ON NON-REDUCIBLE MULTI-PLAYER CONTROL PROBLEMS AND
THEIR NUMERICAL COMPUTATION

Veronika Karl† Frank Pörner‡

November 9, 2018

Abstract. In this article we consider a special class of Nash equilibrium problems that cannot be
reduced to a single player control problem. Problems of this type can be solved by a semi-smooth
Newton method. Applying results from the established convergence analysis we derive superlinear
convergence for the associated Newton method and the equivalent active-set method. We also provide
detailed finite element discretizations for both methods. Several numerical examples are presented to
support the theoretical findings.

AMS Subject Classification: 49M05, 49M15, 65K10, 65K15

Keywords: GNEP, semi-smoothness, Newton method, variational inequality

1. Introduction. We consider a Nash Equilibrium Problem (NEP) in the optimal control setting.
Here, \(N \in \mathbb{N}\) denotes the number of players. The strategy space of all players is given by \(U := \left\{ u^\nu \in L^2(\Omega) \right\}_\nu\). The player \(\nu \in \{1, \ldots, N\}\) is in control of the variable \(u^\nu \in \left\{ u^\nu \in L^2(\Omega) \right\}_\nu\). The strategies of
all players, except the \(\nu\)-th player are denoted by \(u^{-\nu} \in L^2(\Omega)^{N-1}\). Hence, we have the notation \(u := (u^\nu, u^{-\nu})\). Investigating multi-player control problems in the function space setting one usually assumes [6, 7, 12] that the players’ observation areas coincide. An exemplary problem setting for the \(\nu\)-th player’s problem is given by

\[
\min_{u^\nu \in L^2(\Omega)} \frac{1}{2} \|Su - y^\nu_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u^\nu\|_{L^2(\Omega)}^2
\]

s.t. \(u^\nu \in U^\nu_{\text{ad}}\).

where \(U^\nu_{\text{ad}} \subset L^2(\Omega)\) is a bounded convex set. In this setting existence and uniqueness of solutions are quite forward to show by exploiting standard arguments. Indeed the problem can be transformed into a convex single player control problem [7, Proposition 3.10]. However, the situation becomes considerably more complicated if the observation area of the tracking term differs for each player. To be more precise we consider \(\Omega^\nu \subset \Omega\) and assume that the \(\nu\)-th player aims at solving

\[
\min_{u^\nu \in L^2(\Omega)} \frac{1}{2} \|Su - y^\nu_d\|_{L^2(\Omega^\nu)}^2 + \frac{\alpha}{2} \|u^\nu\|_{L^2(\Omega)}^2
\]

s.t. \(u^\nu \in U^\nu_{\text{ad}}\). (1.1)

We will give the precise setting below in Section 2.1. Problems of this type lack of the possibility
to be reduced to a single control problem and therefore require a different treatment. To the best of
our knowledge, until now, there exists no theory regarding the uniqueness of solutions of this kind of
problems. By imposing an assumption on the regularization parameter \(\alpha > 0\) we will show existence
and uniqueness of solutions of the NEP (1.1) in Theorem 2.2. Solving NEPs is not only interesting
for solving the problem itself. Moreover, solving generalized Nash equilibrium problems (GNEPs) that
include inequality constraints like \(Su \leq \psi, \psi \in C(\bar{\Omega})\) require in certain solution methods the solution

*The first author was supported by the German Research Foundation (DFG) within the priority program "Non-smooth and
Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization" (SPP 1962) under grant num-
ber Wa 3626/3-1 and the second author was supported under Wa 3626/1-1.
†Universität Würzburg, Institut für Mathematik, Emil-Fischer-Str. 30, 97074 Würzburg, Germany;
 veronika.karl@mathematik.uni-wuerzburg.de
‡Universität Würzburg, Institut für Mathematik, Emil-Fischer-Str. 30, 97074 Würzburg, Germany;
 frank.poerner@mathematik.uni-wuerzburg.de

arXiv:1805.03009v2 [math.OC] 8 Nov 2018
of a sequence of NEPs [6, 12]. It is a natural approach to apply the semi-smooth Newton method in order to solve these multi-player control problems that are given by the following extension of (1.1)

$$\min_{u^v \in L^2(\Omega)} \frac{1}{2} \|Su - y_d^v\|^2_{L^2(\Omega)} + \frac{\alpha}{2} \|u^v\|^2_{L^2(\Omega)} + \frac{1}{2\rho} \|(\mu + \rho(\psi - \psi))/(\mu + \rho(\psi - \psi))\|^2_{L^2(\Omega)}$$

(1.2)

s.t. $u^v \in U^v_{ad}$

where $\mu \in L^2(\Omega)$ and $\rho > 0$ denotes a penalization parameter. Furthermore $(\cdot)_+ := \max(\cdot, 0)$ in a pointwise almost everywhere sense. Here, we will focus on the studies of the corresponding semi-smooth Newton method. Since the tracking term is again considered on $\Omega_v$ only, the method can be expected to converge superlinear only if $a$ is sufficiently large, see Theorem 4.4.

The outline of this paper is as follows: In Section 2.1 we introduce the reader to non-reducible NEPs and the extended augmented NEP. Here, our main results state existence and uniqueness of solutions, see e.g. [4]. Since

$$u \in L^2(\Omega), \quad y = \sum_{i=1}^N u_i \quad \text{in } \Omega,$$

(2.1)

$$y = 0 \quad \text{on } \partial \Omega.$$

We assume that the operator $A: Y \to H^{-1}(\Omega)$ is linear, bounded and continuously invertible. Note that this is the case for $A := -\Delta$, which is an isomorphism from $H^1_0(\Omega)$ to its dual $H^{-1}(\Omega)$. This can be proven using the Lax-Milgram theorem. In this setting it is convenient to take $Y := H^1_0(\Omega) \cap C(\bar{\Omega})$, see e.g. [4]. Since $u^v \in L^2(\Omega) \mapsto H^{-1}(\Omega)$ the state equation is well-posed. The corresponding solution operator satisfies

$$S: u \mapsto y = A^{-1} \sum_{i=1}^N u_i^v, \quad S: H^{-1}(\Omega)^N \to H^1_0(\Omega) \cap C(\bar{\Omega}).$$

Theorem 2.2 and Theorem 2.4. In Section 3 we collect results from the literature that are necessary for the extended augmented NEP. Here, our main results state existence and uniqueness of solutions, see e.g. [4]. Since

$$u \in L^2(\Omega), \quad y = \sum_{i=1}^N u_i \quad \text{in } \Omega,$$

(2.1)

$$y = 0 \quad \text{on } \partial \Omega.$$
Since for $n = 1$ we have the embedding $H^1_0(\Omega) \hookrightarrow C(\Omega)$, for $n = 2$ we have $H^1(\Omega) \hookrightarrow L^q(\Omega)$ with $1 \leq q < \infty$ and for $n = 3$ we still have $H^1(\Omega) \hookrightarrow L^6(\Omega)$, hence the required assumption on $S$ is satisfied. Due to the linearity of $A^{-1}$ we have

$$Su = \sum_{v=1}^{N} A^{-1}u^v := \sum_{v=1}^{N} S_v u^v, \quad S_v : H^{-1}(\Omega) \to Y \hookrightarrow L^2(\Omega), \quad S_v u^v := A^{-1}u^v.$$

Problems of this type can be reduced to a single convex control problem given by

$$\min_{u \in L^2(\Omega)} \frac{1}{2} \|Su\|_{L^2(\Omega)}^2 + \sum_{v=1}^{N} \left( -(u^v, S_v y^v_d) + \frac{\alpha}{2} \|u^v\|_{L^2(\Omega)}^2 \right)$$

s.t. $u \in U_{\text{ad}}$.

Here $U_{\text{ad}} := U_{\text{ad}}^1 \times \cdots \times U_{\text{ad}}^N$. This easily yields the existence of a unique equilibrium for $\alpha > 0$ [7, Proposition 3.10]. Let us now investigate the case if the tracking type functional for the $v$-th player is considered on a subset $\Omega_v \subseteq \Omega$ only. In this case reduction to a single control problem is not possible and we will refer to this type of problem as a non-reducible NEP. We consider the cost functional

$$f_v(u) := \frac{1}{2} \|Su - y^v_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u^v\|_{L^2(\Omega)}^2$$

and analyze the Nash equilibrium problem

$$\min_{u^v \in L^2(\Omega)} f_v(u) \quad \text{s.t.} \quad u^v \in U_{\text{ad}}^v, \quad (P_v)$$

Our aim is to study under which conditions $(P_v)$ admits a unique solution. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded Lipschitz domain and $\Omega_v \subseteq \Omega$ for $v = 1, ..., N$. For further use we define the characteristic function

$$\chi_v(x) : \Omega \to \mathbb{R}, \quad x \mapsto \begin{cases} 1 & \text{if } x \in \Omega_v, \\ 0 & \text{else}, \end{cases}$$

as well as $U := L^2(\Omega)^N$ and the operator

$$F(u) : U \to U, \quad F^v(u) := D_{u^v}f_v(u) = S_v \chi_v(Su - y^v_d) + \alpha u^v,$$

where $D_{u^v}$ denotes the partial Gâteaux derivative with respect $u^v$. Due to the convexity of the cost functional solutions of the NEP can be characterized via controls $\bar{u} \in U$ that solve the variational inequality

$$(F(\bar{u}), v - \bar{u})_U \geq 0, \quad \forall v \in U_{\text{ad}}$$

\[ \Leftrightarrow \sum_{v=1}^{N} (S_v \chi_v(S\bar{u} - y^v_d) + \alpha \bar{u}^v, v^v - \bar{u}^v) \geq 0, \quad \forall v^v \in U_{\text{ad}}^v \]

(2.3)

We will exploit this relation to prove uniqueness of solutions of problem $(P_1)$. It is well known that (2.3) can be equivalently formulated using the projection operator $P_{U_{\text{ad}}}$ onto the set $U_{\text{ad}}$. A solution $\bar{u} \in U_{\text{ad}}$ of $(P_1)$ can be characterized by the equation

$$0 = \bar{u} - P_{U_{\text{ad}}} (\bar{u} - \gamma F(\bar{u}))$$

(2.4)

for all $\gamma > 0$. Furthermore, this formulation allows us to tackle the problem using a semi-smooth Newton method.

### 2.2. Existence and Uniqueness of Solutions

If the variational inequality (2.3) is uniquely solvable the NEP $(P_v)$ admits a unique solution. It is well known that this is the case if $F$ is strongly monotone [14, Theorem 1.4]. The next theorem states that this is the case if the regularization parameter $\alpha$ is chosen large enough, depending on the sets $\Omega_v$. Let us define the set

$$Z := \bigcup_{v=1}^{N} \Omega_v$$

with associated characteristic function $\chi_Z$. In order to deal with the different sets $\Omega_v$ we need the following assumption.
Assumption 2.1. Assume that the regularization parameter $\alpha$ satisfies the inequality
\[
\alpha > \frac{1}{4} \sum_{i=1}^{N} ||\chi Z(S_v - \chi_v S_v)||_{L^2(\Omega) \to L^2(\Omega)}^2.
\] (2.5)
We will refer to the right hand side of (2.5) as the offset.

Note that for a fixed operator $S$ the offset depends only on the sets $\Omega_v$. Before we use this assumption to prove existence and uniqueness of solutions let us analyze this condition.

