**Control and Cybernetics**

vol. **45** (2016) No. 4

# A Nash equilibrium approach for multiobjective optimal control problems with elliptic partial differential equations*

by

**Axel Dreves**

Universität der Bundeswehr München, Department of Aerospace Engineering, Werner-Heisenberg Weg 39, 85577 Neubiberg, Germany, Axel.Dreves@unibw.de

**Abstract:** We consider the generalized Nash equilibrium as a solution concept for multiobjective optimal control problems governed by elliptic partial differential equations with constraints not only for the control but also for the state variables. In the first part, we present a constructive proof of the existence of a generalized Nash equilibrium via an approximating sequence of suitable finite dimensional discretizations. In the second part, we propose a variant of a potential reduction algorithm for the numerical solution of these discretized problems. In contrast to the existing numerical approaches ours does not require the computation of the control–to–state mapping. Instead we introduce different state variables and guarantee that they become equal at a solution. We prove sufficient conditions for the convergence of our algorithm to a solution. Furthermore, some numerical results showing the applicability are provided.

**Keywords:** multiobjective optimization, generalized Nash equilibrium, optimal control problem, elliptic partial differential equation, finite elements, interior point method

## 1. Introduction

Many physical, economic or financial problems are modeled using partial differential equations (PDEs) and this also is often the setup for different objectives to be commonly optimized. This leads to multiobjective optimal control problems. We want to consider those problems where the different objectives are pursued independently, and hence we introduce for each objective one player who chooses a strategy to optimize his objective function without cooperating with any of the other players. To model this, we use the Nash equilibrium concept for a solution of the multiobjective optimal control problem. However, since the optimization problems of the players are coupled by the common PDE and

*Submitted: November 2014; Accepted: December 2016

we may have further common constraints, we get a dependence on the variables of other players not only in the objective function but also in the feasible set of a player, and hence we do not have standard but generalized Nash equilibrium problems (GNEPs), which were first introduced in a finite dimensional setting in Arrow and Debreu (1954).

While for the multiobjective optimal control problems the concept of Pareto-optimal solutions is used in a number of publications, see, e.g., Lions (1986), Liu, Yang and Whidborne (2001), there is currently not much literature on multiobjective optimal control problems and generalized Nash equilibria. A specific unconstrained setting and a conjugate gradient method for its solution is considered in Ramos, Glowinski and Periaux (2002a,b). Building up on this, Borzi and Kanzow (2013) develop a semismooth Newton method for the solution of a class of multiobjective optimal control problems governed by a linear elliptic PDE. Moreover, an existence result for the solution of such problems, which are standard Nash equilibrium problems, is provided. In Hintermüller and Surowiec (2013), more general problems are considered, since constraints also for the state variables and not only for the control variables are allowed, which leads to an additional coupling in the constraints and hence to GNEPs. In their approach, they suggest to solve a sequence of pure Nash equilibrium problems by adding the state constraints via penalty terms to the objective functions. Thus, they are able to prove an existence result and to derive Karush-Kuhn-Tucker (KKT)-like optimality conditions. Furthermore, the recent papers, Dreves and Gwinner (2016) and Hintermüller, Surowiec and Kämmler (2015), use a Nikaido-Isoda function approach to discuss variational Nash equilibria in the context of GNEPs for multiobjective optimal control problems. The first one develops a relaxation method, and the second one uses a path-following method, which can even deal with parabolic PDEs.

Like the here mentioned articles we will consider linear PDEs only, since we will require convexity for our feasible set. There also exist, however in a different context of special 2-player games, some analyses for problems involving nonlinear PDEs, see Ramos and Roubiček (2007).

All the existing approaches use the so called reduced problem, by defining the solution function of the PDE as a function of the control variables, and inserting this into the objective function. Then, the reduced problem is a problem in the control variables only. This leads to the necessity of computing the derivatives of the reduced objective functions and to the introduction of adjoint variables. In our numerical approach we will not use the adjoint approach, but in a new approach we introduce different state variables for each player and we will guarantee that these states become equal at the solution. This can be used to extend an algorithm for finite dimensional GNEPs to the discretized optimal control problem. Our general strategy here is "first discretize then optimize". Finally, we would like to refer to Hinze et al. (2009) and Tröltzsch (2010) for some standard theory on optimal control problems with a single objective function and PDE constraints.

This paper is organized as follows. In Section 2 we formulate the considered problem and describe a finite element discretization of it. Section 3 examines the relation between the original and the discretized problem and provides convergence of the discretized solutions to a solution of the continuous problem. In Section 4 we suggest a potential reduction algorithm for the solution of the discretized problem, and in Section 5 we discuss sufficient conditions for the convergence of the algorithm. Finally, in Section 6 we present some numerical results obtained by our approach, before we conclude with Section 7.

## 2. Multiobjective Optimal Control Problems and their discretization

In this paper we consider multiobjective optimal control problems of the following form:

Let $V$ be a Hilbert space and $Q$ a nonempty, closed and convex subset of $V$. Let $U^\nu, \nu = 1, \ldots, N$ be reflexive Banach spaces and $U^\nu_{ad} \subseteq U^\nu$ closed, bounded, and convex subsets. Define their product spaces $U := U^1 \times \ldots \times U^N$ and $U_{ad} := U^1_{ad} \times \ldots \times U^N_{ad}$, respectively. For $\nu = 1, \ldots, N$ let $J^\nu : V \times U^\nu \to \mathbb{R}$ be $N$ objective functions. Let a linear elliptic PDE constraint be given, which is stated in variational form as follows. Assume we have $f \in V'$, linear compact operators $B^\nu : U^\nu \to V', \nu = 1, \ldots, N$, the duality pairing $\langle \cdot, \cdot \rangle$ between $V'$ and $V$, and a bilinear form $a : V \times V \to \mathbb{R}$ such that

$$\textit{Find } \ y \in V : \quad a(y, v) = \left\langle \sum_{\nu=1}^{N} B^\nu u^\nu + f, v \right\rangle \quad \forall v \in V$$

is the weak formulation of the elliptic PDE constraint. To emphasize the role of the variable $u^\nu$, we use the notation $u = (u^1, \ldots, u^N) = (u^\nu, u^{-\nu})$, but we do not express a permutation by this notation. Then our multiobjective control problem reads

$$\begin{aligned} \min_{y, u^\nu} J^\nu(y, u^\nu) \qquad s.t. \qquad & a(y, v) = \left\langle \sum_{\nu=1}^{N} B^\nu u^\nu + f, v \right\rangle \quad \forall v \in V, \\ & u^\nu \in U^\nu_{ad}, \\ & y \in Q, \end{aligned} \tag{1}$$

for all $\nu = 1, \ldots, N$. A state $\bar{y}$ together with a vector of controls $(\bar{u}^1, \ldots, \bar{u}^N)$ is called a solution of (1), if they are feasible for all $N$ problems, i.e., $\bar{u}^\nu \in U^\nu_{ad}$ for all $\nu = 1, \ldots, N$, $\bar{y} \in Q$ and $a(\bar{y}, v) = \left\langle \sum_{\nu=1}^{N} B^\nu \bar{u}^\nu + f, v \right\rangle$ for all $v \in V$, and if $J^\nu(\bar{y}, \bar{u}^\nu) \leq J^\nu(y, u^\nu)$ for all other feasible $y$ and $(u^1, \ldots, u^N)$ and for all $\nu = 1, \ldots, N$.

A widely used setting is an open, bounded and convex polyhedral set $\Omega \subset \mathbb{R}^2$, the Hilbert space $V = H^1_0(\Omega)$ with its dual $V' = H^{-1}(\Omega)$ and the Banach spaces

$U^\nu = L^2(\Omega)$ for all $\nu = 1, \ldots, N$. There, the assumed compactness of the operators $B^\nu : U^\nu \to V', \nu = 1, \ldots, N$ is justified by the compact embedding of $L^2(\Omega)$ in $H^{-1}(\Omega)$. This is the setting considered in Hintermüller and Surowiec (2013) and it is more general than the one given in Borzi and Kanzow (2013), since it includes state constraints.

Throughout the paper we will assume that the variational equation

$$\textit{Find } y \in V: \quad a(y,v) = \left\langle \sum_{\nu=1}^{N} B^\nu u^\nu + f, v \right\rangle \quad \forall v \in V$$

has a unique solution for every given control $u$, which can be guaranteed by the Lax-Milgram Theorem if the bilinear form $a$ is bounded and coercive. This solution map will be denoted by $u \in U \mapsto y(u) \in V$. However, we will use this solution map for the theoretical parts only and we will not compute it in our numerical approach.

In order to solve the infinite dimensional problem (1) we will use a finite element discretization to obtain finite dimensional approximations, and the approximations must satisfy a number of assumptions to be introduced in the next section. We introduce for every $h \in (0,1)$

- finite-dimensional subspaces $V_h \subseteq V, V_h' \subseteq V', U_h^\nu \subseteq U^\nu$ and
$$U_h := U_h^1 \times \ldots \times U_h^N;$$
- nonempty, closed and convex subsets $Q_h \subseteq V_h$;
- nonempty, closed and convex subsets $U_{ad,h}^\nu \subseteq U_h^\nu$,
$$U_{ad,h} := U_{ad,h}^1 \times \ldots \times U_{ad,h}^N;$$
- bilinear forms $a^h : V_h \times V_h \to \mathbb{R}$;
- $f_h \in V_h'$, linear operators $B_h^\nu : U_h^\nu \to V_h', \nu = 1, \ldots, N$;
- the duality pairing $\langle \cdot, \cdot \rangle_h$ between $V_h'$ and $V_h$;
- objective functions $J_h^\nu : V_h \times U_h^\nu \to \mathbb{R}$ for all $\nu = 1, \ldots, N$.

