

THE MULTIPLIER-PENALTY METHOD FOR GENERALIZED NASH EQUILIBRIUM PROBLEMS IN BANACH SPACES*

C. KANZOW[†], V. KARL[†], D. STECK[†], AND D. WACHSMUTH[†]

Abstract. This paper deals with generalized Nash equilibrium problems (GNEPs) in Banach spaces. We give an existence result for normalized equilibria of jointly convex GNEPs and then propose an augmented Lagrangian-type algorithm for their computation. A thorough convergence analysis is conducted which considers the existence of subproblem solutions as well as the feasibility and optimality of limit points. We then apply our investigations to differential economic games and multiobjective optimal control problems governed by linear partial differential equations. Numerical results are provided to demonstrate the practical performance of the method.

Key words. generalized Nash equilibrium problem, Banach space, normalized equilibrium, augmented Lagrangian method, optimal control, differential game

AMS subject classifications. 49M, 65K, 90C, 91A

DOI. 10.1137/17M114114X

1. Introduction. We consider the following generalized Nash equilibrium problem (GNEP). Let $N \in \mathbb{N}$ be the number of players, each in control of a variable $x^\nu \in X_\nu$, where X_ν is a (real) Banach space. We write $X := X_1 \times \cdots \times X_N$ for the strategy space of all players. In the following, $x^{-\nu}$ denotes the strategies of all players except the ν th player, and $X_{-\nu}$ the corresponding strategy space. We use the notations $x = (x^\nu, x^{-\nu})$ and $X = X_\nu \times X_{-\nu}$ to emphasize the role of player ν 's variable x^ν . Each player ν attempts to solve the optimization problem

$$(1) \quad \underset{x^\nu \in X_\nu}{\text{minimize}} \quad \theta_\nu(x^\nu, x^{-\nu}) \quad \text{subject to} \quad (x^\nu, x^{-\nu}) \in \mathcal{F}.$$

Here, $\theta_\nu: X \rightarrow \mathbb{R}$ denotes the objective or utility function of player ν and $\mathcal{F} \subseteq X$ is a nonempty closed convex set. We will also assume that the objective functions $\theta_\nu(\cdot, x^{-\nu})$ are convex and continuously differentiable for any given $x^{-\nu}$; in this setting, the GNEP (1) is usually called *jointly convex*. Note that if $\mathcal{F} = \mathcal{F}_1 \times \cdots \times \mathcal{F}_N$ for closed convex sets $\mathcal{F}_\nu \subseteq X_\nu$, then (1) reduces to the standard Nash equilibrium problem (NEP) where each player attempts to solve the following:

$$(2) \quad \underset{x^\nu \in X_\nu}{\text{minimize}} \quad \theta_\nu(x^\nu, x^{-\nu}) \quad \text{subject to} \quad x^\nu \in \mathcal{F}_\nu.$$

The study of NEPs and GNEPs is of significant practical relevance since these problems arise in many application contexts, such as economics, network design, electromagnetics, and aerodynamics. For more information, we refer the reader to the books [2, 3, 26], the papers [15, 18, 53], and the references therein. The two survey papers [15, 18]

*Received by the editors July 28, 2017; accepted for publication (in revised form) November 27, 2018; published electronically March 14, 2019.

<http://www.siam.org/journals/siopt/29-1/M114114.html>

Funding: This research 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 grants KA 1296/24-1 and Wa 3626/3-1.

[†]University of Würzburg, Institute of Mathematics, Campus Hubland Nord, Emil-Fischer-Str. 30, 97074 Würzburg, Germany (kanzow@mathematik.uni-wuerzburg.de, veronika.karl@mathematik.uni-wuerzburg.de, daniel.steck@mathematik.uni-wuerzburg.de, daniel.wachsmuth@mathematik.uni-wuerzburg.de).

in particular contain a broad overview of the finite-dimensional case. However, in many applications of GNEPs, finite-dimensional objects are not enough to describe the problem structure, and a need for infinite-dimensional models arises. This is the case, for instance, in the context of shape optimization (see [53]) or if the underlying model is time dependent.

For GNEPs in Banach spaces, Carlson [11] extended the work of Rosen [48] and provided conditions for the existence and uniqueness of so-called normalized Nash equilibria. Apart from this, the literature on GNEPs in infinite dimensions is rather scattered and most papers only deal with specific problem classes, e.g., [10, 44, 45, 46, 49, 50, 53] for standard NEPs and [13, 23, 24, 25] for GNEPs. In particular, two popular classes of problems are differential games (which arise, e.g., in time-dependent economic models) and multiobjective optimal control problems involving partial differential equations. Many scenarios of these types can be expressed in our framework (1) since they involve joint constraints on, for instance, natural resources, the total pollution in a certain area, or on the state variable in the case of optimal control. More details will be given in section 6.