Let us assume that $S_v: H^{-1}(\Omega) \to H_0^1(\Omega) \cap C(\hat{\Omega})$. As already mentioned this is the case for the operator $S_v = (-\Delta)^{-\frac{1}{2}}$. It is well known that the solution operator $S_v$ is continuous. Hence, we know that the number
\[
C := \max_{\nu=1,...,N} \sup_{||w||_{L^2(\Omega)} = 1} ||S_v w||_{L^\infty(\Omega)} < \infty
\]
exists. Now we obtain
\[
||\chi Z(S_v - \chi_v S_v)||_{L^2(\Omega) \to L^2(\Omega)} = \sup_{||w||_{L^2(\Omega)} = 1} ||\chi Z(S_v - \chi_v S_v)w||_{L^2(\Omega)}
\]
\[
= \sup_{||w||_{L^2(\Omega)} = 1} \left( \int_Z (1 - \chi_v)^2 (S_v w)^2 \, dx \right)^{\frac{1}{2}}
\]
\[
\leq \sup_{||w||_{L^2(\Omega)} = 1} ||S_v w||_{L^\infty(\Omega)} \left( \int_Z (1 - \chi_v)^2 \, dx \right)^{\frac{1}{2}}
\]
\[
\leq C \sqrt{\text{meas}(Z \setminus \Omega_v)}.
\]

Hence,
\[
\sum_{i=1}^{N} ||\chi Z(S_v - \chi_v S_v)||_{L^2(\Omega) \to L^2(\Omega)}^2 \leq C^2 \sum_{i=1}^{N} \text{meas}(Z \setminus \Omega_v).
\]

Thus, we can interpret the offset from (2.5) as the maximum difference of the set $Z = \bigcup_v \Omega_v$ and the sets $\Omega_v$. If $\Omega_v = \Omega$ for all $v$ this offset is obviously zero and we are in the setting of a reducible NEP. However, if the offset is too large, the existence of minimizers can not be guaranteed by our theory for all $\alpha > 0$. Let us now start to exploit Assumption 2.1.

Theorem 2.2. Let Assumption 2.1 be satisfied. Then there exists an unique solution of the non-reducible NEP $(P)$. 

Proof. As already mentioned it is enough to show that the operator $F$ defined in (2.2) is strongly monotone. A calculation reveals for arbitrary $u, v \in U$
\[
(F(u) - F(v), u - v)_U = \sum_{i=1}^{N} \langle S_v^* \chi_v (S u - y_v) + \alpha u^v - S_v^* \chi_v (S v - y_v) - \alpha v^v, u^v - v^v \rangle_{L^2(\Omega)}
\]
\[
= \sum_{i=1}^{N} \langle (S u - S v) \chi_v S_v(u^v - v^v), u - v \rangle_{L^2(\Omega)} + \alpha \|u - v\|_U^2
\]
\[
= \sum_{i=1}^{N} \langle (S u - S v) \chi_v S_v(u^v - v^v), u - v \rangle_{L^2(\Omega)} + \alpha \|u - v\|_U^2.
\]

We now use the decomposition
\[
\sum_{i=1}^{N} \chi_1 S_v = \sum_{i=1}^{N} S_v - \sum_{i=1}^{N} (S_v - \chi_1 S_v)
\]
and Young’s inequality to obtain the following estimate

\[
\left( F(u) - F(v), u - v \right)_U = \| Su - Sv \|_{L^2(\Omega)}^2 - \left( Su - Sv, \sum_{v=1}^{N} (Sv - \chi_v Sv) (u^v - v^v) \right)_{L^2(\Omega)} + \alpha \| u - v \|_U^2
\geq - \frac{1}{4} \left[ \sum_{v=1}^{N} \| \chi Z (Sv - \chi_v Sv) (u^v - v^v) \|_{L^2(\Omega)}^2 \right] + \alpha \| u - v \|_U^2
\geq - \frac{1}{4} \left( \sum_{v=1}^{N} \| \chi Z (Sv - \chi_v Sv) \|_{L^2(\Omega)}^2 \right) \left( \sum_{v=1}^{N} \| u^v - v^v \|_{L^2(\Omega)}^2 \right) + \alpha \| u - v \|_U^2
= \left( \alpha - \frac{1}{4} \sum_{v=1}^{N} \| \chi Z (Sv - \chi_v Sv) \|_{L^2(\Omega)}^2 \right) \| u - v \|_U^2.
\]

Due to our assumption on \( \alpha \) we now conclude that the operator \( F \) is strongly monotone. \( \square \)

**Remark 1.** For getting existence of a solution of the non-reducible NEP \((P_\alpha)\) without requiring its uniqueness, it would be enough to claim monotonicity of \( F \) only, see [14, Theorem 1.4]. Thus, we can relax Assumption 2.1 slightly, by assuming that \( \alpha \) satisfies

\[
\alpha \geq \frac{1}{4} \sum_{v=1}^{N} \| \chi Z (Sv - \chi_v Sv) \|_{L^2(\Omega)}^2.
\]

The condition on the regularization parameter \( \alpha \) is needed to guarantee the existence of a unique solution of \((P_\alpha)\). If \( \alpha \) is chosen too small the resulting operator \( F \) might not be strongly monotone. It is quite interesting that for \( \alpha > 0 \) the operator \( F \) is still strongly monotone if all the domains \( \Omega_v \) coincide, but not necessarily equal to the domain \( \Omega \).

**Corollary 2.3.** If \( \Omega_v = \Omega_0 \subseteq \Omega \) for all \( v = 1, \ldots, N \), then the NEP \((P_\alpha)\) is uniquely solvable for all \( \alpha > 0 \).

### 2.3. The Augmented NEP

In this section we want to extend our result from the former section. Let \( \Omega \subseteq \mathbb{R}^n \) with \( n \in \{1, 2, 3\} \) be a bounded Lipschitz domain. Problems of this type are arise during the process of solving generalized Nash equilibrium problems (GNEPs) where the individual problem is given by

\[
\min_{u^v \in L^2(\Omega)} f_v(u^v) := \frac{1}{2} \| Su - y^v_d \|_{L^2(\Omega)}^2 + \alpha \| u^v \|_{L^2(\Omega)}^2
\text{ s.t. } u^v \in U^v_{ad}
\]

\[(\text{SNEP})\]

by applying an augmented Lagrange method, see [12]. Here, \( \psi \in C(\bar{\Omega}) \) defines an additional upper bound for the state \( y = S u \). To guarantee the existence of Lagrange multipliers it is necessary to have \( \psi \in C(\bar{\Omega}) \), see [13]. Solving \((\text{SNEP})\) with an augmented Lagrange method requires a sequence of solutions of the following Nash equilibrium problem, where \( \alpha, \rho > 0 \)

\[
\min_{u^v \in L^2(\Omega)} f^A_v(u, \mu, \rho) := \frac{1}{2} \| Su - y^v_d \|_{L^2(\Omega)}^2 + \alpha \| u^v \|_{L^2(\Omega)}^2 + \frac{1}{2\rho} \| \mu(u) \|_{L^2(\Omega)}^2
\text{ s.t. } u^v \in U^v_{ad}
\]

\[(\text{PAE})\]

with \( \mu(u) := (\mu + \rho (S u - \psi))_+ \) and \( \mu \geq 0 \) is assumed to be a function in \( L^2(\Omega) \). Defining

\[
F(u) := \left( D_{u^1} f^A_1(u, \mu, \rho), \ldots, D_{u^N} f^A_N(u, \mu, \rho) \right)
\]
it is again the convexity of the cost functional that allows us to characterize the solution of the NEP via controls $\bar{u} \in U$ that solve the variational inequality
\[
(F(\bar{u}), \nu - \bar{u})_U \geq 0, \quad \forall \nu \in U_{ad}
\]

\[
\iff \sum_{\nu=1}^N (S_{\nu}^e(\chi_{\nu}(\bar{u} - y_{ad}^\nu)) + \beta(u)) + \alpha \bar{u}^\nu, \nu^\nu - \bar{u}^\nu) \geq 0, \quad \forall \nu^\nu \in U_{ad}^\nu.
\]

Similar as above an equivalent formulation using the projection operator can be established, see (2.4). From Theorem 2.2 we know that the mapping
\[
u \mapsto \left( D_{\nu^1} f_1(u^1), \ldots, D_{\nu^N} f_N(u^N) \right)
\]
is strongly monotone if Assumption 2.1 is satisfied. Furthermore, we know that the function
\[
u \mapsto \frac{1}{2\alpha} \|\bar{\nu}(u)\|^2_{L^2(\Omega)}
\]
is convex and its derivative is monotone. Hence $F$ is strongly monotone and Theorem 2.2 can easily be adapted to that case.

**Theorem 2.4.** Let Assumption 2.1 be satisfied. Then there exists an unique solution of problem $(P_{\alpha}^{AL})$. Further, if $\Omega_{\nu} = \Omega_0 \subseteq \Omega$ for all $\nu = 1, \ldots, N$, then the NEP $(P_{\alpha}^{AL})$ is uniquely solvable for all $\alpha > 0$.

We will deepen our studies of problem $(P_{\alpha}^{AL})$ in Section 4. Here, we will among others derive the corresponding Newton iteration that allows us to solve the problem numerically with superlinear convergence.

### 3. Semi-smooth Newton Method.
This section aims at collecting important notations and results from literature in order to introduce the semi-smooth Newton method and state the well known theorem that yields superlinear convergence of just this method. We complete this section by contributing Lemma 3.5 that proves semi-smoothness of $u \mapsto \max(a, u)$ from $L^p(\Omega)$ to $L^p(\Omega)$ even if $a \in L^r(\Omega)$, with $1 \leq p \leq r < q \leq \infty$.

To simplify our notation we define
\[
y_q := L^q(\Omega)^N,
\]
for some $q \in [1, \infty)$. Recall that $U = L^2(\Omega)^N$, hence we have $Y_q \hookrightarrow U$ for $q \geq 2$. We want to apply Newton’s method to an equation similar to (2.4). Note that due to the regularization term we can always reformulate our necessary optimality condition to
\[
0 = \bar{u} - P_{U_{ad}}(\bar{u} - \gamma(\bar{F} + \alpha \bar{u}))
\]
with a function $\bar{F}$. From now on we always set $\gamma := \alpha^{-1} > 0$ to simplify our equation. Hence, we are interested in finding zeros of functions $G : U \rightarrow U$ defined as
\[
G(u) := u - P_{U_{ad}} \left( -\frac{1}{\alpha} \bar{F}(u) \right).
\]

We make the following assumption on $\bar{F}$ in order to be able to apply the semi-smooth Newton method, which is introduced in the next section, see Definition 3.1.

**Assumption 3.1.** We assume that $\bar{F}$ from (3.1) satisfies $\bar{F} : U \rightarrow Y_q$ with $q > 2$ such that each component $\bar{F}^\nu$ is semi-smooth and locally Lipschitz from $U$ to $L^q(\Omega)$ for all $u$.

### 3.1. The Semi-Smooth Newton Method.
Applying semi-smooth Newton methods requires the notion of semi-smoothness or Newton-differentiable functions. In this chapter let $U$ denote an arbitrary Banach space.
Definition 3.1 (Newton derivative). Let \( U, Y \) be Banach spaces. The mapping \( G : U \to Y \) is called Newton differentiable or semi-smooth if there exists a linear and continuous mapping \( D_N G : U \to \mathcal{L}(U, Y) \) such that
\[
\|G(u + v) - G(u) - D_N G(u + v)\|_Y = o(\|v\|_U)
\]
for every \( u \in U \). The mapping \( D_N G \) is called the Newton derivative of \( G \).