Then, the discretized problem reads

$$\begin{aligned} \min_{y_h, u_h^\nu} J_h^\nu(y_h, u_h^\nu) \quad s.t. \quad & a^h(y_h, v_h) = \left\langle \sum_{\nu=1}^{N} B_h^\nu u_h^\nu + f_h, v_h \right\rangle_h \quad \forall v_h \in V_h, \\ & u_h^\nu \in U_{ad,h}^\nu, \\ & y_h \in Q_h, \end{aligned} \tag{2}$$

for all $\nu = 1, \ldots, N$. A solution of (2) is a state $\bar{y}_h$ together with a vector of controls $(\bar{u}_h^1, \ldots, \bar{u}_h^N)$ that is feasible for all $N$ problems and further satisfies $J_h^\nu(\bar{y}_h, \bar{u}_h^\nu) \le J_h^\nu(y_h, u_h^\nu)$ for all other feasible $y_h$ and $(u_h^1, \ldots, u_h^N)$ and for all $\nu = 1, \ldots, N$.

Note that in our numerical examples $\langle \cdot, \cdot \rangle_h$ will be an approximation of the $L^2(\Omega)$ scalar product using a quadrature formula.

# 3. The relation of the original and the discretized problem

Here we will clarify the relation between the problems (1) and (2). Furthermore, we will present a constructive existence result for a generalized Nash equilibrium of (1). Note that one can also use the approach from Hintermüller and Surowiec (2013) to obtain a solution of the problem (1) by a sequence of penalized Nash equilibrium problems.

Let $y_h := y_h(u_h)$ be the unique solution of the discretized variational equation

$$\textit{Find } y_h \in V_h: \quad a^h(y_h, v_h) = \left\langle \sum_{\nu=1}^{N} B_h^\nu u_h^\nu + f_h, v_h \right\rangle_h \quad \forall v_h \in V_h. \tag{3}$$

Introducing the sets

$$\begin{aligned} X^\nu(u^{-\nu}) &:= \{u^\nu \in U_{ad}^\nu \mid y(u^\nu, u^{-\nu}) \in Q\}, \\ X_h^\nu(u_h^{-\nu}) &:= \{u_h^\nu \in U_{ad,h}^\nu \mid y_h(u_h^\nu, u_h^{-\nu}) \in Q_h\}. \end{aligned}$$

leads to a compact reduced form of our original problem (1):

$$\min_{u^\nu} J^\nu(y(u), u^\nu) \quad s.t. \quad u^\nu \in X^\nu(u^{-\nu}) \quad \forall \nu = 1, \ldots, N, \tag{4}$$

which only depends on the control variables. Analogously, we obtain for the discretized problem (2)

$$\min_{u_h^\nu} J_h^\nu(y_h(u_h), u_h^\nu) \quad s.t. \quad u_h^\nu \in X_h^\nu(u_h^{-\nu}) \quad \forall \nu = 1, \ldots, N. \tag{5}$$

Now we can introduce the assumptions that we need for the analysis of the relation between these problems. We adapt and weaken the assumptions given in Haslinger and Neittaanmäki (1996) in the context of our setting. In the following, $y_h \to y$ means strong convergence, while weak convergence is denoted by $y_h \rightharpoonup y$. By $\|\cdot\|$ we denote the norm in $V$, by $\|\cdot\|_U$ the norm in $U$. For the families $\{V_h\}, \{U_{ad,h}\}, \{Q_h\}, \{a^h\}, \{f_h\}, \{B_h\}$ and $\{J_h^\nu\}$ we assume the following

- (H1) $\exists m > 0: a^h(v_h, v_h) \geq m\|v_h\|^2 \quad \forall v_h \in V_h, \forall h \in (0,1)$;
- (H2) $\exists M > 0: |a^h(y_h, v_h)| \leq M\|y_h\| \, \|v_h\| \quad \forall y_h, v_h \in V_h, \forall h \in (0,1)$;
- (H3) for any $\{v_h\}, \{y_h\}$ with $v_h, y_h \in V_h$ such that $y_h \rightharpoonup y, v_h \to v$ in $V$ we have $a^h(y_h, v_h) \to a(y,v)$ and $a^h(v_h, y_h) \to a(v,y)$ as $h \to 0+$;
- (H4) $\forall v \in V \, \exists \{v_h\}, v_h \in V_h: v_h \to v$ in $V$;
- (H5) $\exists C > 0: |\langle f_h, v_h \rangle_h| \leq C\|v_h\| \quad \forall v_h \in V_h, \forall h \in (0,1)$;
- (H6) for any $\{v_h\}, v_h \in V_h$ such that $v_h \rightharpoonup v$ in $V$ we have $\langle f_h, v_h \rangle_h \to \langle f, v \rangle$ as $h \to 0+$;

(H7) for any $\{u_h\}, u_h \in U_{ad,h}$ such that $\|u_h\|_U$ is bounded, $\exists \bar{C} > 0$ such that $\left|\left\langle \sum_{\nu=1}^N B_h^\nu u_h^\nu, v_h \right\rangle_h\right| \leq \bar{C}\|v_h\| \quad \forall v_h \in V_h, \forall h \in (0,1)$;

(H8) for any $\{u_h\}, \{v_h\}$, where $u_h \in U_{ad,h}, v_h \in V_h$ such that $u_h \rightharpoonup u$ in $U_{ad}$ and $v_h \rightharpoonup v$ in $V$, we have $\left\langle \sum_{\nu=1}^N B_h^\nu u_h^\nu, v_h \right\rangle_h \to \left\langle \sum_{\nu=1}^N B^\nu u^\nu, v \right\rangle$ as $h \to 0+$;

(H9) for any $\{y_h\}, y_h \in Q_h$ such that $y_h \to y$ in $V$ we have $y \in Q$;

(H10) for arbitrary $\nu = 1, \ldots, N$ and for all $u^\nu \in X^\nu(u^{-\nu})$ and all sequences $\{u_h^{-\nu}\}$ with $u_h^{-\nu} \rightharpoonup u^{-\nu}$ in $U_{ad}^{-\nu}$ and $X_h^\nu(u_h^{-\nu}) \neq \emptyset$, there exists a sequence $\{u_h^\nu\}$ with $u_h^\nu \in X_h^\nu(u_h^{-\nu})$ such that $u_h^\nu \to u^\nu$ in $U_{ad}^\nu$ as $h \to 0+$;

(H11) for all $\nu = 1, \ldots, N$ and any $\{u_h^\nu\}, \{y_h\}$, where $u_h^\nu \in U_h^\nu, y_h \in V_h$ such that $u_h^\nu \rightharpoonup u^\nu$ in $U_{ad}^\nu$ and $y_h \to y$ in $V$ we have $\liminf_{h\to 0+} J_h^\nu(y_h, u_h^\nu) \geq J^\nu(y, u^\nu)$ and moreover $J_h^\nu(y_h, u_h^\nu) \to J^\nu(y, u^\nu)$ as $h \to 0+$, if $u_h \to u$ in $U_{ad}$.

Note that in (H8) we already exploited the compactness assumption on $B^\nu$, $\nu = 1, \ldots, N$ in order to assume only $u_h \rightharpoonup u$ instead of $u_h \to u$. Using a standard finite element approach, i.e., defining $V_h$ and $U_{ad,h}$ via piecewise polynomial functions and using some quadrature formula for the discrete scalar product including the function $f$, one can show (H1) to (H8) as in the single objective case of PDE optimization for a number of commonly used settings, including the ones in the numerical section. Furthermore, (H9) can be obtained by using linear interpolation of the discrete pointwise defined state constraints. This can also be used to show the most critical assumption (H10) for our numerical examples, where we use piecewise linear functions for the discretization and the controls are active on disjoint subregions of $\Omega$ only. For (H11) we can choose $J_h^\nu \equiv J^\nu$ for all $\nu = 1, \ldots, N$, and we need some continuity properties of these functions, which are satisfied for the objective functions used in our examples in the numerical part.

Next we show a result, which is motivated by Haslinger and Neittaanmäki (1996, Theorem 10.3), and the proof uses the same techniques but with weaker assumptions.

Lemma 1 *Let (H1)-(H8) hold, and let a sequence $\{u_h\}, u_h \in U_h$ be given, such that $u_h \rightharpoonup u$ as $h \to 0+$. Then $y_h(u_h) \to y(u)$ in $V$.*

Proof By assumptions (H1), (H2) and the Lax-Milgram Theorem, $y_h := y_h(u_h)$ is the unique solution of the discretized variational equality (3). With (H1), (H5) and (H7) we can find positive constants $m, C, \bar{C}$ such that

$$m\|y_h\|^2 \stackrel{(H1)}{\leq} a^h(y_h, y_h) = \left\langle \sum_{\nu=1}^N B_h^\nu u_h^\nu + f_h, y_h \right\rangle_h \stackrel{(H5),(H7)}{\leq} \bar{C}\|y_h\| + C\|y_h\|.$$

This implies

$$\|y_h\| \leq \frac{\bar{C} + C}{m},$$

and hence the sequence $\{y_h\}$ is bounded in $V$. Therefore, we can find a $y \in V$ and a subsequence, which will be denoted by $\{y_h(u_h)\}$ again, that weakly converges to $y$, i.e., $y_h(u_h) \rightharpoonup y$ as $h \to 0+$.

Next, we will show that $y = y(u)$, meaning $a(y,v) = \left\langle \sum_{\nu=1}^N B^\nu u^\nu + f, v \right\rangle$ for all $v \in V$. For arbitrary $v \in V$ we get by (H4) a sequence $\{v_h\}, v_h \in V_h$ such that $v_h \to v$. With $u_h \rightharpoonup u, v_h \to v, y_h(u_h) \rightharpoonup y$ we can use (H3), (H6) and (H8) to obtain

$$\begin{aligned} a(y,v) &\overset{(H3)}{=} \lim_{h\to 0+} a^h(y_h, v_h) = \lim_{h \to 0+} \left\langle \sum_{\nu=1}^N B_h^\nu u_h^\nu + f_h, v_h \right\rangle_h \\ &\overset{(H6),(H8)}{=} \left\langle \sum_{\nu=1}^N B^\nu u^\nu + f, v \right\rangle. \end{aligned}$$

Since $v \in V$ was arbitrary, we have $y = y(u)$. By the uniqueness of the solution $y(u)$ the whole sequence $\{y_h(u_h)\}$ converges weakly to $y(u)$.

