control Opt., 41 (2003), pp. 1477–1497.

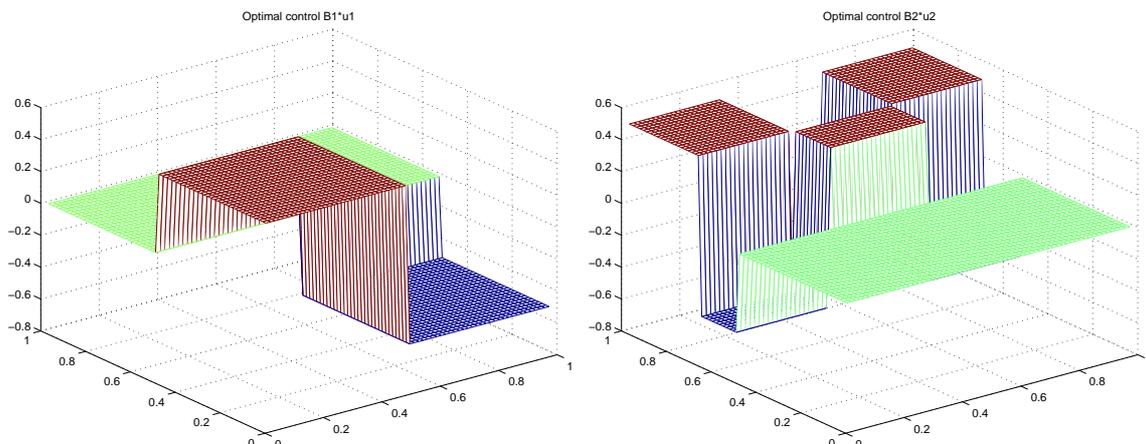


Figure 2: Computed optimal controls  $B_1u_1, B_2u_2$  for  $N = 64$  with  $\nu_1 = \nu_2 = 0$

- [4] A. BORZI AND V. SCHULZ, *Computational optimization of systems governed by partial differential equations*, SIAM book series on Computational Science and Engineering 08, SIAM Philadelphia 2012.
- [5] A. CHINCHULUUN, P.M. PARDALOS, A. MIGDALAS, AND L. PITSOULIS (EDS.), *Pareto Optimality, Game Theory, and Equilibria*, Springer, 2008.
- [6] F.H. CLARKE, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983 (reprinted by SIAM, Philadelphia, PA, 1990).
- [7] T. DE LUCA, F. FACCHINEI, AND C. KANZOW, *A semismooth equation approach to the solution of nonlinear complementarity problems*, Math. Program., 75 (1996), pp. 407–439.
- [8] S. DEMPE, *Foundations of Bilevel Programming*, Kluwer Academic Press, Dordrecht, 2002.
- [9] J.-A. DÉSIDÉRI, *Cooperation and competition in multidisciplinary optimization, Application to the aero-structural aircraft wing shape optimization*. Computational Optimization and Applications; DOI: 10.1007/s10589-011-9395-1.
- [10] M. EHRGOTT, *Multicriteria Optimization*, Springer, 2nd edition 2005.
- [11] F. FACCHINEI AND C. KANZOW, *Generalized Nash equilibrium problems*, Annals Oper. Res., 175 (2010), pp. 177–211.
- [12] F. FACCHINEI AND J. SOARES, *A new merit function for nonlinear complementarity problems and a related algorithm*, SIAM J. Optim., 7 (1997), pp. 225–247.

- [13] A. FISCHER, *A special Newton-type optimization method*, Optim., 24 (1992), pp. 269–284.
- [14] F. FORGÓ, J. SZÉP, AND F. SZIDAROVSKY, *Introduction to the Theory of Games. Concepts, Methods, Applications*, Kluwer Academic Press, Dordrecht, The Netherlands, 1999.
- [15] N. GAUGER, C. ILIC, S. SCHMIDT AND V. SCHULZ, *Non-parametric aerodynamic shape optimization*, in G. Leugering, S. Engell, A. Griewank, M. Hinze, R. Rannacher, V. Schulz, M. Ulbrich and S. Ulbrich (eds.), *Constrained Optimization and Optimal Control for Partial Differential Equations*, International Series of Numerical Mathematics, vol. 160, pp. 289-300, Birkhäuser 2012.
- [16] W. HACKBUSCH, *Elliptic Differential Equations*, Springer-Verlag, New York, 1992.
- [17] M. HINTERMÜLLER UND M. HINZE, *A SQP-semi-smooth Newton-type algorithm applied to control of the instationary Navier-Stokes system subject to control constraints*, SIAM J. Optim., 16 (2006), pp. 1177–1200.
- [18] M. HINZE AND F. TRÖLTZSCH, *Discrete concepts versus error analysis in pde constrained optimization*, GAMM-Mitt. 33 (2010), pp. 148-162.
- [19] K. ITO AND K. KUNISCH, *Lagrange Multiplier Approach to Variational Problems and Applications*, SIAM, Philadelphia, 2008.
- [20] J.L. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin, 1971.
- [21] D. KINDERLEHRER AND G. STAMPACCHIA, *An Introduction to Variational Inequalities and their Applications*, Academic Press, New York, 1980 (reprinted by SIAM, Philadelphia, 2000).
- [22] J.L. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin, 1971.
- [23] J.L. LIONS, *Contrôle de Pareto de systemes distribues: Le cas d'evolution*. Comptes Rendus de l'Academie des Sciences, Serie I, 302 (1986), pp. 413–417.
- [24] G.P. LIU, J.B. YANG, AND J.F. WHIDBORNE, *Multiobjective Optimisation and Control*, Research Studies Press LTD, 2001.
- [25] K. MALANOWSKI, *Convergence of approximations vs. regularity of solutions for convex, control-constrained optimal-control problems*, Appl. Math. Optim., 8 (1981), pp. 69–95.
- [26] C. MEYER AND A. RÖSCH, *Superconvergence properties of optimal control problems*, SIAM J. Control Opt., 43 (2004), pp. 970–985.

- [27] L. QI, *Convergence analysis of some algorithms for solving nonsmooth equations*, Math. Oper. Res., 18 (1993), pp. 227–244.
- [28] L. QI AND J. SUN, *A nonsmooth version of Newton’s method*, Math. Program., 58 (1993), pp. 353–368.
- [29] A.M. RAMOS, R. GLOWINSKI, AND J. PERIAUX, *Nash equilibria for the multiobjective control of linear partial differential equations*, Journal of Optimization Theory and Applications, 112 (2002), pp. 457–498.
- [30] A.M. RAMOS, R. GLOWINSKI, AND J. PERIAUX, *Pointwise control of the Burgers equation and related Nash equilibrium problems: Computational approach*. Journal of Optimization Theory and Applications, 112 (2002), pp. 499–516.
- [31] A. RÖSCH, *Error estimates for linear-quadratic control problems with control constraints*, Optimization Methods and Software, 21 (2006), pp. 121–134.
- [32] W. STADLER, *Fundamentals of multicriteria optimization*, In Multicriteria optimization in engineering and in the sciences, volume 37 of Math. Concepts Methods Sci. Engrg., pages 1–25. Plenum, New York, 1988.
- [33] E. SÜLI, *Convergence of finite volume schemes for Poisson’s equation on nonuniform meshes*, SIAM J. Numer. Anal., 28 (1991), pp. 1419–1430.
- [34] F. TRÖLTZSCH, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, AMS, 2010.
- [35] M. ULBRICH, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization in Function Spaces*. MOS-SIAM Series in Optimization 11, Philadelphia, 2011.