Let us present a well-known class of functions which are semi-smooth.

Lemma 3.2. Let \( U, Y \) denote Banach spaces. Every Fréchet differentiable function \( F : U \to Y \) with continuous Fréchet derivative \( F' \) is Newton differentiable with Newton derivative \( F' \).

The semi-smooth Newton method for finding a solution of \( G(u) = 0 \) is given in the following algorithm.

Algorithm 1 Semi-smooth Newton method

Choose \( u_0 \in U \)
For \( k = 0, 1, 2, \ldots \) repeat:
1: Compute \( \delta_k \) by solving \( D_N G(u_k)\delta_k = -G(u_k) \).
2: Set \( u_{k+1} := u_k + \delta_k \).

Let \( \bar{u} \) solve \( G(u) = 0 \), where \( G : U \to Y \). Let us assume that the mappings \( D_N G(u_k) \in \mathcal{L}(U, Y) \) are invertible. Applying the definition of a Newton step we get
\[
\|u_{k+1} - \bar{u}\|_U \leq \|D_N G(u_k)^{-1}\|_{Y \to U} \|D_N G(u_k)(u_k - \bar{u}) - (G(u_k) - G(\bar{u}))\|_Y.
\]

Based on this estimate it is well known and easy to see ([11, Theorem 8.16]), that Algorithm 1 converges superlinearly to a solution \( \bar{u} \) of \( G(u) = 0 \) if the following two conditions hold:

1. Approximation condition: The mapping \( G : U \to Y \) is Newton-differentiable at \( \bar{u} \) with Newton derivative \( D_N G : U \to \mathcal{L}(U, Y) \).
2. Regularity condition: There exists a constant \( c > 0 \) and an \( \varepsilon > 0 \) such that for every \( u \) satisfying \( \|u - \bar{u}\|_U \leq \varepsilon \) all \( D_N G(u) \) are invertible and \( \|D_N G(u)^{-1}\|_{Y \to U} \leq c \) holds.

To end this section we will recall some properties of Newton differentiable functions, see [8, Theorem 2.10].

Lemma 3.3. Let \( U, Y, Z, U_i, Y_i \) be Banach spaces.

a) If the operators \( G_i : U \to Y_i \) are Newton differentiable at \( u \) then \( (G_1, G_2) \) is Newton differentiable at \( u \).

b) If \( G_i : U \to Y_i, i = 1, 2 \) are Newton differentiable at \( u \) then \( G_1 + G_2 \) is Newton differentiable at \( u \) with Newton derivative \( D_N (G_1 + G_2) \).

c) Let \( G_i : Y \to Z \) and \( G_2 : U \to Y \) be Newton differentiable at \( G_2(u) \) and \( u \), respectively. Assume that \( D_N G_1 \) is bounded near \( y = G_2(u) \) and that \( G_2 \) is Lipschitz-continuous near \( u \). Then \( G = G_1 \circ G_2 \) is Newton differentiable with
\[
D_N G(u) = M_1 M_2 \quad \text{with} \quad M_1 := D_N G_1(G_2(u)), \ M_2 := D_N G_2(u).
\]

3.2. Semi-Smoothness of the Projection Operator. For our later application we will need semi-smoothness of the mapping
\[
u \mapsto (\mu + \rho(Su - \psi))^+,
\]
see Section 4.1. Since \( \mu \) is only a \( L^2(\Omega) \) function we cannot expect from the known result [20, Theorem 4.4] that the mapping
\[
\max(0, \mu + \rho(Su - \psi)) = \mu - \rho \psi + \max(-\mu + \rho \psi, Su)
\]
is semi-smooth from $L^q(\Omega)$ to $L^2(\Omega)$. In [11, Example 8.12] Ito and Kunisch investigated the semi-smoothness of superposition operators

$$F: L^q(\Omega) \to L^p(\Omega), \quad F(u)(x) = f(u(x)) \text{ for a.e. } x \in \Omega,$$

where $1 \leq p < q \leq \infty$ and $f: \mathbb{R} \to \mathbb{R}$ is semi-smooth and globally Lipschitz continuous. However, due to the dependence of $a$ and $b$ on the $x$-variable the mapping $u \mapsto \max(a, \min(u, b))$ cannot be built via superposition. Nevertheless, since the regularity of the functions $a$ and $b$ isn’t needed in the proof one can apply similar arguments.

**Theorem 3.4.** Let $a, b \in L^1(\Omega)$ with $a \leq b$ and $1 \leq r < q \leq \infty$. The mapping $m: L^q(\Omega) \to L^p(\Omega)$, $u \mapsto \max(a, \min(u, b))$ is semi-smooth with Newton derivative

$$L^q(\Omega) \ni h(u)(x) = \begin{cases} 0 & \text{if } u(x) \geq b(x), \\ 1 & \text{if } u(x) \in (a(x), b(x)), \\ 0 & \text{if } u(x) \leq a(x), \end{cases} \quad (3.3)$$

where $s$ is chosen such that $\frac{1}{p} = \frac{1}{q} + \frac{1}{q}$ holds.

**Proof.** A similar proof can be found in the PhD-Thesis [19]. Let $u \in L^q(\Omega)$ be arbitrary and $(s_k)_k \subset L^q(\Omega)$ be a (strong) nullsequence. Furthermore, define $u_k := u + s_k$ and $d_k := h(u_k)$. We have to check property (3.2).

First we extract a subsequence $(s_{k_l})_{l \in I}$ with an index set $I$ such that $s_{k_l}(x) \to 0$ for almost all $x \in \Omega$. To shorten the notation we furthermore define $v := m(u)$ and $v_k := m(u_k)$. We now use It is known [8, Example 2.5] that the mapping $\tilde{m}: \mathbb{R} \to \mathbb{R}, x \mapsto \max(a, \min(x, b))$ with $a, b \in \mathbb{R}$ is semi-smooth. Hence, we obtain

$$s_{k_l}(x)^{-1}(v_{k_l}(x) - v(x) - d_{k_l}(x)s_{k_l}(x)) \to 0$$

for almost all $x \in \Omega$. The quotient on the left side is understood to be zero whenever $s_{k_l}(x) = 0$. Now we use that the projection $m$ is nonexpansive and obtain

$$|v_{k_l}(x) - v(x) - d_{k_l}(x)s_{k_l}(x)| \leq |v_{k_l}(x) - v(x)| + |d_{k_l}(x)s_{k_l}(x)|$$

$$\leq |u(x) + s_{k_l}(x) - u(x)| + |s_{k_l}(x)|$$

$$\leq 2|s_{k_l}(x)|.$$

By applying Lebesgue’s dominated convergence theorem we obtain

$$s_{k_l}^{-1}(v_{k_l} - v - d_{k_l}s_{k_l}) \to 1$$

in $L^r(\Omega)$ for all $r \in [1, \infty)$. Hence, by applying Hölder’s inequality we get with $\frac{1}{p} = \frac{1}{s} + \frac{1}{q}$

$$\frac{\|v_{k_l} - v - d_{k_l}s_{k_l}\|_{L^p(\Omega)}}{\|s_{k_l}\|_{L^q(\Omega)}} \leq \|s_{k_l}^{-1}(v_{k_l} - v - d_{k_l}s_{k_l})\|_{L^r(\Omega)} \to 1. \quad (3.4)$$

Since this argumentation can be repeated for any subsequence of $(s_{k_l})_k$ the limit in (3.4) holds in fact for the whole sequence.

In the same manner we obtain the following result.

**Lemma 3.5.** Let $a \in L^1(\Omega)$ and $1 \leq p \leq r < q \leq \infty$. The mapping $m: L^q(\Omega) \to L^p(\Omega)$, $u \mapsto \max(a, u)$ is semi-smooth with Newton derivative

$$L^q(\Omega) \ni h(u)(x) = \begin{cases} 1 & \text{if } u(x) > a(x), \\ 0 & \text{if } u(x) \leq a(x), \end{cases} \quad (3.5)$$

where $s$ is chosen such that $\frac{1}{p} = \frac{1}{s} + \frac{1}{q}$ holds.
Note that the norm gap $p < q$ is indispensable for Newton differentiability of the projection operator, see for instance [11, Example 8.14]. Hence, the functions defined in (3.3) and (3.5) can in general not serve as a Newton derivative for $m : L^2(\Omega) \to L^2(\Omega)$, see [5, Proposition 4.1]. This causes trouble proving superlinear convergence since we cannot expect that the approximation condition holds. To bridge this norm gap, one needs additional structure. For problems that involve partial differential equations this structure is often given by smoothing properties of the corresponding solution operators. To finish, let us briefly comment on the semi-smoothness of the projection operator $P_{\text{Uad}} : L^q(\Omega)^N \to L^2(\Omega)^N$. The mapping

$$\Pi_v : L^q(\Omega)^N \to L^q(\Omega), \ u \mapsto u^v.$$ 

is linear and Fréchet differentiable, hence semi-smooth by Lemma 3.2. Applying the chain rule (Lemma 3.3 c)) we now obtain that

$$P_{\text{Uad}}^v(u) = \max(\Pi_v(u), b^v)$$

is a composition of semi-smooth functions, hence semi-smooth from $L^q(\Omega)^N \to L^2(\Omega)$, see Lemma 3.4. Using Lemma 3.3 a) we obtain that $P_{\text{Uad}}$ is semi-smooth from $L^q(\Omega)^N \to L^2(\Omega)^N$.

### 3.3. Convergence Analysis

To simplify our notation let us introduce the following notation. Let $d \in U$ with components $d^v \in L^2(\Omega)$ and $M \in \mathcal{L}(U, U)$. We define the product $d \cdot M = dM \in \mathcal{L}(U, U)$ in a component-wise manner

$$(d \cdot M(u))^v := d^v M^v(u) \in L^2(\Omega). \tag{3.6}$$

Hence, $d \cdot M : U \to U$. In a similar way we define $d \cdot u \in U$ for some $u \in U$. The chain rule from Lemma 3.3 c) allows us to show semi-smoothness of $G$ from (3.1).

**Lemma 3.6.** Let Assumption 3.1 be satisfied. Then the operator $G : U \to U$ is Newton differentiable with Newton derivative

$$D_N G(u) = \text{Id} + \frac{1}{\alpha} \chi_I(u) D_N F(u) \in \mathcal{L}(U, U),$$

where the components of $\chi_I(u)$ are given as

$$(\chi_I(u))^v(x) := \begin{cases} 0 & \text{if } -\frac{1}{\alpha} F(u)^v(x) \geq b^v(x), \\ 1 & \text{if } -\frac{1}{\alpha} F(u)^v(x) \in (a^v(x), b^v(x)), \\ 0 & \text{if } -\frac{1}{\alpha} F(u)^v(x) \leq a^v(x), \end{cases} \tag{3.7}$$

for almost all $x \in \Omega$.

**Proof.** This follows directly by Assumption 3.1, the chain rule (Lemma 3.3 c)) and Lemma 3.2. Further, the representation of the derivative follows immediately with the chain rule where the derivative of $P_{\text{Uad}}$ can be deduced from [8, Example 2.5].