To finish the proof we have to show strong convergence. For $y(u) \in V$ we can find by (H4) a sequence $\{z_h\}, z_h \in V_h$ with $z_h \to y(u)$. Then, using (H1) we have

$$\begin{aligned} m\|y_h(u_h) - z_h\|^2 &\overset{(H1)}{\leq} a^h(y_h(u_h) - z_h, y_h(u_h) - z_h) \\ &= \left\langle \sum_{\nu=1}^N B_h^\nu u_h^\nu + f_h, y_h(u_h) - z_h \right\rangle_h - a^h(z_h, y_h(u_h) - z_h). \end{aligned}$$

Since $u_h \rightharpoonup u, y_h(u_h) \rightharpoonup y(u)$ and $z_h \to y(u)$, the first term on the right hand side has by (H6) and (H8) the limit

$$\left\langle \sum_{\nu=1}^N B^\nu u^\nu + f, y(u) - y(u) \right\rangle = 0.$$

The limit of the second second term is, by (H3)

$$a(y(u), y(u) - y(u)) = 0.$$

Therefore, we have

$$\|y_h(u_h) - y(u)\| \leq \|y_h(u_h) - z_h\| + \|z_h - y(u)\| \to 0,$$

as $h \to 0+$, which proves the strong convergence of $\{y_h(u_h)\}$ to $y(u)$ in $V$. □

With this Lemma we can show the following approximation result for the finite element discretization which extends Haslinger and Neittaanmäki (1996, Theorem 10.4) to the multiobjective case.

Theorem 1 *Let (H1)-(H11) hold and let* $(\hat{u}_h^1, \dots, \hat{u}_h^N, \hat{y}_h)$ *be a solution of the discretized problem (5). Then there exists a subsequence* $(\hat{u}_{h'}^1, \dots, \hat{u}_{h'}^N, \hat{y}_{h'})$ *and further* $(\hat{u}^1, \dots, \hat{u}^N) \in U_{ad}$ *and* $\hat{y} \in V$ *such that* $\hat{y}_{h'} \to \hat{y}, \hat{u}_{h'}^\nu \to \hat{u}^\nu$ *for all* $\nu = 1, \dots, N$ *and* $(\hat{u}^1, \dots, \hat{u}^N, \hat{y})$ *is a solution of the problem (4).*

Proof Let $(\hat{u}_h^1, \dots, \hat{u}_h^N, \hat{y}_h)$ be a solution of the discretized problem (5). Since $U$ is bounded and $U_{ad,h} \subseteq U_h \subseteq U$, we can find a weakly convergent subsequence

$$\hat{u}_{h'} := (\hat{u}_{h'}^1, \dots, \hat{u}_{h'}^N) \rightharpoonup (\hat{u}^1, \dots, \hat{u}^N) =: \hat{u} \in U_{ad}.$$

Now, using Lemma 1 we obtain

$$\hat{y}_{h'} = y_{h'}(\hat{u}_{h'}) \to y(\hat{u}) =: \hat{y}$$

in $V$. Furthermore by (H9) we have $\hat{y} = y(\hat{u}) \in Q$, and hence also

$$\hat{u}^\nu \in X^\nu(\hat{u}^{-\nu}) = \{u^\nu \in U_{ad}^\nu \mid y(u^\nu, \hat{u}^{-\nu}) \in Q\}$$

for all $\nu = 1, \dots, N$. $(\hat{u}_{h'}^1, \dots, \hat{u}_{h'}^N, \hat{y}_{h'})$ being a solution of the discretized problem (5) implies for all $\nu = 1, \dots, N$:

$$J_{h'}^\nu(y_{h'}(\hat{u}_{h'}^\nu, \hat{u}_{h'}^{-\nu}), \hat{u}_{h'}^\nu) \le J_{h'}^\nu(y_{h'}(w_{h'}^\nu, \hat{u}_{h'}^{-\nu}), w_{h'}^\nu) \quad \forall w_{h'}^\nu \in X_{h'}^\nu(\hat{u}_{h'}^{-\nu}). \tag{6}$$

Let $\nu \in \{1, \dots, N\}$ be fixed. Using (H10) we can find for any $u^\nu \in X^\nu(\hat{u}^{-\nu})$ and the sequence $\hat{u}_{h'}^{-\nu} \rightharpoonup \hat{u}^{-\nu}$, which satisfies $\hat{u}_{h'}^\nu \in X_{h'}^\nu(\hat{u}_{h'}^{-\nu}) \neq \emptyset$, a sequence $\{u_{h'}^\nu\}$ with $u_{h'}^\nu \in X_{h'}^\nu(\hat{u}_{h'}^{-\nu})$ and $u_{h'}^\nu \to u^\nu$. Using Lemma 1 once again, we get $y_{h'}(u_{h'}^\nu, \hat{u}_{h'}^{-\nu}) \to y(u^\nu, \hat{u}^{-\nu})$.

Now, taking the limit $h' \to 0+$ in (6) and using $u_{h'}^\nu \in X_{h'}^\nu(\hat{u}_{h'}^{-\nu})$ together with (H11) we obtain

$$\begin{aligned} J^\nu(y(\hat{u}^\nu, \hat{u}^{-\nu}), \hat{u}^\nu) &\overset{(H11)}{\le} \liminf_{h' \to 0+} J_{h'}^\nu(y_{h'}(\hat{u}_{h'}^\nu, \hat{u}_{h'}^{-\nu}), \hat{u}_{h'}^\nu) \\ &\overset{(6)}{\le} \liminf_{h' \to 0+} J_{h'}^\nu(y_{h'}(u_{h'}^\nu, \hat{u}_{h'}^{-\nu}), u_{h'}^\nu) \\ &\overset{(H11)}{=} J^\nu(y(u^\nu, \hat{u}^{-\nu}), u^\nu). \end{aligned}$$

Since $u^\nu \in X^\nu(\hat{u}^{-\nu})$ was arbitrary, and the arguments hold for all $\nu = 1, \dots, N$, this implies that $(\hat{u}^1, \dots, \hat{u}^N, \hat{y})$ is a solution of the problem (4) and hence completes the proof. □

Note that instead of (H1) - (H8) we can also assume the assertion of Lemma 1 to prove this theorem. Therefore we get the following existence result.

Proposition 1 *Let (H9)-(H11) hold and assume that for any given sequence* $\{u_h\}, u_h \in U_h$ *with* $u_h \rightharpoonup u$ *in* $U$ *as* $h \to 0+$ *we have* $y_h(u_h) \to y(u)$ *in* $V$*. Then there exists a solution of (4), if every discretized problem (5) has a solution.*

It remains to prove the existence of a solution for the discretized finite-dimensional GNEP (5). Here we can use some known results, see Facchini and Kanzow (2007, Theorem 4.1) going back to Arrow and Debreu (1954) and Ichishi (1983), which we state in our notation.

Lemma 2 *Let a finite dimensional GNEP be given. Assume that the functions* $J_h^\nu$ *are continuous and* $J_h^\nu(y_h(\cdot, u_h^{-\nu}), \cdot, u_h^{-\nu})$ *are quasi-convex on* $X_h^\nu(u_h^{-\nu})$ *for every* $\nu = 1, \dots, N$. *Further, let nonempty, convex, and compact sets* $K_\nu \subseteq \mathbb{R}^{dim(U_{ad,h}^\nu)}$ *for* $\nu = 1, \dots, N$ *exist, such that for every* $u_h$ *with* $u_h^\nu \in K_\nu$ *for every* $\nu$, *the sets* $X_h^\nu(u_h^{-\nu})$ *are nonempty, closed, and convex subsets of* $K_\nu$. *Moreover let* $X_h^\nu$, *as a point-to-set map, be upper and lower semicontinuous. Then there exists a generalized Nash equilibrium.*

Note that the notation of upper and lower semicontinuity is in the sense of Berge, see Berge (1963). Using this Lemma we can prove the following existence result.

Theorem 2 *Assume that the functions* $J_h^\nu$ *are continuous and* $J_h^\nu(y_h(\cdot, u_h^{-\nu}), \cdot)$ *are quasi-convex on* $X_h^\nu(u_h^{-\nu})$ *for every* $\nu = 1, \dots, N$. *Let* $U_{ad,h}^\nu$ *be nonempty, convex, and compact and assume that there is a continuous and in each* $u_h^\nu$, $\nu = 1, \dots, N$ *(separately not jointly) convex function* $g_h : U_{ad,h} \to \mathbb{R}^\ell$ *for some* $\ell \in \mathbb{N}$ *such that*

$$X_h^\nu(u_h^{-\nu}) = \{u_h^\nu \in U_{ad,h}^\nu \mid g_h(u_h^\nu, u_h^{-\nu}) \le 0\}.$$

*Further, let for all* $\nu = 1, \dots, N$ *the sets* $X_h^\nu(u_h^{-\nu})$ *satisfy a Slater condition for all* $u_h^{-\nu} \in U_{ad,h}^{-\nu}$, *i.e.,*

$$\forall u_h^{-\nu} \in U_{ad,h}^{-\nu} \ \exists u_h^\nu \in U_{ad,h}^\nu : \ g_h(u_h^\nu, u_h^{-\nu}) < 0.$$

*Then the GNEP (5) has a solution.*

Proof We want to use Lemma 2 with $K_\nu := U_{ad,h}^\nu$. By assumption, the sets $U_{ad,h}^\nu$ are nonempty, convex and compact and $X_h^\nu(u_h^{-\nu}) \subseteq U_{ad,h}^\nu$. By the continuity and convexity of $g$, the set $X_h^\nu(u_h^{-\nu})$ is closed and convex and by the assumed Slater condition also nonempty. Furthermore, the continuity of $g_h$ implies the closedness of the point-to-set map $X_h^\nu$ by Hogan (1973, Theorem 10) and by the compactness of $U_{ad,h}$ this implies upper semicontinuity. By Hogan (1973, Theorem 12) the Slater condition and the continuity and convexity of $g_h$ imply that the point-to-set map $X_h^\nu$ is an open mapping in the sense of Hogan (1973), which is equivalent to lower semicontinuity. Therefore, Lemma 2 shows the existence of a solution of the GNEP (5). □

To close this section let us remark on two issues:

- Let us consider the case, where the discretized variational equation is a linear equation system, the discretized objective functions are quadratic and the feasible sets $U^\nu_{ad,h}$ for the controls are box constrained, hence nonempty, convex and compact. Further assume that the constraints on the state can be expressed by $y_h(u_h) \geq \psi$. Then we only need the Slater condition to be satisfied, which becomes

$$\forall u_h^{-\nu} \in U_{ad,h}^{-\nu} \, \exists u_h^\nu \in U_{ad,h}^\nu : y_h(u_h^\nu, u_h^{-\nu}) > \psi.$$

By the compactness of $U_{ad,h}$, this is similar to the *strict uniform feasible response constraint qualification SUFR* from Hintemüller and Surowiec (2013) in our particular context of this finite dimensional GNEP. However, it does not imply SUFR, since we do not have a uniform bound here.