Due to Assumption 3.1 the operator $G$ from (3.1) already satisfies the approximation condition. It remains to check on the regularity condition. For a given iterate $u_k$ let us define $\chi_{I_k} := \chi_I(u_k) \in U$. Let us now consider the bilinear form

$$a(w, v) := \left( \left( \text{Id} + \frac{1}{\alpha} \chi_{I_k} D_N F(u_k) \chi_{I_k} \right) w, v \right)_U. \tag{3.8}$$

We make the following assumption.

**Assumption 3.7.** Assume that the bilinear form (3.8) is coercive for all $u_k \in U$, i.e. there exists a constant $c > 0$ such that for all $w \in U$ it holds $a(w, w) \geq c \|w\|_U^2$.

In fact, Assumption 3.7 is obviously satisfied if $D_N F(u_k)$ is positive semidefinite with respect to the scalar product in $U$, i.e.,

$$(D_N F(u_k) w, w)_U \geq 0 \quad \forall w \in U$$

holds. Furthermore, it is well known [3, Proposition 4.1.6] that if $\tilde{F} : U \to Y_d$ is Gâteaux differentiable for all $u \in U$ and monotone, then the Gâteaux derivative $D\tilde{F}$ is positive semidefinite for every $u$. In
the next section we will explicitly show that the needed assumptions for superlinear convergence are satisfied for our NEP. Furthermore, please note that the structure of the first part of the bilinear form is very similar to the structure of the Newton derivative of $G$. However, the additional characteristic function available in the bilinear form allows us to prove superlinear convergence. This is part of the next theorem.

**Theorem 3.8.** Let $\bar{u}$ solve $G(u) = 0$ with $G$ as given in (3.1). Let Assumption 3.1 and 3.7 be satisfied. If $\|u_0 - \bar{u}\|_U$ is sufficiently small the iterates $u_k$ from Algorithm 1 converge superlinear to $\bar{u}$.

**Proof.** The proof uses standard arguments for semi-smooth Newton methods. For the readers convenience it can therefore be found in the appendix.

4. Newton Iteration for the Non-Reducible NEP. We now want to study the semi-smooth Newton method applied to problem $(P^\text{AL}_\nu)$.

4.1. Newton Iteration and Convergence Result. We aim at solving

$$G(u) := u - Pu_{ad} \left( -\frac{1}{\alpha} F(u) \right) = 0,$$  

(4.1)

where $\tilde{F} : U \to Y$ and the $\nu$-th component $\tilde{F}^\nu$ is given as the adjoint state

$$\tilde{F}^\nu(u) = p^\nu(u) := S^\nu \left( \chi_\nu(Su - y_\nu^*) + \bar{p}(u) \right).$$

**Lemma 4.1.** The operator $u \mapsto (S^\nu \left( \chi_\nu(Su - y_\nu^*) + \bar{p}(u) \right))_{\nu=1}^N$ satisfies Assumption 3.1.

**Proof.** We note that $S^\nu : H^{-1}(\Omega) \to (H^{-1})^* = H_0^1(\Omega)$. Due to embedding theorems we obtain that $p^\nu(u) = S^\nu \left( \chi_\nu (Su - y_\nu^*) + (\mu + \rho(Su - \psi))_+ \right)$ maps from $U$ to $L^q(\Omega)$ with some $q > 2$. Splitting the adjoint state in two parts

$$p^\nu(u) := S^\nu \chi_\nu (Su - y_\nu^*) + S^\nu \left( (\mu + \rho(Su - \psi))_+ \right),$$

we see clearly that the first part is continuously Fréchet differentiable, hence semi-smooth due to Lemma 3.2 and Lipschitz continuous from $U$ to $L^q(\Omega)$. Recall that we need the norm gap in order to prove semi-smoothness of the projection operator. For the second part, we know from Lemma 3.5 and the regularity conditions on $S$ that the mapping $u \mapsto \max(0, \mu + \rho(Su - \psi))$ is semi-smooth from $L^2(\Omega)^N$ to $L^2(\Omega)$. Further it is well known that it is Lipschitz continuous from $L^2(\Omega)^N$ to $L^2(\Omega)$. Since $S^\nu$ maps linear to $L^q(\Omega)$ we gain semi-smoothness and Lipschitz continuity of the whole second part.

Let us analyze the problem in more detail. The Newton-derivative $D_N G$ of $G$ at the point $u_k$ is of the form

$$D_N G(u_k) = \text{Id} + \frac{1}{\alpha} \chi_{\mathcal{Y}_k} D_N \tilde{F}(u_k).$$

In order to compute the derivative of $\tilde{F}^\nu(u_k)$ recall

$$\tilde{p}(u_k) = (\mu + \rho(Su_k - \psi))_+.$$

A Newton derivative of $\tilde{p}(u_k)$ is given by

$$D_N \tilde{p}(u_k) = \chi_{\mathcal{Y}_k} \rho S \quad \text{with} \quad \mathcal{Y}_k := \{ x \in \Omega : (\mu + \rho(Su_k - \psi))(x) > 0 \}.$$

Hence, we obtain that the $\nu$-th component of the Newton derivative $D_N \tilde{F}(u_k)$ is given by

$$D_N \tilde{F}(u_k)^\nu = S^\nu \left( \chi_\nu S + \rho \chi_{\mathcal{Y}_k} S \right).$$

We end up with the following theorem.

**Theorem 4.2.** Let $G$ be given as in (4.1). A suitable Newton derivative of $G$ at $u_k$ in direction $h \in U$ is given by

$$(D_N G(u_k) h)^\nu = h_\nu + \frac{1}{\alpha} \chi_{\mathcal{Y}_k} \left( S^\nu \left( \chi_\nu S + \chi_{\mathcal{Y}_k} \rho S \right) h \right).$$
Let us analyze this problem in more detail. In particular we want to provide a finite element discretization. Let us recall the sets $I^v_k$ and $J^v_k$ and also define the sets $A^{v,a}_k$ and $A^{v,b}_k$

$$A^{v,a}_k := \left\{ x \in \Omega : -\frac{1}{\alpha} F^v(u_k) \leq u^v_a \right\}, \quad A^{v,b}_k := \left\{ x \in \Omega : -\frac{1}{\alpha} F^v(u_k) \geq u^v_b \right\},$$

$$I^v_k := \left\{ x \in \Omega : -\frac{1}{\alpha} F^v(u_k) \in (u^v_a, u^v_b) \right\}, \quad J^v_k := \left\{ x \in \Omega : (\mu + \rho(Su_k - \psi)) > 0 \right\}. \quad (4.2)$$

Following the lines of Theorem 3.8 we obtain that the following equality holds

$$u^{v}_{k+1}(x) = \begin{cases} 
  u^v_a(x) & \text{if } x \in A^{v,a}_k, \\
  -\frac{1}{\alpha} S^v(\chi S(u_k - y^\nu_d) + \chi \lambda \nu (\mu + \rho(Su_k - \psi))) (x) & \text{if } x \in I^v_k, \\
  u^v_b(x) & \text{if } x \in A^{v,b}_k.
\end{cases}$$

Thus, on the set $I^v_k$ we obtain

$$\chi I^v_k \left(u^v_{k+1} + \frac{1}{\alpha} (S^v(\chi S + \chi \lambda \nu \rho S)u^v_k)\right) = 0. \quad (4.3)$$

Let us introduce the function $u^{v,T}_{k+1} \in U$ with components $u^{v,T}_{k+1} := \chi I^v_k u^v_k$ for $\nu = 1, ..., N$. Hence, we can write $u^{v,T}_{k+1} = \chi A^v_k u_{k+1}$ with $\chi A^v_k := \chi I^v_k$ defined in (3.7). In a similar way we define $\chi A^a_k$ and $\chi A^b_k$. Using this definitions we can now write (4.3) as a linear equation for the $v$-th component of $u^{v,T}_{k+1}$ and we obtain

$$u^{v,T}_{k+1} + \frac{1}{\alpha} \chi I^v_k \left(S^v(\chi S + \chi \lambda \nu \rho S)u^v_k\right) = -\frac{1}{\alpha} \chi I^v_k \left(S^v((\chi S + \chi \lambda \nu \rho S)(\chi A^a_k u_a + \chi A^b_k u_b) - \chi y^\nu_d + \chi \lambda \nu (\mu - \rho \psi))\right).$$

The Newton step can now be written in the following compact form.

**Lemma 4.3.** The solution $u_{k+1}$ of one step of the semi-smooth Newton method is given by

$$u_{k+1} = u^{T}_{k+1} + \chi A^a_k u_a + \chi A^b_k u_b,$$

where $u^{T}_{k+1}$ is given as the solution of the linear system

$$\left(\text{Id} + \chi A_k T^k\right) u^{T}_{k+1} = \chi g_k, \quad (4.4)$$

with the operator $T^k : U \to U$ and function $g_k \in U$ given by

$$(T^k h)^v := \frac{1}{\alpha} S^v((\chi S + \chi \lambda \nu \rho S)h),$$

$$(g_k)^v := -\frac{1}{\alpha} \chi I^v_k \left(S^v((\chi S + \chi \lambda \nu \rho S)(\chi A^a_k u_a + \chi A^b_k u_b) - \chi y^\nu_d + \chi \lambda \nu (\mu - \rho \psi))\right).$$

Here $\text{Id} : U \to U$ denotes the identity mapping.

The complete semi-smooth Newton method is given in the following algorithm.

**Algorithm 2** Semi-smooth Newton method for problem ($P^A_v$)

1: Set $k = 0$, choose $u_0$ in $L^2(\Omega)^N$
2: repeat
3: Set $A^{v,a}_k$, $A^{v,b}_k$, $T^v_k$ and $J^v_k$ as defined in (4.2)
4: Solve for $u^{v,T}_{k+1} \in L^2(\Omega)^N$ by solving (4.4)
5: Set $u_{k+1} := u^{v,T}_{k+1} + \chi A^a_k u_a + \chi A^b_k u_b$
6: Set $k := k + 1$
7: until $A^{v,a}_k = A^{v,a}_{k-1}$, $A^{v,b}_k = A^{v,b}_{k-1}$, $T^v_k = T^v_{k-1}$ and $J^v_k = J^v_{k-1}$. 

NON-REDUCIBLE MULTI-PLAYER CONTROL PROBLEMS
Theorem 4.4 (Convergence of the semi-smooth Newton method). Let Assumption 2.1 hold and let \( \bar{u} \) denote the solution of \((P^{AL}_v)\). Then the semi-smooth Newton method from Algorithm 2 has the following properties

a) Let \( \|u_0 - \bar{u}\|_{L^2(\Omega)}^N \) be sufficiently small. Then the iterates \( u_k \) converge for \( k \to \infty \) superlinearly to \( \bar{u} \) which is the solution of \((P^{AL}_v)\).

b) Let \( u_k \) be generated by Algorithm 2 such that the stopping criterion from step 7 is satisfied. Then \( u_k \) is a solution of \((4.1)\).

Proof. a) Due to Theorem 3.8 it remains to check if Assumption 3.7 is satisfied. Thus, we consider the bilinear form

\[ a(w, v) := \left( \left( \text{Id} + \frac{1}{\alpha} \chi_{I_k} M \chi_{I_k} \right) w, v \right)_U, \]

where \((Mw)\nu = S^\nu(S + \rho \chi_{k^1})w\). Using the decomposition from the proof of Theorem 2.2 we obtain

\[
\begin{align*}
&\left( \left( \text{Id} + \frac{1}{\alpha} \chi_{I_k} M \chi_{I_k} \right) w, v \right)_U = \sum_{v=1}^{N} \left( w^v + \frac{1}{\alpha} \chi_{I_k} S^v(S + \rho \chi_{k^1}) \chi_{I_k} w, w^v \right)_{L^2(\Omega)} \\
&= \|w\|_{U}^2 + \frac{1}{\alpha} \sum_{v=1}^{N} \left( \chi_{I_k} S^v w, \chi_{I_k} S^v w \right)_{L^2(\Omega)} + \frac{2}{\alpha} \sum_{v=1}^{N} \left( \chi_{I_k} S^v w, \chi_{I_k} S^v w \right)_{L^2(\Omega)} \\
&\geq \|w\|_{U}^2 + \frac{1}{\alpha} \left( \sum_{v=1}^{N} \|S^v \chi_{I_k} w\|_{L^2(\Omega)}^2 \right) - \frac{1}{\alpha} \left( \sum_{v=1}^{N} \left( S^v \chi_{I_k} w, \sum_{v=1}^{N} (S^v - \chi_{k^1}) \chi_{I_k} w^v \right)_{L^2(\Omega)} \right) \\
&\geq \left( 1 - \frac{1}{\alpha} \sum_{v=1}^{N} \|S^v \chi_{I_k} w\|_{L^2(\Omega)} \right) \|w\|_{U}^2.
\end{align*}
\]

Choosing \( \alpha \) as in Assumption 2.1 we get the desired result from Theorem 3.8.

b) We know that the solution of \((4.4)\) is unique for fixed sets \( A_k^{\nu,a}, A_k^{\nu,b}, T_k^\nu \) and \( Y_k \).

We set \( A_k^{\nu,a} := A_k^{\nu,a}, A_k^{\nu,b} := A_k^{\nu,b}, T_k^\nu := T_{k+1}^\nu \) and \( Y_k := Y_{k+1} \) in \((4.4)\) and get

\[
\begin{align*}
&u_{k+1}^{\nu,T} + \frac{\chi_{I_k+1,v}}{\alpha} \left( S^v(S + \rho \chi_{k^1}) - \chi_{I_k+1} \right) u_{k+1}^{\nu,T} = \\
&- \frac{\chi_{I_k+1,v}}{\alpha} \left( S^v(S + \rho \chi_{k^1}) - \chi_{I_k+1} \right) u_{k+1}^{\nu,T} = 0.
\end{align*}
\]

Together with \( u_{k+1} = u_a \) on \( A_k^{\nu} \) and \( u_{k+1} = u_b \) on \( A_k^{\nu} \) we get

\[
u_k^{\nu} - P_{\nu,k+1} \left( - \frac{1}{\alpha} p^\nu(u_{k+1}) \right) = 0.
\]

Hence, \( u_{k+1} \) is a solution of \((4.1)\).
Again, we can drop the assumption on $\alpha$ if the sets $\Omega_v$ coincide.

**Corollary 4.5.** Let $\Omega_v = \Omega_0 \subseteq \Omega$ for all $v = 1, \ldots, N$ and let $\alpha > 0$. Then the associated Newton method to the NEP $(P^A_v)$ converges superlinear.

Considering the non-reducible NEP $(P_v)$ just results in setting $M := S^*_v \chi_v S$ instead of $S^*_v (\chi_v S + \rho \chi \chi_S)$. Hence, the convergence result from Theorem 4.4 transfers one by one to this kind of problem.

### 4.2. Implementation.

Let us now focus on the details of an implementation using finite elements. To illustrate the implementation we focus on problem $(P^r)$ where $S$ denotes the solution operator of (2.1) with $A := -\Delta$. Using standard methods the corresponding optimality system is given by

\[-\Delta \bar{g} = a^v \quad \text{in } \Omega \]

\[-\Delta \bar{p}^v = \chi_v (\bar{y} - y^d) + \bar{\mu} \quad \text{in } \Omega \] (4.5a)

\[(\bar{p}^v + \alpha \bar{a}^v, v^v - a^v) \geq 0 \quad \forall v^v \in U^v_{\text{adh}} \] (4.5b)

\[\bar{\mu} = (\mu + \rho (\bar{y} - \psi))_+ \] (4.5c)

where the state and adjoint equation satisfy suitable boundary conditions. We are interested in a finite element discretization; so let us define the finite dimensional space $\mathcal{V}_h := \text{span} \{ \phi_1, \ldots, \phi_m \}$. The index $h$ indicates the underlying discretization and the functions $\phi_j$ denote the basis functions. Let us now consider a discretized version of (4.5). We define the bilinear form

\[a(w, v) := \int_{\Omega} \nabla w \nabla v \, dx.\]

Then, the discretized version of (4.5) is given by the solution $(y_h, u_h, p_h)$ of the system

\[a(y_h, v_h) = \left( \sum_{v=1}^{N} u^v_h, v_h \right) \quad \forall v_h \in \mathcal{V}_h, \]

\[a(p^v_h, v_h) = (\chi_v (y_h - y^d) + (\mu + \rho(y_h - \psi))_+, v_h) \quad \forall v_h \in \mathcal{V}_h, \]

\[u^v_h = P_{[u^v_h, u^v_h]} \left( -\frac{1}{\alpha} p^v_h \right). \] (4.6)

Since for a given $u_h$ there exists an unique $y_h(u_h)$ and an unique adjoint states $p^v_h(u_h)$, system (4.6) can be reduced to the single equation

\[u^v_h = P_{[u^v_h, u^v_h]} \left( -\frac{1}{\alpha} p^v_h(u_h) \right) \quad \forall v_h \in \mathcal{V}_h. \]

Again we define the active and inactive sets for the discrete function $u_{k,h}$:

\[A^v_{k,h} := \left\{ x \in \Omega : \frac{1}{\alpha} p^v_h(u_{k,h}) \leq u^v_h \right\}, \quad A^{u,v}_{k,h} := \left\{ x \in \Omega : \frac{1}{\alpha} p^v_h(u_{k,h}) \geq u^v_h \right\}, \]

\[I^v_{k} := \left\{ x \in \Omega : -\frac{1}{\alpha} p^v_h(u_{k,h}) \in (u^v_h, u^v_h) \right\}, \quad Y^v_{k} := \left\{ x \in \Omega : (\mu + \rho (S_{u_{k,h}} - \psi)) > 0 \right\}. \]

We now define the functions $u^I_{k+1,h} := \chi_{I^v_k} u_{k+1,h}$, where $\chi_{I^v_k} := \chi_{I^v_k}(u_k)$. Following the lines of the proof of Section 4.1 we can establish a linear equation for the components of $u^I_{k+1,h}$:

\[u^I_{k+1,h} + \frac{1}{\alpha} \chi_{I^v_k} \left( S^*_v (\chi_v S + \chi_S \rho S) u^I_{k+1,h} \right) \]

\[= \frac{1}{\alpha} \chi_{I^v_k} \left( S^*_v (\chi_v S + \chi_S \rho S) (\chi_{A^v_{k,h}} u_a + \chi_{A^u,v_{k,h}} u_b) - \chi_v y^d + \chi_{Y^v_{k}} (\mu + \rho \psi) \right). \]

We want to solve this system by testing it with a function $v_h \in \mathcal{V}_h$. Note that $v^I_{k+1,h} \not\in \mathcal{V}_h$ in general, but it can be calculated as a projection $u^I_{k+1,h} = \chi_{I^v_k} \bar{u}^I_{k+1,h}$ of a function $\bar{u}^I_{k+1,h} \in \mathcal{V}_h$, see (4.6).
In the following denote \( u_h \in \mathbb{R}^m \) the coefficient vector of a function \( u_h \in \mathcal{V}_h \), where \( m \) denotes the dimension of the space \( \mathcal{V}_h \).

Furthermore, we assume that \( u^e_h, u^a_h \in \mathcal{V}_h \). We can reformulate the Newton step as a linear system in the coefficient vectors \( \tilde{u}_{k+1,h}^e \).

**Lemma 4.6.** The coefficient vectors \( \tilde{u}_{k+1,h}^e \) for \( 1 \leq v \leq N \) satisfy the linear system

\[
\begin{pmatrix}
E_{1,1} & E_{1,2} & \ldots & \ldots & E_{1,N} \\
E_{2,1} & E_{2,2} & \ldots & \ldots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
E_{N,1} & \ldots & \ldots & E_{N,N-1} & E_{N,N}
\end{pmatrix}
\begin{pmatrix}
\tilde{u}_{k+1,h}^e \\
\tilde{u}_{k+1,h}^a
\end{pmatrix}
= \begin{pmatrix}
C_1 \\
C_2 \\
\vdots \\
C_{N-1} \\
C_N
\end{pmatrix},
\]

where

\[
E_{i,j} := \begin{cases}
M_{T_k}^{ij} + \frac{1}{h} M_{T_k}^{ij} K^{-1} M_{T_k}^{1,k} + \frac{\mu}{h} M_{T_k}^{ij} K^{-1} M_{T_k}^{b,k} - \mu_k \frac{1}{h} M_{T_k}^{ij} K^{-1} M_{T_k}^{b,k} & \text{if } i = j \\
\frac{1}{h} M_{T_k}^{ij} K^{-1} M_{T_k}^{1,k} + \frac{\mu}{h} M_{T_k}^{ij} K^{-1} M_{T_k}^{b,k} & \text{else}
\end{cases}
\in \mathbb{R}^{m \times m},
\]

as well as

\[
C_i := -\frac{1}{\mu} M_{T_i}^{1,k} K^{-1} \left[ M_{T_i}^{ij} (\mu - \rho \nabla) - M_i \rho \right] - \mu_k \frac{1}{h} M_{T_k}^{ij} K^{-1} \sum_{i=1}^{N} \left( M_{A_k^{ij}} u^e_h + M_{A_k^{ij}} u^a_h \right) \in \mathbb{R}^m,
\]

and matrices \( K, M, M_{T_k}, M_{A_k^{ij}}, M_{A_k^{ij}} \) and \( \mu_k \) of the size \( \mathbb{R}^{m \times m} \) with

\[
K_{ij} := \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j, \quad (M_i)_{ij} := \int_{\Omega} \phi_i \phi_j,
\]

\[
(M_{T_k})_{ij} := \int_{T_k} \phi_i \phi_j, \quad (M_{A_k^{ij}})_{ij} := \int_{T_k} A^{ij} \phi_i \phi_j,
\]

\[
(M_{A_k^{ij}})_{ij} := \int_{T_k} \phi_i \phi_j, \quad (M_{\mu_k})_{ij} := \int_{\Omega} \phi_i \phi_j,
\]

where \( \phi_i, \phi_j \) denote the finite element basis functions of \( \mathcal{V}_h \).

We can reconstruct the state and the adjoint states using the coefficient vectors \( \tilde{u}_{k+1,h}^e \).