- One can also consider the more general setting, where the objective functions may depend directly on the control variables of the other players, i.e., $J^\nu(y, u^\nu, u^{-\nu})$. Introducing the additional assumption that $u^{-\nu}$ occurs in the objective function only together with a compact linear operator $C^{-\nu}$, meaning that we actually have $J^\nu(y, u^\nu, C^{-\nu}u^{-\nu})$, we can obtain all the results above, with a simple modification at the end of the proof of Theorem 1.

# 4. The algorithm

In this section we are interested in a method to solve the discretized problems. We propose a variant of a potential reduction algorithm, which was earlier used to solve finite dimensional GNEPs, see Dreves et al. (2011), or quasi-variational inequality problems, see Facchinei, Kanzow and Sagratella (2014). Since we have the same state variable $y$ for all players, we are not in the standard form of a GNEP. Now, the idea is to introduce a state variable $y^\nu$ for each player $\nu = 1, \ldots, N$, which will be all equal at a solution, since we assume a unique solution of the variational equation. By this approach we can use some concepts and ideas from the GNEP theory.

Since we always use the discretized problem for some fixed $h$ in this and the next section, we will skip the discretization index $h$ from now on, i.e., we will write $y, u^\nu, \ldots$ instead of $y_h, u_h^\nu, \ldots$.

Let $u^1, \ldots, u^N, y^1, \ldots, y^N$ be discretized so that $u^\nu \in \mathbb{R}^{n_\nu}$ and $y^\nu \in \mathbb{R}^{n'}$ for all $\nu = 1, \ldots, N$. Define also $n := N \cdot n' + \sum_{\nu=1}^N n_\nu$ as the total number of primal variables. Let the discretization of the variational equality result in an equation of the form $\bar{h}(y^\nu, u^\nu, u^{-\nu}) = 0$ for a function $\bar{h} : \mathbb{R}^{n'+n_1+\ldots+n_N} \to \mathbb{R}^{p'}$. If we use, for example, a Ritz-Galerkin ansatz with nodal basis functions, i.e., basis functions $\phi_i$ that are 1 at node $i$ and zero at all the other nodes, we get an affine linear function $\bar{h}$ here. Furthermore, we assume the existence of a function $g^\nu : \mathbb{R}^{n'+n_\nu} \to \mathbb{R}^{m_\nu}$ that describes the constraints $u^\nu \in U^\nu_{ad}, y^\nu \in Q$ by $g^\nu(y^\nu, u^\nu) \leq 0$. Hence, our discretized problem becomes

$$\min_{y^\nu, u^\nu} J^\nu(y^\nu, u^\nu) \quad s.t. \quad \begin{aligned} \bar{h}(y^\nu, u^\nu, u^{-\nu}) &= 0, \\ g^\nu(y^\nu, u^\nu) &\le 0, \end{aligned} \tag{7}$$

for $\nu = 1, \ldots, N$. We will use the notation $\nabla_{u^\nu} \bar{h}(y^\nu, u^\nu, u^{-\nu}) \in \mathbb{R}^{n_\nu \times p'}$ for the matrix that has in column $i$ the partial gradient of the component function $h_i$ with respect to $u^\nu$, whereas the notation $J_{u^\nu} \bar{h} \in \mathbb{R}^{p' \times n_\nu}$ is used for the Jacobian of $\bar{h}$ with respect to $u^\nu$. Upon setting $p := Np'$ and $m := m_1 + \ldots + m_N$ we can define the function $H : \mathbb{R}^{n+p+m+m} \to \mathbb{R}^{n+p+m+m}$ by

$$H((u^1, y^1, \ldots, u^N, y^N), (\mu^1, \ldots, \mu^N), (\lambda^1, \ldots, \lambda^N), (w^1, \ldots, w^N)) :=$$
$$\begin{pmatrix} \left[\nabla_{(u^\nu, y^\nu)} J^\nu(y^\nu, u^\nu) + \nabla_{(u^\nu, y^\nu)} \bar{h}(y^\nu, u^\nu, u^{-\nu})\, \mu^\nu + \nabla_{(u^\nu, y^\nu)} g^\nu(y^\nu, u^\nu)\, \lambda^\nu\right]_{\nu=1}^N \\ \left[\bar{h}(y^\nu, u^\nu, u^{-\nu})\right]_{\nu=1}^N \\ \left[g^\nu(y^\nu, u^\nu) + w^\nu\right]_{\nu=1}^N \\ \left[\lambda^\nu \circ w^\nu\right]_{\nu=1}^N, \end{pmatrix}$$

with the componentwise Hadamard product $w \circ \lambda$. For a short notation we define

$$\begin{aligned}
x &:= (u^1, y^1, u^2, y^2, \ldots, u^N, y^N), \\
\mu &:= (\mu^1, \ldots, \mu^N), \\
\lambda &:= (\lambda^1, \ldots, \lambda^N), \\
w &:= (w^1, \ldots, w^N), \\
z &:= (x, \mu, \lambda, w), \\
h(x) &:= \left[\bar{h}(y^\nu, u^\nu, u^{-\nu})\right]_{\nu=1}^N, \\
g(x) &:= \left[g^\nu(y^\nu, u^\nu)\right]_{\nu=1}^N, \\
F(x, \mu, \lambda) &:= \left[\nabla_{(u^\nu, y^\nu)}(J^\nu(y^\nu, u^\nu) + h(y^\nu, u^\nu, u^{-\nu})\, \mu^\nu + g^\nu(y^\nu, u^\nu)\, \lambda^\nu)\right]_{\nu=1}^N \\
\Lambda &:= diag(\lambda), \\
W &:= diag(w), \\
E_h(x) &:= \begin{pmatrix} \nabla_{(u^1, y^1)} \bar{h}(y^1, u^1, u^{-1}) & & \\ & \ddots & \\ & & \nabla_{(u^N, y^N)} \bar{h}(y^N, u^N, u^{-N}) \end{pmatrix}.
\end{aligned}$$

Note that we do not define $E_g(x)$ similarly to $E_h(x)$, since by the structure of the function $g$ we have $E_g(x) = Jg(x)^\top$. Using our notation we obtain

$$H(x, \mu, \lambda, w) = \begin{pmatrix} F(x, \mu, \lambda) \\ h(x) \\ g(x) + w \\ \lambda \circ w \end{pmatrix}.$$

Now, every solution of the concatenated KKT system of the problem (7) is a solution of the constrained equation

$$H(z) = 0, \quad z \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m_+ \times \mathbb{R}^m_+. \tag{8}$$

For the solution of this constrained equation we want to give a short description of the interior point algorithm from Dreves et al. (2011), which is based on the algorithm first proposed in Monteiro and Pang (1999). In contrast to Dreves et al. (2011), we now have equality constraints in our setting. Let

$$Z_I := \{z = (x, \mu, \lambda, w) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m_{++} \times \mathbb{R}^m_{++} \mid g(x) + w > 0\}$$

denote the set of all strictly feasible points, for which the last $2m$ components of $H$ are positive. Then we define for a given real number $\zeta > m$ the potential function $\psi : Z_I \to \mathbb{R}$ by

$$\psi(z) := \zeta \log(\|H(z)\|^2) - \sum_{i=n+p+1}^{n+p+2m} \log(H_i(z)).$$

With $a^\top := (0^\top_{n+p}, 1^\top_{2m})$ we can now present the algorithm from Dreves et al. (2011) including equality constraints.

**Algorithm 1: Inexact Potential Reduction Algorithm for GNEPs**

(S.0): Choose $z^0 \in Z_I$ and $\beta, \gamma, \in (0,1), \zeta > m$. Set $k := 0$.

(S.1): If $H(z^k) = 0$ then STOP.

(S.2): Choose $\sigma_k \in [0,1), \eta_k \geq 0$ and compute a solution $d^k$ of

$$\left\| H(z^k) - JH(z^k)d - \sigma_k \frac{a^\top H(z^k)}{\|a\|^2} a \right\| \leq \eta_k \|H(z^k)\| \qquad \text{and}$$
$$\nabla\psi(z^k)^\top d^k < 0.$$

(S.3): Compute a stepsize $t_k := \max\{\beta^i \mid i = 0, 1, 2, \ldots\}$ such that

$$z^k + t_k d^k \in Z_I \quad \text{and}$$
$$\psi(z^k + t_k d^k) \leq \psi(z^k) + \gamma t_k \nabla\psi(z^k)^\top d^k.$$

(S.4): Set $z^{k+1} := z^k + t_k d^k$, $k := k + 1$, and go to (S.1).

If $JH(z^k)$ is nonsingular, we can always find a solution $d^k$ of the linear equation system in (S.2). By Monteiro and Pang (1999) this solution is a direction of descent for the potential function $\psi$, which leads to the name Potential Reduction Algorithm. Since $Z_I$ is an open set and $d^k$ is a direction of descent, according to the line search rule, and hence the entire Algorithm 1 is well-defined, if $JH(z^k)$ is nonsingular for all $z^k$. The next result follows from Dreves et al. (2011, Theorem 4.3), whose proof can be extended to our setting with the additional equations. It requires a non restrictive technical assumption $\limsup_{k\to\infty} \sigma_k < 1$, which guarantees some uniform descent.