**Corollary 4.7.** The coefficient vector of the state \( y_{k+1,h} \) satisfies

\[
y_{k+1,h} = K^{-1} \sum_{v=1}^{N} \left( M_{T_k}^{1,k} \tilde{u}_{k+1,h}^e + M_{A_k^{ij}} u^e_h + M_{A_k^{ij}} u^a_h \right)
\]

and the coefficient vector of the adjoint state \( p_{k+1,h}^v \) can be computed by

\[
p_{k+1,h}^v = K^{-1} \left( M_v (y_{k+1,h} - y^v) + \mu \rho (y_{k+1,h} - y^v) \right).
\]

The control \( u^e_{k+1,h} \) can be computed by

\[
u_{k+1,h}^e = \chi_{T_k} u_{k+1,h}^e + \chi_{A_k^{ij}} u^e_h + \chi_{A_k^{ij}} u^a_h.
\]

We only need the adjoint states to update our active sets, hence kinks and discontinuities in the control will not be accumulated during the algorithm. This is an advantage over the discrete version of the active-set method. However, the expressions arising in the Newton method are more complicated than the expressions in the active-set method.
5. Active-Set Method. In this section we want to introduce an active-set method which is equivalent to the semi-smooth Newton method. For additional information regarding active-set methods, we want to refer to [2, 5, 9, 10, 18] and the references therein.

5.1. Equivalence to Active-Set Method. Let us establish the relation between the semi-smooth Newton method and the active-set method. For the sake of simplicity we consider the setting from Section 4.2. We consider the problem’s first-order optimality conditions (4.5). Reformulating (4.5c) by applying the projection formula one has to solve systems of this type in the active-set method which is defined below.

**Algorithm 3** Active-set method for problem \((P_{\nu}^{AL})\)

1: Set \(k = 0\), choose \((y_0, u_0, p_0) \in Y \times L^2(\Omega)^N \times L^2(\Omega)^N\)

2: repeat

3: Set \(A_k^{v,a}, A_k^{v,b}, T_k^a\) and \(\mathcal{Y}_k\) as defined in (4.2)

4: Solve for \((y_{k+1}, u_{k+1}, p_{k+1}) \in Y \times L^2(\Omega)^N \times L^2(\Omega)^N\) by solving

\[
-\Delta y_{k+1} = \sum_{i=1}^{N} u_{k+1}^i \text{ in } \Omega, \quad (5.1a)
\]

\[
-\Delta p_{k+1}^u = \chi_y (y_{k+1} - y_d^u) + \chi_y (\mu + \rho (y_{k+1} - \psi)) \text{ in } \Omega, \quad (5.1b)
\]

\[
u_{k+1}^i + \chi_{T_k^a} \left( \frac{1}{\alpha} p_{k+1}^v \right) = \chi_{A_k^{v,a}} u_{k+1}^i + \chi_{A_k^{v,b}} u_{k+1}^i, \quad (5.1c)
\]

5: Set \(k = k+1\)

6: until \(A_k^{v,b} = A_{k-1}^{v,b}, A_k^{v,a} = A_{k-1}^{v,a}, T_k^a = T_{k-1}^a\) and \(\mathcal{Y}_k = \mathcal{Y}_{k-1}\)

Here, we assume that the state and the adjoint equation satisfy suitable boundary conditions. Considering the active-set method from Algorithm 3, equation (5.1b) yields the identity

\[
\rho_{k+1}^v = S^v \left( \chi_y (y_{k+1} - y_d^u) + \chi_y (\mu + \rho (y_{k+1} - \psi)) \right).
\]

Inserting this identity in (5.1c) and exploiting \(y_{k+1} = Su_{k+1}\) from (5.1a) we get

\[
u_{k+1}^i + \frac{\chi_{T_k^a}}{\alpha} (S^v (\chi_y S + \chi_y \rho S) u_{k+1}) = \chi_{A_k^{v,a}} u_{k+1}^i + \chi_{A_k^{v,b}} u_{k+1}^i + \frac{\chi_{T_k^a}}{\alpha} (S^v (\chi_y y_d^v + \chi_y (-\mu + \rho \psi))).
\]

Using the representation \(u_{k+1} = \chi_{A_k^{v,a}} u_{k+1} + \chi_{A_k^{v,b}} u_{k+1} + \chi_{T_k^a} u_{k+1}\) we get \(u_{k+1} = u_0\) on \(A_k^{v,a}\) and \(u_{k+1} = u_b\) on \(A_k^{v,b}\). On the set \(T_k^a\) we have for all \(v = 1, \ldots, N\)

\[
\chi_{T_k^a} u_{k+1}^v + \frac{\chi_{T_k^a}}{\alpha} (S^v (\chi_y S + \chi_y \rho S) u_{k+1}) = \frac{\chi_{T_k^a}}{\alpha} (S^v (\chi_y y_d^v + \chi_y (-\mu + \rho \psi)))
\]

\[
\chi_{T_k^a} u_{k+1}^v + \frac{\chi_{T_k^a}}{\alpha} (S^v (\chi_y S + \chi_y \rho S) u_{k+1})
\]

which coincides with a Newton step from (4.4). In a similar way we can start with the Newton method and derive the active-set method. Hence, both methods are equivalent.

5.2. Implementation. Let us also present a numerical implementation of the active-set method. To formulate the active-set method we need to introduce the matrix \(M \in \mathbb{R}^{m \times m}, (M)_{ij} := \left[ \int_{\Omega} \phi_i \phi_j \right]_{ij} \).

**Lemma 5.1.** One step of the active-set method from Algorithm 3 can be computed by solving the system

\[
\begin{pmatrix}
K & E_1 & 0 \\
E_2 & 0 & E_3 \\
0 & E_4 & E_5
\end{pmatrix}
\begin{pmatrix}
y \\
u \\
p
\end{pmatrix}
=
\begin{pmatrix}
0 \\
l_1 \\
l_2
\end{pmatrix}
\]
where \( E_1 := (-M \cdots -M) \in \mathbb{R}^{m \times Nm} \) and

\[
E_2 := \begin{pmatrix} -M_1 - \rho M y_k \\ \vdots \\ -M_N - \rho M y_k \end{pmatrix} \in \mathbb{R}^{Nm \times m}, \quad E_3 := \begin{pmatrix} K \\ \vdots \\ K \end{pmatrix} \in \mathbb{R}^{Nm \times Nm},
\]

\[
E_4 := \begin{pmatrix} M \\ \vdots \\ M \end{pmatrix} \in \mathbb{R}^{Nm \times Nm}, \quad E_5 := \begin{pmatrix} \alpha^{-1} M T^1_k \\ \vdots \\ \alpha^{-1} M T^N_k \end{pmatrix} \in \mathbb{R}^{Nm \times Nm},
\]

as well as

\[
u := \begin{pmatrix} u_{k+1,h}^1 \\ \vdots \\ u_{k+1,h}^N \end{pmatrix} \in \mathbb{R}^m, \quad \nu := y_{k+1,h} \in \mathbb{R}^m, \quad p := \begin{pmatrix} p_{k+1,h}^1 \\ \vdots \\ p_{k+1,h}^N \end{pmatrix} \in \mathbb{R}^m,
\]

and right hand side

\[
l_1 := \begin{pmatrix} -M_1 y_{d}^1 + M y_k (\mu - \rho \psi) \\ \vdots \\ -M_N y_{d}^N + M y_k (\mu - \rho \psi) \end{pmatrix}, \quad l_2 := \begin{pmatrix} M A_{1,1} u_{d}^1 + M A_{1,1} u_{d}^1 \\ \vdots \\ M A_{N,1} u_{d}^N + M A_{N,1} u_{d}^N \end{pmatrix}, \quad 0 \in \mathbb{R}^m
\]

with the notation used in Lemma 4.6.

Let us now compare the discrete Newton step (4.7) and the discrete active-set method (5.2). The entries on the diagonal of the matrix on the left hand side of (4.7) \( E_{i,i} \) are symmetric. However, for \( N > 1 \) the resulting system is not symmetric. Note that the matrix (4.7) should not be computed explicitly due to the appearance of \( K^{-1} \). Still it is possible to compute its matrix-vector multiplication. This makes it impossible to apply a direct solver or a preconditioner which is based on decomposition, i.e. LU-factorisation. However, it can be solved by iterative methods, i.e. GMRES or BiCGSTAB. The resulting system for the active-set method (5.2) is not symmetric even for \( N = 1 \), but it can be solved by a direct solver with a preconditioner, i.e. incomplete LU-factorisation.

6. Numerical Examples. The matrices are computed using DOLFIN [16, 17], which is part of the open-source computing platform FEniCS [1, 15]. The arising linear systems are solved with NumPy and SciPy.

6.1. Example 1 - Four Player Game. We consider a four player game like \( (P_\nu^{AL}) \) on the domain \( \Omega = (0, 1)^2 \) with observation domains

\[
\Omega_1 := \left( 0, \frac{1}{2} \right) \times \left( 0, \frac{1}{2} \right), \quad \Omega_2 := \left( \frac{1}{2}, 1 \right) \times \left( 0, \frac{1}{2} \right), \\
\Omega_3 := \left( \frac{1}{2}, 1 \right) \times \left( \frac{1}{2}, 1 \right), \quad \Omega_4 := \left( 0, \frac{1}{2} \right) \times \left( \frac{1}{2}, 1 \right).
\]

In this example we assume that \( S \) is the solution mapping of the state equation \( -\Delta \nu = \sum_{\nu=1}^{N} \nu \nu \) with homogeneous Dirichlet boundary conditions. The desired states are given by constant functions

\[
y_{d}^1 := 0, \quad y_{d}^2 := 1, \quad y_{d}^3 := 2, \quad y_{d}^4 := 3
\]

and we choose \( \psi(x_1, x_2) := -2x_1 + 2x_2 + 2 \), where \((x_1, x_2) \in \Omega\). For the approximation of the multiplier we set \( u_0, p_0 \) and \( \mu \) equal zero as well as \( y_0 = 10.0 \), \( \alpha = 10^{-5} \), and \( \rho = 10 \). Let us introduce the quantity

\[
\kappa(u_k) := \log \left( \frac{\| u_{k+1} - u_k \|_U}{\| u_k - u_{k-1} \|_U} \right) \left( \log \left( \frac{\| u_k - u_{k-1} \|_U}{\| u_{k-1} - u_{k-2} \|_U} \right) \right)^{-1},
\]
which is an approximation for the numerical order of convergence. If the sequence \(\{u_k\}_k \subset U\) converges superlinear we expect \(\kappa(u_k) \in (1, 2)\) for \(k\) large enough. Note that we do not have an exact solution available to compute the order of convergence, but in practice \(\kappa(u_k)\) will give a good approximation. We use a regular triangulation with different mesh sizes \(h\). We applied both, the semi-smooth Newton method and the active-set method to this type of problem. The system that arises if the active-set method is applied has been solved directly by using the \texttt{spsolve} method from the \texttt{scipy.sparse.linalg} library. The Newton equation instead has to be solved by an iterative method. Here we make use of the \texttt{gmres} method from the same library and use a tolerance of \(10^{-12}\). Since both methods are equivalent it is not surprising that the approximated order of convergence \(\kappa\) and the change of the active sets coincide for both methods. Table 1 shows the computed results dependent on \(h\) for the active-set and the semi-smooth Newton method, respectively. Clearly, the computed orders of convergence support the superlinear convergence.

We are using linear finite elements for the controls, adjoints and state variable. Let us quickly comment on our stopping criterion step 7 of Algorithm 2 or step 6 of Algorithm 3. Both algorithms stop when the active and inactive sets coincide. Due to the use of linear finite elements we compare the values on the nodes to check this condition. Let us illustrate this on the example of the set \(A_{k+1}^{a, d}\), which is defined by the inequality \(a^{-1} p_h^b(u_k, h) \leq u_{k+1}^a\). We now count all the nodes which lie in the symmetric difference of \(A_{k+1}^{a, d}\) and \(A_{k}^{a, d}\). If this returns zero, we conclude that \(A_{k+1}^{a, d} \approx A_{k}^{a, d}\) holds good enough. We count these nodes for all the active and inactive sets in each iteration and sum them up. This calculation can be found in the row labeled "nodes".

Figure 1: (Example 1) Computed state and the state constraint (transparent).

| \(k\) | \(\kappa(u_k)\) | nodes | opt AS | opt N | \(\|u_k - P_{L_{ad}}(\frac{1}{n} p_h)\|_{L^2(\Omega)}\) | \(\kappa(u_k)\) | nodes | opt AS | opt N | \(\|u_k - P_{L_{ad}}(\frac{1}{n} p_h)\|_{L^2(\Omega)}\) |
|------|----------------|-------|-------|-------|-------------------------|------|-------|-------|-------|-------------------------------|
| 1    | 2490           | 1.8e-13 | 3.1e-07 | 87    | 9543                     | 1.8e-13 | 4.9e-07 | 84    | 9543 | 1.8e-13 | 4.9e-07 | 84 |
| 2    | 818            | 8.8e-14 | 1.7e-07 | 74    | 3286                     | 8.5e-14 | 1.7e-07 | 73    | 3286 | 8.5e-14 | 1.7e-07 | 73 |
| 3    | 417            | 5.2e-14 | 1.3e-07 | 66    | 1617                     | 5.3e-14 | 1.2e-07 | 65    | 1617 | 5.3e-14 | 1.2e-07 | 65 |
| 4    | 1.3912         | 317    | 5.0e-14 | 1.0e-07 | 61 | 1.3883                   | 1256 | 5.0e-14 | 1.5e-07 | 59 |
| 5    | 1.0706         | 179    | 5.1e-14 | 1.0e-07 | 57 | 1.0391                   | 701  | 5.2e-14 | 1.5e-07 | 55 |
| 6    | 0.3861         | 102    | 5.3e-14 | 1.4e-07 | 53 | 0.4398                   | 380  | 5.3e-14 | 1.3e-07 | 52 |
| 7    | 1.0543         | 48     | 5.1e-14 | 7.8e-08 | 51 | 0.7993                   | 138  | 5.5e-14 | 1.3e-07 | 49 |
| 8    | 2.4179         | 13     | 5.2e-14 | 9.5e-08 | 48 | 2.7034                   | 78   | 5.4e-14 | 1.0e-07 | 47 |
| 9    | 1.2989         | 6      | 5.1e-14 | 1.1e-07 | 45 | 1.3089                   | 22   | 5.4e-14 | 1.0e-07 | 44 |
| 10   | 1.4880         | 2      | 5.3e-14 | 1.1e-07 | 42 | 1.4354                   | 5    | 5.4e-14 | 7.8e-08 | 42 |
| 11   | 1.7432         | 0      | 5.2e-14 | 7.8e-08 | 35 | 1.6903                   | 1    | 5.5e-14 | 7.5e-08 | 35 |
| 12   |                |        |        |       |                           | 1.9467  | 0      | 5.4e-14 | 7.8e-08 | 19 |

Table 1: (Example 1) Computed order of convergence \(\kappa(u_k)\), change of nodes of the respective active sets, optimality of the problem, i.e., \(\|u_k - P_{L_{ad}}(\frac{1}{n} p_h)\|_{L^2(\Omega)}\) for the Newton method (opt N) and the active-set method (opt AS) and number of GMRES iterations for solving the Newton system.
6.2. Example 2 - Four Player Game with Known Exact Solution. Next, we aim at solving $\mathcal{P}^{AL}_K$, where $S$ denotes the solution operator of

$$-\Delta y = \sum_{\nu=1}^N u^\nu + f \quad \text{in } \Omega, \quad y = 0 \quad \text{on } \partial \Omega,$$

where $f$ denotes a function in $L^2(\Omega)$. This setting differs slightly from the one presented above. However, it is easy to see that this does not have any impact on our convergence analysis. We investigate a four player game on the domain $\Omega = (-1,1)^2$ with observations domains

$$\Omega_1 := (-1,0) \times (-1,0), \quad \Omega_2 := (0,1) \times (-1,0),$$
$$\Omega_3 := (-1,0) \times (0,1), \quad \Omega_4 := (0,1) \times (0,1).$$

First, with $(x_1,x_2) \in \Omega$ we set the optimal state

$$\theta(x_1,x_2) := \sin(2\pi x_1) \sin(2\pi x_2).$$

With $c_1^1 := (0.5,-0.5,0.5,-0.5)$ and $c_2^2 := (0.5,0.5,-0.5,-0.5)$ we set

$$r_0 := r_0(x_1,x_2) := \sqrt{(x_1 + c_1^1)^2 + (x_2 + c_2^2)^2}$$

and define for $v = 1,\ldots,N$ the optimal adjoint states via

$$\rho_v := (-1)(-r_v^2 + 0.25)/(16r_v^4 - 8r_v^2 + 1).$$

Choosing a regularization parameter $\alpha$ and setting the control constraints $u_0^\alpha := -1.0$ and $u_0^\alpha := 20$ for all $v$ we construct the optimal control via $\hat{u}^\nu := PL_{u^\nu} \left( -\frac{1}{\pi} \rho^\nu \right)$. Due to the construction of the adjoint states we obtain $\hat{u}^\nu = 0$ in $\Omega \setminus \Omega_v$. We set $f := -\Delta \bar{y} - \sum_{\nu=1}^N \bar{u}^\nu$ so that $\theta$ and $\bar{u}^\nu$ satisfy the state equation. It remains to construct $\hat{y}_d^\nu, \nu = 1,\ldots,N$. Due to the adjoint equation we obtain

$$\hat{y}_d^\nu := \begin{cases} g + \Delta y^\nu + (\mu + \rho(g - \psi)) & \text{in } \Omega_v, \\ 0 & \text{else.} \end{cases}$$

For our numerical experiments we use $\rho := 10.0, \mu := 0$ and $\psi := 2.0$. In order to solve this problem we apply the active-set method using the initial values $(y_0, u_0, p_0) := (10,0,0)$. Due to the knowledge of the exact solution the rate $R$ and order of convergence $\kappa$ can be estimated via

$$\lim_{k \to \infty} \frac{\|u_{k+1} - \bar{u}\|_U}{\|u_k - \bar{u}\|_U} = R,$$
$$\kappa^{ex}(u_k) = \left( \log \frac{\|u_{k+1} - \bar{u}\|_U}{\|u_k - \bar{u}\|_U} \right) \left( \log \frac{\|u_k - \bar{u}\|_U}{\|u_{k-1} - \bar{u}\|_U} \right)^{-1}.$$

We solved the problems for $h \approx 0.02$ which corresponds to approximately $1.6 \cdot 10^4$ degrees of freedom and used a tolerance of $10^{-6}$ for the gmres method. For determining the rate of convergence we compute in each iteration $R(u_k) := \frac{\|u_{k+1} - \bar{u}\|_U}{\|u_k - \bar{u}\|_U}$ and denote the corresponding value of the active-set method by $R_{AS}(u_k)$ and the one of the semi-smooth Newton method by $R_N(u_k)$. Let us check on the convergence properties corresponding to different regularization parameters $\alpha$. For $\alpha < 0.002$ neither the active-set method nor the semi-smooth Newton method converged. For $\alpha = 0.1$ the upper constraint $u_0^\alpha$ is not active. Hence, we choose $\alpha \in [0.002,0.1]$. As can be seen in Table 2 the computed values $R(u_k)$ for both methods imply superlinear convergence until only very few nodes in the active set change. The result is strengthened by the corresponding computed orders of convergence, i.e., $\kappa^{ex}_N(u_k)$ for the active-set method and $\kappa^{ex}_N(u_k)$ for the semi-smooth Newton method. In contrast to Example 1 the results of the semi-smooth Newton method now differ slightly from the active-set method. This may be due to the arising expression $Sf$ in the Newton equation, which requires an additional solution of the corresponding PDE. Indeed, by introducing $\hat{y}_d^\nu := y_d^\nu - Sf$ we obtain that the tracking type term

$$\frac{1}{2} \|y - y_d^\nu\|_{L^2(\Omega)}^2 = \frac{1}{2} \|Su + Sf - y_d^\nu\|_{L^2(\Omega)}^2 = \frac{1}{2} \|Su - y_d^\nu\|_{L^2(\Omega)}^2$$

can be treated in the same way as in Section 4. The same approach has to be followed for the state constraint $\psi$, where we set $\hat{y}_d^\nu := y_d^\nu - Sf$. Since $y_d^\nu$ and $\hat{y}_d^\nu$ appear in every iteration on the right hand side of the Newton equation, the error, which arises due to the solution of $Sf$ will be accumulated over the iterations. Finally, Figure 2 depicts the sum of the computed controls and the computed state.
\[ \alpha = 0.002 \]

| \( k \) | \( R_{AS}(u_k) \) | \( \kappa^{\alpha}_{AS}(u_k) \) | nodes AS | \( R_{N}(u_k) \) | \( \kappa^{\alpha}_{N}(u_k) \) | nodes N | gmres |
|---|---|---|---|---|---|---|---|
| 1 | 0.6210 | 94091 | 0.6211 | 94091 | 23 |
| 2 | 0.5949 | 25816 | 0.5948 | 25106 | 149 |
| 3 | 0.8378 | 49534 | 0.8397 | 49726 | 155 |
| 5 | 0.2105 | 22188 | 0.2124 | 19272 | 113 |
| 6 | 0.1897 | 8368 | 0.2349 | 11216 | 86 |
| 7 | 0.0585 | 32 | 0.0049 | 56 | 104 |
| 8 | 1.0005 | -0.0002 | 0.9736 | 0 | 67 |

\[ \alpha = 0.005 \]

| \( k \) | \( R_{AS}(u_k) \) | \( \kappa^{\alpha}_{AS}(u_k) \) | nodes AS | \( R_{N}(u_k) \) | \( \kappa^{\alpha}_{N}(u_k) \) | nodes N | gmres |
|---|---|---|---|---|---|---|---|
| 1 | 0.6574 | 41917 | 0.6574 | 41909 | 18 |
| 2 | 0.5222 | 32196 | 0.5221 | 32132 | 106 |
| 4 | 0.4825 | 14404 | 0.4824 | 14492 | 102 |
| 5 | 0.3000 | 19696 | 0.2998 | 19656 | 99 |
| 6 | 0.0190 | 120 | 0.0184 | 120 | 79 |
| 7 | 0.7412 | 0.0756 | 0.0772 | 8 | 61 |
| 8 | 1.000 | 2.7e-06 | 0 | 3 |

\[ \alpha = 0.01 \]

| \( k \) | \( R_{AS}(u_k) \) | \( \kappa^{\alpha}_{AS}(u_k) \) | nodes AS | \( R_{N}(u_k) \) | \( \kappa^{\alpha}_{N}(u_k) \) | nodes N | gmres |
|---|---|---|---|---|---|---|---|
| 1 | 0.5903 | 29083 | 0.5904 | 29083 | 10 |
| 2 | 0.5924 | 32196 | 0.5923 | 32196 | 10 |
| 4 | 0.6011 | 1046 | 0.6011 | 1046 | 10 |
| 5 | 0.0190 | 0 | 0.0014 | 0 | 11 |

\[ \alpha = 0.1 \]

| \( k \) | \( R_{AS}(u_k) \) | \( \kappa^{\alpha}_{AS}(u_k) \) | nodes AS | \( R_{N}(u_k) \) | \( \kappa^{\alpha}_{N}(u_k) \) | nodes N | gmres |
|---|---|---|---|---|---|---|---|
| 1 | 0.4016 | 12182 | 0.4016 | 12182 | 10 |
| 3 | 0.6011 | 12870 | 0.5579 | 12870 | 10 |
| 4 | 0.0014 | 0 | 0.0014 | 0 | 11 |

Table 2: (Example 2) Computed rates \( R_{AS}(u_k) \), \( R_{N}(u_k) \) and order of convergence \( \kappa^{\alpha}_{AS}(u_k) \), \( \kappa^{\alpha}_{N}(u_k) \), change of nodes of the respective active sets and number of gmres iterations for solving the Newton system.