Theorem 3 *Assume that $JH(z)$ is nonsingular for all $z \in Z_I$ and that the sequences $\{\sigma_k\}$ and $\{\eta_k\}$ satisfy the conditions*

$$\limsup_{k\to\infty} \sigma_k < 1 \qquad \textit{and} \qquad \lim_{k\to\infty} \eta_k = 0.$$

*Let $\{z^k\}$ be any sequence generated by Algorithm 1. Then the following assertions hold:*
*(a) The sequence $\{H(z^k)\}$ is bounded.*
*(b) Any accumulation point of $\{z^k\}$ is a solution of (8).*

By $z^0 \in Z_I$ and due to the step size rule in (S.3) of Algorithm 1 all iterates $z^k$ are contained in $Z_I$. Therefore, the nonsingularity of $JH(z)$ on $Z_I$ guarantees that Algorithm 1 is well defined. Note that none of the solutions of the system $H(z) = 0$ does belong to $Z_I$, so that the nonsingularity of $JH(z)$ on $Z_I$ does not imply the nonsingularity of the Jacobian $JH$ at a KKT point. Further, note that setting $\eta_k = 0$ for all $k \in \mathbb{N}_0$ we have an exact version of this algorithm. In our numerical section we observe that such a version is faster for our test problems than using an iterative solver, like gmres, for the approximate solution of the equation system in (S.2). However, if the discretized problems are very large, but manageable by the memory constraints, this may change. Beside some weak technical assumption, the convergence of the inexact algorithm requires only the nonsingularity of $JH(z)$. Furthermore, let us mention that if one aims to solve the discretized problem at a high precision, one can use the hybrid method suggested in Dreves et al. (2014), which combines the potential reduction algorithm with an LP-Newton method to inherit its local quadratic rate of convergence.

## 5. Nonsingularity conditions

Algorithm 1 is well defined only, if the Jacobian of the function $H$ is nonsingular for all iterates $z \in Z_I$. The Jacobian of $H$ is in our setting given by

$$JH(z) = \begin{pmatrix} J_x F(x,\mu,\lambda) & E_h(x) & Jg^\top(x) & 0 \\ Jh(x) & 0 & 0 & 0 \\ Jg(x) & 0 & 0 & I_m \\ 0 & 0 & W & \Lambda \end{pmatrix},$$

and we are interested in conditions guaranteeing the nonsingularity of this matrix only for all $z \in Z_I$. For an easier reading we often drop the arguments of the functions in the following, if they are clear from the context. As a first step we want to use the condition given in Dreves et al. (2011), where, however, no equality constraints were used. To do so, we repeat the definition of a $P_0$-matrix.

Definition 1 *A matrix $M \in \mathbb{R}^{n\times n}$ is called a $P_0$-matrix if* $\det(M_{\alpha\alpha}) \geq 0$ *for all $\alpha \subseteq \{1, 2, \ldots, n\}$.*

Note that the class of $P_0$-matrices strictly includes the positive semidefinite matrices, see Cottle, Pang and Stone (1992).

Lemma 3 *The matrix $JH(z)$ is nonsingular for all $z \in Z_I$, if the matrices $J_xF$ and $Jh(J_xF)^{-1}E_h$ are nonsingular and the matrix*

$$Jg\,[(J_xF)^{-1} - (J_xF)^{-1}\,E_h\,(Jh\,(J_xF)^{-1}\,E_h)^{-1}\,Jh\,(J_xF)^{-1}]\,Jg^\top$$

*is a $P_0$-matrix.*

Proof If we use the function $\bar{F}(x, \mu, \lambda) = \begin{pmatrix} F(x, \mu, \lambda) \\ h(x) \end{pmatrix}$ to describe all equalities in the KKT system, we are in the setting of Dreves et al. (2011, Theorem 4.6). Therefore the matrix $JH(z)$ is nonsingular for all $z \in Z_I$, if $J_{(x,\mu)}\bar{F}$ is nonsingular and

$$(Jg, 0)(J_{(x,\mu)}\bar{F}(x, \mu, \lambda))^{-1}\begin{pmatrix} Jg^\top \\ 0 \end{pmatrix}$$

is a $P_0$-matrix. Let us first show that the matrix $J_{(x,\mu)}\bar{F} = \begin{pmatrix} J_xF & E_h \\ Jh & 0 \end{pmatrix}$ is nonsingular. If we consider the equation system

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = J_{(x,\mu)}\bar{F}\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} J_xF & E_h \\ Jh & 0 \end{pmatrix}\begin{pmatrix} d_1 \\ d_2 \end{pmatrix},$$

and use the nonsingularity of $J_xF$, we get from the first equation

$$d_1 = -(J_xF)^{-1}\,E_h\,d_2$$

and hence from the second one

$$0 = Jh\,d_1 = -(Jh\,(J_xF)^{-1}\,E_h)d_2.$$

The assumed nonsingularity of $Jh\,(J_xF)^{-1}\,E_h$ implies $d_2 = 0$ and thus also $d_1 = 0$ follows. This proves that $J_{(x,\mu)}\bar{F}$ is nonsingular.

Furthermore, we have to show that $(Jg, 0)(J_{(x,\mu)}\bar{F}(x, \mu, \lambda))^{-1}\begin{pmatrix} Jg^\top \\ 0 \end{pmatrix}$ is a $P_0$-matrix. For this purpose, we will make use of the following general algebraic representation of the inverse of a special blockmatrix, which can be easily verified:

$$\begin{pmatrix} A & B \\ C & 0 \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} - A^{-1}B(CA^{-1}B)^{-1}CA^{-1} & A^{-1}B(CA^{-1}B)^{-1} \\ (CA^{-1}B)^{-1}CA^{-1} & -(CA^{-1}B)^{-1} \end{pmatrix}. \quad (9)$$

With this formula, we obtain

$$\begin{aligned} & (Jg, 0) \begin{pmatrix} J_xF & E_h \\ Jh & 0 \end{pmatrix}^{-1} \begin{pmatrix} Jg^\top \\ 0 \end{pmatrix} \\ = \;& Jg[(J_xF)^{-1} - (J_xF)^{-1}E_h(Jh(J_xF)^{-1}E_h)^{-1}Jh(J_xF)^{-1}]Jg^\top. \end{aligned}$$

By assumption, this is a $P_0$-matrix, which completes the proof. □

It is possible to reduce the equation system $JH\,d = b$ to a smaller one, in the following way: by splitting the system we obtain

$$\begin{aligned} J_xF\,d_1 + E_h\,d_2 + Jg^\top d_3 &= b_1, \\ Jh\,d_1 &= b_2, \\ Jg\,d_1 + d_4 &= b_3, \\ W\,d_3 + \Lambda\,d_4 &= b_4. \end{aligned}$$

Now, the third equation gives

$$d_4 = b_3 - Jg\,d_1$$

and, since we assume that we are in $Z_I$ and hence all entries of the diagonal matrix $W$ are positive, we get from the last equality

$$d_3 = W^{-1}(b_4 - \Lambda\,d_4) = W^{-1}b_4 - W^{-1}\Lambda\,b_3 + W^{-1}\Lambda\,Jg\,d_1.$$

By inserting this into the first equality we get a system in $d_1$ and $d_2$ only, namely

$$\begin{pmatrix} J_xF + Jg^\top W^{-1}\Lambda\,Jg & E_h \\ Jh & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} b_1 - Jg^\top(W^{-1}\,b_4 - W^{-1}\Lambda\,b_3) \\ b_2 \end{pmatrix}.$$

Hence, we have shown the following

LEMMA 4 *The matrix $JH(z)$ is nonsingular for all $z \in Z_I$, if and only if*

$$\begin{pmatrix} J_xF + Jg^\top W^{-1}\Lambda\,Jg & E_h \\ Jh & 0 \end{pmatrix}$$

*is nonsingular for all $z \in Z_I$.*

Now we are in the position to show a further nonsingularity result.

THEOREM 4 *The matrix $JH(z)$ is nonsingular for all $z \in \Omega$, if the matrices $J_xF + Jg^\top W^{-1}\Lambda\,Jg$ and $Jh\,(J_xF + Jg^\top W^{-1}\Lambda\,Jg)^{-1}E_h$ are nonsingular.*

PROOF The assertion follows from the previous Lemma 4 and the formula of the inverse given in (9), which can be used since, by assumption

$$A = J_xF + Jg^\top W^{-1}\Lambda\, Jg \quad \text{and} \quad CA^{-1}B = Jh\,(J_xF + Jg^\top W^{-1}\Lambda\, Jg)^{-1}E_h$$

are nonsingular. □

COROLLARY 1 *Assume that $J_xF$ is positive semidefinite and*

$$d^\top J_xF\, d > 0 \quad \forall d \in \{d \in \mathbb{R}^n \setminus \{0\} \mid Jg\, d = 0\}.$$

*Then, nonsingularity of $Jh\,(J_xF + Jg^\top W^{-1}\Lambda\, Jg)^{-1}E_h$ implies that $JH(z)$ is nonsingular for all $z \in Z_I$.*

PROOF Since $W$ and $\Lambda$ are diagonal matrices with positive entries, the matrix $Jg^\top W^{-1}\Lambda\, Jg$ is positive semidefinite for all $z \in Z_I$. Then, positive semidefiniteness of $J_xF$, together with

$$d^\top J_xF\, d > 0 \quad \forall d \in \{d \in \mathbb{R}^n \setminus \{0\} \mid Jg\, d = 0\}$$

is sufficient for nonsingularity of $J_xF + Jg^\top W^{-1}\Lambda\, Jg$. Thus, the assertion follows from Theorem 4. □

For the rest of this section let us consider problems with objective functions of the form

$$J^\nu(y^\nu, u^\nu) = \frac{1}{2}\|y^\nu - z^\nu\|^2_{L^2(\Omega)} + \frac{1}{2}\|u^\nu\|^2_{L^2(\Omega)}$$

for given $z^\nu \in L^2(\Omega)$. Further, we will assume that we have a Ritz-Galerkin discretization, resulting in the discrete objective functions of the form

$$J^\nu(y^\nu, u^\nu) = \frac{1}{2}(y^\nu - z^\nu)^\top P_y(y^\nu - z^\nu) + \frac{1}{2}(u^\nu)^\top P_\nu u^\nu$$

with symmetric positive definite matrices $P_y, P_\nu, \nu = 1, \ldots, N$, and the variational equality from the constraints is an affine linear equation

$$A\, y^\nu = \sum_{\nu=1}^{N} M_\nu\, u^\nu + F$$

for some matrices $A, M_\nu, \nu = 1, \ldots, N$ with $A$ being positive definite.