Figure 2: (Example 2) Left: Computed sum of controls, right: computed state.
Appendix. We present the proof of Theorem 3.8.

Proof. Due to Assumption 3.1 the operator $G$ satisfies the approximation condition. We are left to check the regularity condition. By Lemma 3.6 we know that a Newton derivative of $G$ at the point $u_k$ is given by

$$ D_N G(u_k) = \text{Id} + \frac{1}{\alpha} \chi_{\mathcal{I}_k} M $$

where $\chi_{\mathcal{I}_k}$ is defined as in (3.7) and $M := D_N \tilde{F}(u_k)$. Note that $\chi_{\mathcal{I}_k}$ depends on $u_k$. We want to apply the Lax-Milgram theorem to obtain boundedness of $D_N G(u_k)^{-1}$. However, a direct application to the bilinear form $(D_N G(u_k) \bar{w}, v)_{\Omega}$ is not possible since it is not coercive in general. We consider instead an equivalent formulation which satisfies the assumptions of the Lax-Milgram theorem.

We denote by $u_{k+1} := u_k + \delta_k$ the next iterate of the semi-smooth Newton method. A Newton step is given by

$$ D_N G(u_k) \delta_k = -G(u_k) $$

$$ \Leftrightarrow \left( \text{Id} + \frac{1}{\alpha} \chi_{\mathcal{I}_k} M \right) \delta_k = -u_k + P_{\mathcal{I}_k} \left( -\frac{1}{\alpha} \tilde{F}(u_k) \right) $$

$$ \Leftrightarrow \left( \text{Id} + \frac{1}{\alpha} \chi_{\mathcal{I}_k} M \right) u_{k+1} = P_{\mathcal{I}_k} \left( -\frac{1}{\alpha} \tilde{F}(u_k) \right) + \frac{1}{\alpha} \chi_{\mathcal{I}_k} M u_k. $$

Using this representation we see that

$$ u_{k+1}(x) = \begin{cases} \ u_{k}^u(x) & \text{if } x \in \mathcal{A}_k^{u,b}, \\ \left( -\frac{1}{\alpha} \tilde{F}(u_k) - \frac{1}{\alpha} M \delta_k \right)(x) & \text{if } x \in \mathcal{I}_k, \\ u_{k}^v(x) & \text{if } x \in \mathcal{A}_k^{v,b}. \end{cases} $$

where the sets $\mathcal{A}_k^{u,b}, \mathcal{A}_k^{v,b}, \mathcal{I}_k^{\pm}$ are defined by

$$ \mathcal{I}_k^{\pm} := \left\{ x \in \Omega : -\frac{1}{\alpha} \tilde{F}(u_k)^{\pm}(x) \leq u_{k}^{\pm}(x) \right\}, $$

$$ \mathcal{A}_k^{u,b} := \left\{ x \in \Omega : -\frac{1}{\alpha} \tilde{F}(u_k)^{+}(x) \in (u_{k}^u(x), u_{k}^v(x)) \right\}, $$

$$ \mathcal{A}_k^{v,b} := \left\{ x \in \Omega : -\frac{1}{\alpha} \tilde{F}(u_k)^{-}(x) \geq u_{k}^v(x) \right\}. $$

Similar to $\chi_{\mathcal{I}_k}$ we define the sets $\chi_{\mathcal{A}_k^{u,b}}$ and $\chi_{\mathcal{A}_k^{v,b}}$ in a componentwise manner based on the set $\mathcal{A}_k^{u,b}$ and $\mathcal{A}_k^{v,b}$. Using the decomposition $u_{k+1} = \chi_{\mathcal{I}_k} u_{k+1} + (1 - \chi_{\mathcal{I}_k}) u_{k+1}$ and exploiting the identities

$$ P_{\mathcal{I}_k} \left( -\frac{1}{\alpha} \tilde{F}(u_k) \right) = \chi_{\mathcal{A}_k^{u,b}} u_a + \chi_{\mathcal{A}_k^{b}} u_b + \chi_{\mathcal{I}_k} \left( -\frac{1}{\alpha} \tilde{F}(u_k) \right), $$

$$ (1 - \chi_{\mathcal{I}_k}) u_{k+1} = \chi_{\mathcal{A}_k^{u,b}} u_a + \chi_{\mathcal{A}_k^{b}} u_b, $$

we have the equivalent formulation

$$ \left( \text{Id} + \frac{1}{\alpha} \chi_{\mathcal{I}_k} M \right) \chi_{\mathcal{I}_k} u_{k+1} $$

$$ = P_{\mathcal{I}_k} \left( -\frac{1}{\alpha} \tilde{F}(u_k) \right) + \frac{1}{\alpha} \chi_{\mathcal{I}_k} M u_k - \left( \text{Id} + \frac{1}{\alpha} \chi_{\mathcal{I}_k} M \right) \left( 1 - \chi_{\mathcal{I}_k} \right) u_{k+1} $$

$$ = \chi_{\mathcal{I}_k} \left( -\frac{1}{\alpha} \tilde{F}(u_k) \right) + \frac{1}{\alpha} \chi_{\mathcal{I}_k} M u_k - \frac{1}{\alpha} \chi_{\mathcal{I}_k} M \left( 1 - \chi_{\mathcal{I}_k} \right) u_{k+1} $$

$$ = \chi_{\mathcal{I}_k} \left( -\frac{1}{\alpha} \tilde{F}(u_k) \right) + \frac{1}{\alpha} \chi_{\mathcal{I}_k} M u_k - \frac{1}{\alpha} \chi_{\mathcal{I}_k} M \left( \chi_{\mathcal{A}_k^{u,b}} u_a + \chi_{\mathcal{A}_k^{b}} u_b \right). $$
Applying the decomposition $u_{k+1} = \chi_{I_k} u_{k+1} + \chi_{A_k^i} u_a + \chi_{A_k^b} u_b$ we finally reach at

$$
\left(\text{Id} + \frac{1}{\alpha} \chi_{I_k} M \chi_{I_k}\right) u_{k+1} = \chi_{I_k} \left(- \frac{1}{\alpha} \tilde{F}(u_k)\right) + \frac{1}{\alpha} \chi_{I_k} M u_k \nonumber \\
- \frac{1}{\alpha} \chi_{I_k} M \left(\chi_{A_k^i} u_a + \chi_{A_k^b} u_b\right) \nonumber \\
+ \chi_{A_k^i} u_a + \chi_{A_k^b} u_b.
$$

Defining the bilinear form $a(w, v) = \left(\left(\text{Id} + \frac{1}{\alpha} \chi_{I_k} M \chi_{I_k}\right) w, v\right)_U$ we get

$$
a(w, v) = (w, v)_U + \frac{1}{\alpha} (\chi_{I_k} M \chi_{I_k} w, v)_U \leq c \|w\|_U \|v\|_U.
$$

Further, with Assumption 3.7 we see that $a(w, v)$ satisfies the conditions of the Lax-Milgram Theorem, which yields boundedness of $\|D_N G(u_k)^{-1}\|_{U \rightarrow U}$.

References.

[1] M. S. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes, and G. N. Wells. The FEniCS project version 1.5. Archive of Numerical Software, 3(100):9–23, 2015.

[2] M. Bergounioux, K. Ito, and K. Kunisch. Primal-dual strategy for constrained optimal control problems. SIAM J. Control Optim., 37(4):1176–1194, 1999.

[3] R. S. Burachik and A. N. Iusem. Set-valued Mappings and Enlargements of Monotone Operators, volume 8 of Springer Optimization and Its Applications. Springer, New York, 2008.

[4] E. Casas. Second order analysis for bang-bang control problems of PDEs. SIAM J. Control Optim., 50(4):2355–2372, 2012.

[5] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. SIAM J. Optim., 13(3):865–888 (2003), 2002.

[6] M. Hintermüller and T. Surowiec. A PDE-constrained generalized Nash equilibrium problem with pointwise control and state constraints. Pac. J. Optim., 9(2):251–273, 2013.

[7] M. Hintermüller, T. Surowiec, and A. Kämmler. Generalized Nash equilibrium problems in Banach spaces: theory, Nikaido-Isoda-based path-following methods, and applications. SIAM J. Optim., 25(3):1826–1856, 2015.

[8] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. Optimization with PDE Constraints, volume 23 of Mathematical Modelling: Theory and Applications. Springer, New York, 2009.

[9] K. Ito and K. Kunisch. Semi-smooth Newton methods for state-constrained optimal control problems. Systems Control Lett., 50(3):221–228, 2003.

[10] K. Ito and K. Kunisch. Semi-smooth Newton methods for variational inequalities of the first kind. M2AN Math. Model. Numer. Anal., 37(1):41–62, 2003.

[11] K. Ito and K. Kunisch. Lagrange Multiplier Approach to Variational Problems and Applications, volume 15 of Advances in Design and Control. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[12] C. Kanzow, V. Karl, D. Steck, and D. Wachsmuth. The Multiplier-Penalty method for Generalized Nash equilibrium problems in Banach spaces. Preprint SPP1962-028 of priority program "Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization" (SPP 1962), 2017.
[13] C. Kanzow, D. Steck, and D. Wachsmuth. An augmented Lagrangian method for optimization problems in Banach spaces. *SIAM J. Control Optim.*, 56(1):272–291, 2018.

[14] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and Their Applications*, volume 31 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. Reprint of the 1980 original.

[15] A. Logg, K.-A. Mardal, G. N. Wells, et al. *Automated Solution of Differential Equations by the Finite Element Method*. Springer, 2012.

[16] A. Logg and G. N. Wells. Dolfin: Automated finite element computing. *ACM Transactions on Mathematical Software*, 37(2), 2010.

[17] A. Logg, G. N. Wells, and J. Hake. *DOLFIN: a C++/Python Finite Element Library*, chapter 10. Springer, 2012.

[18] G. Stadler. Elliptic optimal control problems with $L^1$-control cost and applications for the placement of control devices. *Comput. Optim. Appl.*, 44(2):159–181, 2009.

[19] D. Steck. *Lagrange Multiplier Methods for Constrained Optimization and Variational Problems in Banach Spaces*. PhD thesis, Universität Würzburg, 2018. to appear.

[20] M. Ulbrich. *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, volume 11 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.