COROLLARY 2 *For discrete problems having the above form, the Jacobian $JH(z)$ is nonsingular for all $z \in Z_I$.*

PROOF By the properties of the discretization, the matrices $P_y, P_1, \ldots, P_N$ are symmetric positive definite and therefore also

$$J_x F = \text{blockdiag}(P_1, P_y, P_2, P_y, \ldots, P_N, P_y)$$

is symmetric positive definite. Together with the positive semidefiniteness of $Jg^\top W^{-1}\Lambda\, Jg$ this shows that $J_x F + Jg^\top W^{-1}\Lambda\, Jg$ is symmetric positive definite. Since $Jg^\top W^{-1}\Lambda\, Jg$ has the same blockdiagonal structure as $J_x F$, we can define the inverse

$$(J_x F + Jg^\top W^{-1}\Lambda\, Jg)^{-1} =: \text{blockdiag}(Q_1, Q_y, Q_2, Q_y, \ldots, Q_N, Q_y),$$

which has also symmetric positive definite blocks $Q_y, Q_1, \ldots, Q_N$. By Corollary 1 we only have to show that $Jh\,(J_x F + Jg^\top W^{-1}\Lambda\, Jg)^{-1} E_h$ is nonsingular. Using the fact that $A$ is symmetric we can compute

$$Jh = \begin{pmatrix} -M_1 & A & -M_2 & 0 & \ldots & \ldots & -M_N & 0 \\ -M_1 & 0 & -M_2 & A & \ldots & \ldots & -M_N & 0 \\ \vdots & \vdots & \vdots & \vdots & & & \vdots & \vdots \\ -M_1 & 0 & -M_2 & 0 & \ldots & \ldots & -M_N & A \end{pmatrix},$$

$$E_h = \begin{pmatrix} -M_1^\top & & & \\ A & & & \\ & -M_2^\top & & \\ & A & & \\ & & \ddots & \\ & & & -M_N^\top \\ & & & A \end{pmatrix}$$

and then

$$\hat{M} := Jh\,(J_x F + E_g\, W^{-1}\Lambda\, Jg)^{-1} E_h =$$
$$\begin{pmatrix} M_1 Q_1 M_1^\top + A Q_y A & M_2 Q_2 M_2^\top & \ldots & M_N Q_N M_N^\top \\ M_1 Q_1 M_1^\top & M_2 Q_2 M_2^\top + A Q_y A & \ldots & M_N Q_N M_N^\top \\ \vdots & \vdots & & \vdots \\ M_1 Q_1 M_1^\top & M_2 Q_2 M_2^\top & \ldots & M_N Q_N M_N^\top + A Q_y A \end{pmatrix}.$$

Now, let a vector $v = (v_1^\top, \ldots, v_N^\top)^\top$ be given, such that $\hat{M}v = 0$. By subtracting the $i$-th row from the first one we get

$$A Q_y A(v_1 - v_i) = 0 \quad \forall i = 2, \ldots, N,$$

and therefore, by the positive definiteness of $A$ and $Q_y$, we obtain $v_1 = v_i$ for all $i = 2, \ldots, N$. This implies

$$0 = v^\top \hat{M} v = N v_1^\top A Q_y A v_1 + N \sum_{\nu=1}^{N} v_1^\top M_\nu Q_\nu M_\nu^\top v_1.$$

Since $Q_y, Q_\nu$ and $A$ are positive definite and $M_\nu Q_\nu M_\nu^\top$ is positive semidefinite, $v_1 = 0$ must hold. Hence we have $v = 0$ and this shows that the matrix $Jh\,(J_xF + Jg^\top W^{-1}\,\Lambda\,Jg)^{-1}\,E_h$ is nonsingular and by Corollary 1 the matrix $JH$ is nonsingular for all $z \in Z_I$. □

At the end of this section we want to show that for the exact version of Algorithm 1, i.e., the version with $\eta_k = 0$ for all $k \in \mathbb{N}$, the states $y^\nu, \nu = 1, \ldots, N$, are equal at all iterates, provided they are equal at the beginning. When we solve the linear equation system in (S.2) of Algorithm 1, one part to consider is

$$Jh(x)\,d = h(x) \quad \Leftrightarrow \quad Ad_{y^\nu} - \sum_{\nu=1}^{N} M_\nu d_{u^\nu} = Ay^\nu - \sum_{\nu=1}^{N} M_\nu u^\nu - F \quad \forall \nu = 1, \ldots, N.$$

Subtracting the $i$-th equation from the first one for all $i = 2, \ldots, N$, yields

$$A(d_{y^1} - d_{y^i}) = A(y^1 - y^i),$$

and by the nonsingularity of $A$ we obtain $d_{y^1} - d_{y^i} = y^1 - y^i$ for all $i = 2, \ldots, N$. Hence, we have shown the following

Lemma 5 *If the starting values for all $y^\nu, \nu = 1, \ldots, N$ are equal, these values stay equal at all iterates of the exact version of Algorithm 1. In particular, we get equal values $y^1 = \ldots = y^N$ at the solution of the discretized problem.*

# 6. Numerical results

To show the numerical applicability of our approach we made a straightforward implementation of our Algorithm 1 in Matlab. As parameters we choose $\beta = 0.5$, $\gamma = 0.01, \zeta = 2m$ and $\sigma_0 = 0.9$, and $\sigma_k = 0.1, k \in \mathbb{N}$. Further, we set $\eta_k = 0$ for all $k \in \mathbb{N}_0$, since we observed that for our test problems this exact version is faster. We used a minimum value of $10^{-14}$ for $\lambda, w$ to ensure feasibility. As a first test example we used the one described in Borzi and Kanzow (2013). Although we do not have state constraints here and hence more efficient methods for the solution of this problem exist, we used it since it has a known analytical solution, and hence the computed errors are observable. Further, we tested a problem coming from a preprint version of Hintermüller and Surowiec (2013), which is described to have one of the major difficulties that can arise when solving GNEPs, meaning nontrivial biactive sets, and a third example having low multiplier regularity for the state constraints. As starting vectors we always used $x^0 = 0, \mu^0 = 0, \lambda^0 = 10$ and $w^0 = \max\{10, 5 - g(x^0)\}$ and we stopped the algorithm if $\|H(z)\|_\infty < 10^{-10}$.

The problems always have the form

$$\min_{y^\nu, u^\nu} \frac{1}{2}\|y^\nu - z^\nu\|^2_{L^2(\Omega)} + \frac{\alpha_\nu}{2}\|\chi_{B_\nu} u^\nu\|^2_{L^2(\Omega)} \quad s.t. \quad \begin{aligned} & -\Delta y^\nu = \sum_{\nu=1}^{N} \chi_{B_\nu} u^\nu + f \quad in\ \Omega, \\ & y^\nu = 0 \quad on\ \partial\Omega, \\ & l_\nu \le u^\nu \le r_\nu \quad a.e.\ in\ \Omega, \\ & \psi_l \le y^\nu \le \psi_r \quad a.e.\ in\ \Omega, \end{aligned}$$

for $\nu = 1, \dots, N$. The weak formulation of the PDE is given by: Find $y \in H_0^1(\Omega)$ :

$$\int_\Omega \nabla y^\nu \cdot \nabla v \, dx = \int_\Omega \left( \sum_{\nu=1}^{N} \chi_{B_\nu} u^\nu \right) \cdot v \, dx + \int_\Omega f \cdot v \, dx \quad \forall v \in H_0^1(\Omega).$$

We discretize the problem by a regular triangulation of $\Omega \subset \mathbb{R}^2$ using linear functions on each triangle. Let $N_I^h$ be the set of interior nodes in $\Omega$ and $N_T^h$ be the set of all nodes and further $\phi_i^h$, $i = 1, \dots, N_T^h$ be the nodal basis functions, i.e., the functions $\phi_i^h$ are 1 at node $i$ and 0 at all the other nodes. Then we define the matrices $A_h = \left( \int_\Omega \nabla \phi_i^h \cdot \nabla \phi_k^h \, dx \right)_{i,k \in N_I^h}$ and $M_h = \left( \int_\Omega \phi_i^h \cdot \phi_k^h \, dx \right)_{i,k \in N_T^h}$. Further, we compute

$$F_h^i \approx \int_\Omega f \cdot \phi_i^h \, dx \quad \forall i = 1, \dots, N_I^h$$

by a quadrature formula, which is exact for polynomials up to degree 2. Hence, our discretized version of the weak formulation of the PDE becomes a linear equation system with $N_I^h$ equations

$$A_h \, y_h^\nu - \sum_{\nu=1}^{N} (M_h)_{N_I^h \times N_\nu^h} \, u_h^\nu - F_h = 0,$$

where $N_\nu^h$ contains all nodes in $B_\nu$ for $\nu = 1, \dots, N$. Furthermore, the objective functions become quadratic and are given by

$$J_h^\nu(y_h^\nu, u_h^\nu) = \frac{1}{2}(y_h^\nu - z_h^\nu)^\top (M_h)_{N_I^h \times N_I^h} (y_h^\nu - z_h^\nu) + \frac{\alpha_\nu}{2} (u_h^\nu)^\top (M_h)_{N_\nu^h \times N_\nu^h} \, u_h^\nu,$$

where the $i$-th component of $z_h^\nu$ is $z^\nu$ evaluated at the node $i$.

Example 1 *This is the example from Borzi and Kanzow (2013). Here we have* $N = 2, \alpha_1 = \alpha_2 = 1$, $\Omega = (0,1) \times (0,1), l_1 = l_2 = -0.5, r_1 = r_2 = 0.5, \psi_l = -\infty, \psi_r = +\infty$. *Further,* $B_1 = ]0,1[ \times ]0, 0.5[\,, B_2 = ]0,1] \times ]0.5, 1[\,$, *and*

$$\begin{aligned} z^1(x) &= \sin(\pi x_1)\sin(\pi x_2) + 8\pi^2 \sin(2\pi x_1)\sin(2\pi x_2), \\ z^2(x) &= \sin(\pi x_1)\sin(\pi x_2) + 18\pi^2 \sin(3\pi x_1)\sin(3\pi x_2), \\ f(x) &= 2\pi^2 \sin(\pi x_1)\sin(\pi x_2) - \chi_{B_1} \max\{l_1, \min\{r_1, \sin(2\pi x_1)\sin(2\pi x_2)\}\} \\ &\quad -\chi_{B_2} \max\{l_2, \min\{r_2, \sin(3\pi x_1)\sin(3\pi x_2)\}\}. \end{aligned}$$

Table 1. Iteration numbers for Examples 1, 2 and 3

| | $h = 1/16$ | $h = 1/32$ | $h = 1/64$ | $h = 1/128$ |
|---|---|---|---|---|
| Example 1 | 21 | 22 | 25 | 25 |
| Example 2 | 26 | 31 | 30 | 33 |
| Example 3 | 39 | 47 | 50 | 53 |

Table 2. Error in the discrete $L^2(\Omega)$-norm for Example 1

| | $h = 1/16$ | $h = 1/32$ | $h = 1/64$ | $h = 1/128$ |
|---|---|---|---|---|
| $\lvert\bar{u}^1 - u_h^1\rvert_0$ | 0.0240 | 0.0080 | 0.0030 | 0.0010 |
| $\lvert\bar{u}^2 - u_h^2\rvert_0$ | 0.0258 | 0.0119 | 0.0059 | 0.0035 |
| $\lvert\bar{y} - y_h\rvert_0$ | $1.4e-3$ | $3.7e-4$ | $1.0e-4$ | $3.1e-5$ |

*The solution of this problem can be computed analytically, see Borzi and Kanzow (2013), and is given by*

$$\begin{aligned}
\bar{u}^1(x) &= \max\{l_1, \min\{r_1, \chi_{B_1} \sin(2\pi x_1)\sin(2\pi x_2)\}\},\\
\bar{u}^2(x) &= \max\{l_2, \min\{r_2, \chi_{B_2} \sin(3\pi x_1)\sin(3\pi x_2)\}\},\\
\bar{y}^1(x) &= \bar{y}^2(x) = \sin(\pi x_1)\sin(\pi x_2).
\end{aligned}$$

*Table 1 shows the iteration numbers for different values of $h$ and Table 2 summarizes the errors (measured in the discrete $L^2(\Omega)$-norm $|\cdot|_0$) between the computed discretized and the exact solution. As suggested in Borzi and Kanzow (2013) one can extend the existing results from Meyer and Rösch (2004), Rösch (2006), and Tröltzsch (2010) to obtain the convergence order $\mathcal{O}(h^2)$ for the state variable and $\mathcal{O}(h^{3/2})$ for the controls, if there are no constraints. However, we have active constraints for our controls and we observe in Table 2 a lower rate, in particular, for the control $u^2$. This might be improved by a more accurate solution of the discretized problem, which can be achieved by including the hybrid strategy suggested in Dreves et al. (2014). A detailed analysis of this is part of a future work. Note that also a finite differences discretization as in Borzi and Kanzow (2013) can be used, but the finite element method is more flexible when it comes to more complicated domains.*

Example 2 *This is the example from the preprint version of Hintermüller and Surowiec (2013). We have $N = 4, \Omega = (0,1) \times (0,1), \psi_r = +\infty, \alpha_\nu = 1.1 \cdot$*

$10^{-3}, l_\nu = 0$ *for all* $\nu = 1, \ldots, 4$. *Further, define*

$$B_1 = \left]0, \frac{1}{2}\right[ \times \left]0, \frac{1}{2}\right[, B_2 = \left]\frac{1}{2}, 1\right[ \times \left]0, \frac{1}{2}\right[,$$
$$B_3 = \left]0, \frac{1}{2}\right[ \times \left]\frac{1}{2}, 1\right[, B_4 = \left]\frac{1}{2}, 1\right[ \times \left]\frac{1}{2}, 1\right[$$

*and then*

$$\tilde{B}_1 = \left]\frac{1}{4}, \frac{1}{2}\right[ \times \left]\frac{1}{4}, \frac{1}{2}\right[, \tilde{B}_2 = \left]\frac{1}{2}, \frac{3}{4}\right[ \times \left]\frac{1}{4}, \frac{1}{2}\right[,$$
$$\tilde{B}_3 = \left]\frac{1}{4}, \frac{1}{2}\right[ \times \left]\frac{1}{2}, \frac{3}{4}\right[, \tilde{B}_4 = \left]\frac{1}{2}, \frac{3}{4}\right[ \times \left]\frac{1}{2}, \frac{3}{4}\right[.$$

*Let* $A := \tilde{B}_1 \cup \tilde{B}_2 \cup \tilde{B}_3 \cup \tilde{B}_4$ *and*

$$\begin{aligned}
r_\nu &= \chi_{B_\nu} - \frac{1}{2}\chi_{\tilde{B}_\nu}, \qquad \nu = 1, \ldots, 4, \\
z^\nu(x_1, x_2) &= 500 \min\{\max\{0, \psi(x_1, x_2)\}, 0.2\}, \qquad \nu = 1, \ldots, 4, \\
\psi_l(x_1, x_2) &= \cos\left(2\sqrt{(x_1 - 0.5)^2 + (x_2 - 0.5)^2}\right) - 0.7, \\
f(x_1, x_2) &= -\Delta(\psi)_+ - \frac{1}{2}\chi_A - \chi_{(\Omega \setminus A)}.
\end{aligned}$$

*The total number of iterations until convergence is reported in Table 1. Figure 1 shows the computed optimal controls, the optimal state and the optimal multiplier for the state constraints for player 1 for* $h = 1/64$. *Note that in some vertices of the domain we obtain* $u^\nu = 0$, *which is a result of the used triangulation. Despite of this, the discrete controls* $u^\nu_{64}, \nu = 1, \ldots, N$ *look very similar to the solution given in the preprint of Hintermüller and Surowiec (2013). With our unsophisticated implementation of Algorithm 1 we were not able to solve the problem with mesh size up to* $h = 1/512$ *as it was done in the preprint of Hintermüller and Surowiec (2013). However, exploiting the sparsity structure of the problem and using a computer with more memory this should also be possible with our approach. In contrast to the approach of Hintermüller and Surowiec (2013) we do not use a penalty term and therefore do not have to find an appropriate penalty parameter.*

EXAMPLE 3 *Here we have the data* $N = 2, \Omega = (0,1) \times (0,1), \alpha_\nu = 0.1, l_\nu = -1$, $r_\nu = 1$ *for* $\nu = 1, 2$. *Further, we have* $\psi_l = -\infty, \psi_r = 0$,

$$B_1 = ]0, 1[\times]0, 0.5[, \qquad B_2 = ]0, 1]\times]0.5, 1[,$$

$f \equiv 0$, *and*

$$z^1(x_1, x_2) = 10(\sin(2\pi x_1) + x_2), \qquad z^2(x_1, x_2) = 10(\sin(2\pi x_2) + 2x_1).$$

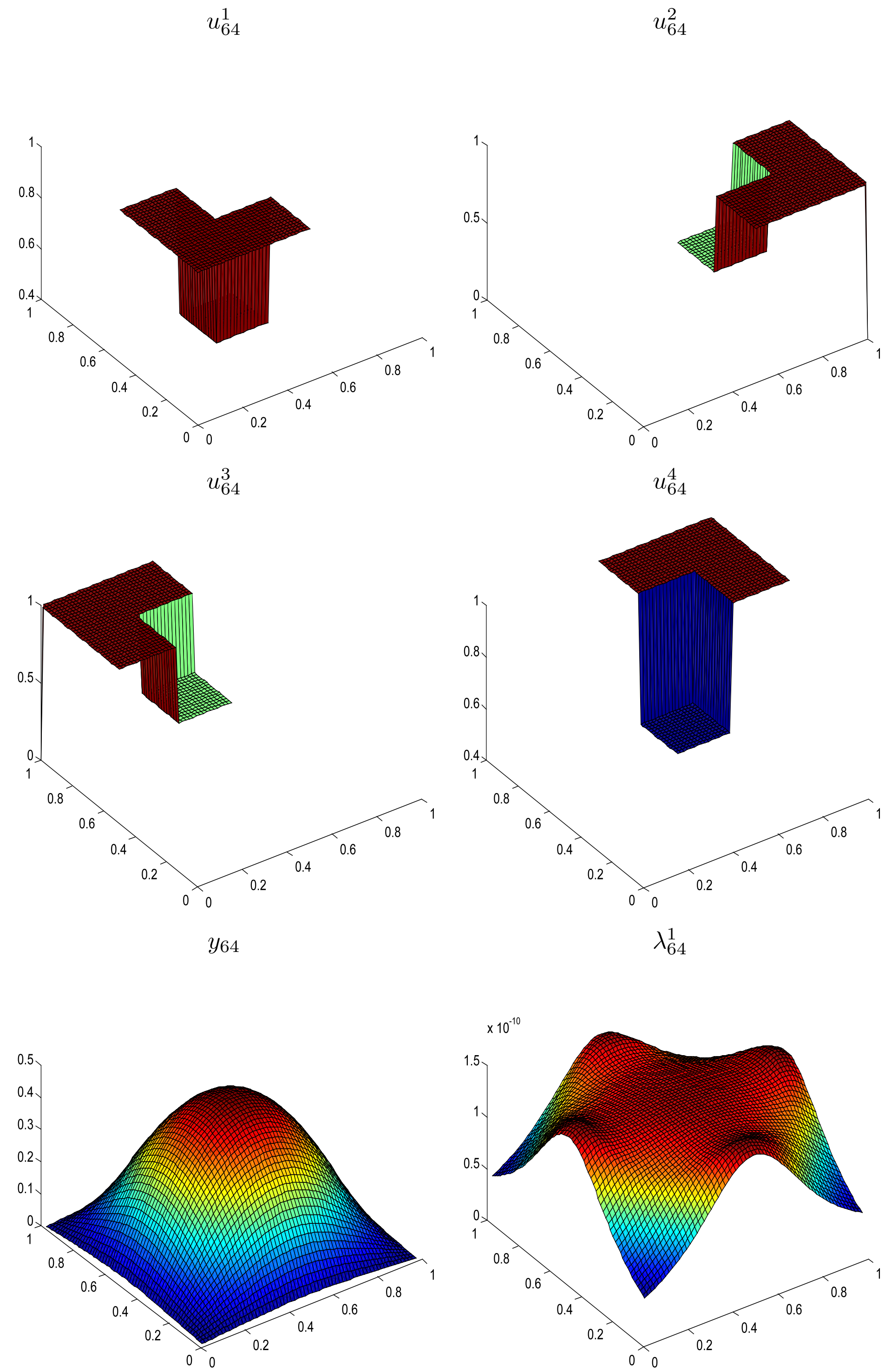


Figure 1. Computed optimal controls, state and multiplier for $h = 1/64$ for Example 2

*The total number of iterations until convergence is reported in Table 1. Figure 2 shows the computed optimal controls, the optimal state and the optimal multiplier for the state constraints for player 1 for $h = 1/64$. Here we can see that the multiplier enjoys only low regularity. As one might have expected, this results in some stronger mesh dependence of the algorithm than what was observed in Example 2, where the multiplier is more regular.*

# 7. Conclusion

We considered multiobjective optimal control problems governed by elliptic PDEs with pointwise constraints on the controls and on the state variables. Using the control-to-state mapping these problems are transformed into GNEPs. We proved convergence of a sequence of solutions of the discretized GNEPs to a solution of the infinite dimensional GNEP. Further, by introducing different state variables for each objective, one can also reformulate the problem as a GNEP. This was exploited in our numerical approach, where we proposed a potential reduction algorithm to find a generalized Nash equilibrium of the finite element discretization of the GNEP, and proved its convergence to a solution, where the introduced state variables are all equal. We presented some numerical results, showing the applicability of our method. In a future research project we are interested in a more efficient implementation that exploits the sparsity structure.

# Acknowledgments

Special thanks go to Professor Joachim Gwinner for the discussions and comments that helped to improve the paper a lot and to an anonymous referee.

# References

Arrow, K.J. and Debreu, G. (1954) Existence of an equilibrium for a competitive economy. *Econometrica* **22**, 265–290.

Berge, C. (1963) *Topological Spaces.* Oliver and Boyd, Edinburgh/London.

Borzi, A. and Kanzow, C. (2013) Formulation and numerical solution of Nash equilibrium multiobjective elliptic control problem. *SIAM J. Control Optim.* **51**, 718–744.

Cottle, R.W., Pang, J.-S. and Stone, R.E. (1992) *The Linear Complementarity Problem.* Academic Press, Boston.

Dreves, A., Facchinei, F., Fischer, A. and Herrich, M. (2014) A new error bound result for generalized Nash equilibrium problems and its algorithmic application. *Comput. Optim. Appl.* **59**, 63–84.

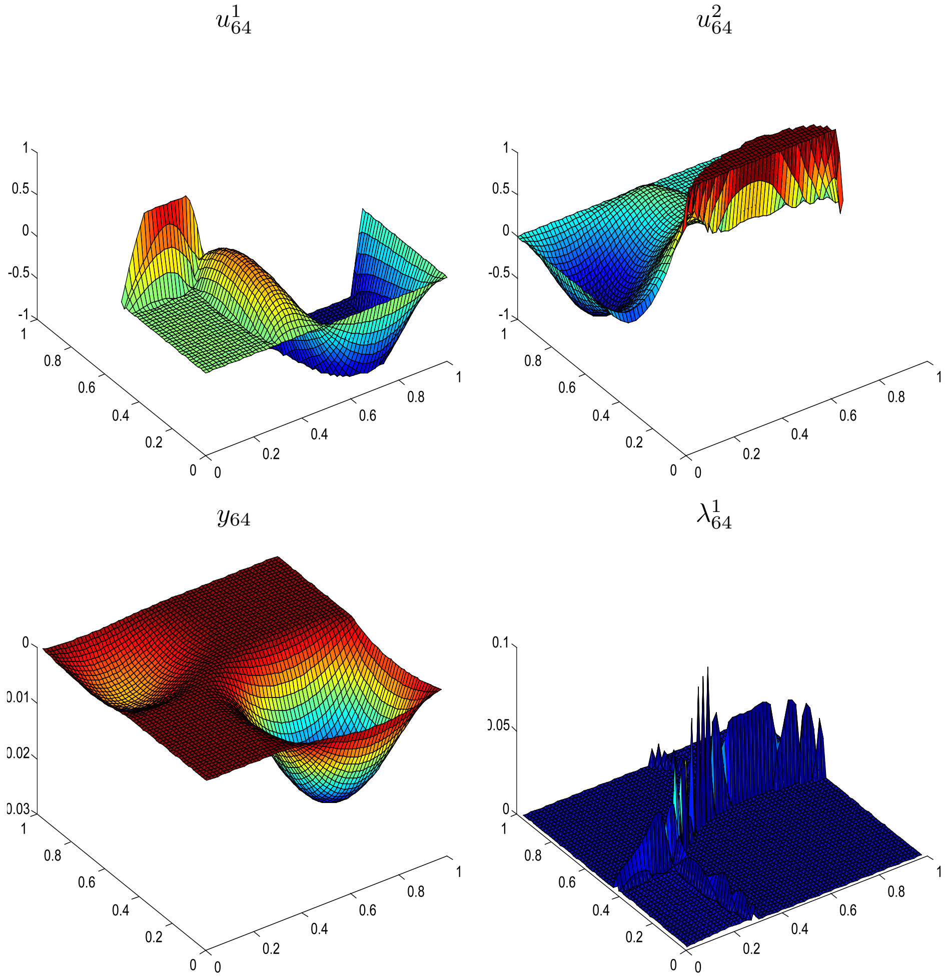


Figure 2. Computed optimal controls, state and multiplier for $h = 1/64$ for Example 3

DREVES, A., FACCHINEI, F. KANZOW, C. AND SAGRATELLA, S. (2011) On the solution of the KKT conditions of generalized Nash equilibrium problems. *SIAM J. Optim.* **21**, 1082–1108.

DREVES, A. AND GWINNER, J. (2016) Jointly convex generalized Nash equilibria and elliptic multiobjective optimal control. *J. Optim. Theory Appl.* **168**, 1065–1086.

FACCHINEI, F. AND KANZOW, C. (2007) Generalized Nash Equilibrium Problems. *4OR*, **5**, 173–210.

FACCHINEI, F., KANZOW, C. AND SAGRATELLA, S. (2014) Solving quasi-variational inequalities via their KKT-conditions. *Math. Program.* **144**, 369–412.

HASLINGER, J. AND NEITTAANMÄKI, P. (1996) *Finite Element Approximation for Optimal Shape, Material and Topology Design.* John Wiley & Sons, England, 2nd edition.

HINTERMÜLLER, M. AND SUROWIEC, T. (2013) A PDE-constrained generalized Nash equilibrium problem with pointwise control and state constraints. *Pac. J. Optim.* **9** (2), 251–273.

HINTERMÜLLER, M., SUROWIEC, T. AND KÄMMLER, A. (2015) Generalized Nash equilibrium problems in Banach spaces: Theory, Nikaido-Isoda-based path-following methods, and applications. *SIAM J. Optim.* **25** (3), 1826–1856.

HINZE, M., PINNAU, R., ULBRICH, M. AND ULBRICH, S. (2009) Optimization with PDE constraints. *Math. Model. Theory Appl.* **23**, Springer, New York.

HOGAN, W.W. (1973) Point-to-set maps in mathematical programming. *SIAM Review* **15**, 591–603.

ICHIISHI, T. (1983) *Game Theory for Economic Analysis.* Academic Press, New York.

LIONS, J.L. (1986) Contrôle de Pareto de systèmes distribués: Le cas d'évolution. *Comptes Rendus de L'Académie des Sciences, Serie I*, **302**, 413–417.

LIU, G.P., YANG, J.B. AND WHIDBORNE, J.F. (2001) *Multiobjective Optimisation and Control.* Research Studies Press LTD.

MEYER, C. AND RÖSCH, A. (2004) Superconvergence properties of optimal control problems. *SIAM J. Control Opt.* **43**, 970–985.

MONTEIRO, R.D.C. AND PANG, J.S. (1999) A potential reduction Newton method for constrained equations. *SIAM J. Optim.* **9**, 729–754.

RAMOS, A.M., GLOWINSKI, R. AND PERIAUX, J. (2002a) Nash equilibria for the multiobjective control of linear partial differential equations. *J. Optim. Theory Appl.* **112**, 457–498.

RAMOS, A.M., GLOWINSKI, R. AND PERIAUX, J. (2002b) Pointwise control of the Burgers equation and related Nash equilibrium problems: Computational approach. *J. Optim. Theory Appl.* **112**, 499–516.

RAMOS, A.M. AND ROUBICEK, T. (2007) Nash Equilibria in Noncooperative PredatorPrey Games. *Appl. Math. Optim.* **56**, 211–241.

Rösch, A. (2006) Error estimates for linear-quadratic control problems with control constraints. *Opt. Methods Softw.* **21**, 121–134.

Tröltzsch, F. (2010) *Optimal Control of partial differential equations.* Grad. Stud. Math. 112, AMS, Providence, RI. Translated from the 2005 German original by J. Sprekels.