To the best of our knowledge, the only paper which considers an algorithmic approach to “generic” jointly convex GNEPs in infinite dimensions is [13], where a relaxation method is presented. The aim of the present paper is to discuss some theoretical background on GNEPs in infinite dimensions and to provide an alternative algorithm. Our main approach is to apply an augmented Lagrangian (or multiplier-penalty) scheme to eliminate some or all of the constraints in (1) and therefore reduce the GNEP to a sequence of “easier” problems. This idea is not completely new since there are multiple penalty-type algorithms for GNEPs in finite and infinite dimensions [16, 21, 24, 31, 42]. On the other hand, some research has been conducted on augmented Lagrangian methods for constrained optimization in Banach spaces [27, 28, 33]. Thus, it is natural to consider an augmented Lagrangian-type method for GNEPs in our general setting (1).

It should be noted that much of the theory for augmented Lagrangian methods in the existing literature cannot readily be adapted to GNEPs in general Banach spaces. In particular, it is well known that regularity properties such as constraint qualifications can be rather subtle in infinite dimensions (as opposed to the finite-dimensional setting in, e.g., [31, 42]). On the other hand, many of the arguments for constrained optimization in Banach spaces [33, 28] rely on descent properties or, more generally, the ability to compare function values in order to get an indicator of optimality. This is in general not possible for NEPs or GNEPs, which implies that a more sophisticated analysis is needed. In the present paper, we solve these issues by using an embedding framework to enable the fulfillment of constraint qualifications, and by using the well-known Nikaido–Isoda (NI) function in order to control the optimality of the iterates.

The paper is organized as follows. In section 2, we deal with some preliminary results such as the existence of solutions of the GNEP and state the KKT conditions of the problem considered. Section 3 contains a detailed description of our proposed multiplier-penalty method. Section 4 is dedicated to the convergence analysis, where we cover solvability of the arising penalized subproblems as well as feasibility and Nash optimality of every weak limit point generated by the presented algorithm. Section 5 leads us to strong convergence of the primal iterates and weak-* convergence of the corresponding multiplier sequence under certain regularity assumptions. In section 6, we discuss applications and the numerical results of our approach in the context of differential games and optimal control problems. Section 7 contains some final remarks.

Notation. Throughout the paper, X and Y are Banach spaces, and we denote strong and weak convergence by \rightarrow and \rightharpoonup , respectively. Moreover, duality pairings are written as $\langle \cdot, \cdot \rangle$, scalar products in Hilbert spaces are written as (\cdot, \cdot) , and norms are denoted by $\|\cdot\|$ with an appropriate subscript to emphasize the corresponding space (e.g., $\|\cdot\|_X$). If S is a nonempty closed convex subset of a Hilbert space, we denote by P_S and d_S the projection and distance function to S , respectively. The partial derivative with respect to x^ν is denoted by D_{x^ν} .

2. Preliminaries. This section is dedicated to proving the existence of solutions of the jointly convex GNEP, and to its KKT conditions. Before we analyze these two topics, we first recall the definitions of generalized Nash equilibria and normalized equilibria.

2.1. Equilibria and normalized equilibria. For the definition of the two types of equilibria, recall that \mathcal{F} is the joint constraint set of all players. For a given point $x^{-\nu} \in X_{-\nu}$, we write

$$\mathcal{F}_\nu(x^{-\nu}) := \{x^\nu \in X_\nu : (x^\nu, x^{-\nu}) \in \mathcal{F}\}$$

for the feasible set of player ν 's optimization problem. Note that this set might be empty for some (or many) $x^{-\nu}$. If we are dealing with a standard NEP, then \mathcal{F}_ν is independent of $x^{-\nu}$ (see (2)). Hence, in that case, \mathcal{F}_ν is always nonempty.

DEFINITION 2.1. Let $\bar{x} \in \mathcal{F}$ be a feasible point. We say that \bar{x} is

(a) a generalized Nash equilibrium or simply a solution of the GNEP if, for every ν ,

$$(3) \quad \theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \leq \theta_\nu(y^\nu, \bar{x}^{-\nu}) \quad \text{for all } y^\nu \in \mathcal{F}_\nu(\bar{x}^{-\nu});$$

(b) a normalized (Nash) equilibrium if

$$(4) \quad \sum_{\nu=1}^N \theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \leq \sum_{\nu=1}^N \theta_\nu(y^\nu, \bar{x}^{-\nu}) \quad \text{for all } y \in \mathcal{F}.$$

Note that every normalized equilibrium is also a generalized Nash equilibrium, which can be seen by inserting points of the form $y := (y^\nu, \bar{x}^{-\nu})$ into (4). The converse, however, is not true in general. For NEPs, both concepts are equivalent.

The existence of Nash equilibria is a rather delicate topic, even when restricted to standard NEPs or finite-dimensional problems. Most existence results [1, 5, 15, 40] assume

- (i) some kind of compactness of the set \mathcal{F} , and
- (ii) an appropriate continuity property for the functions θ_ν .

In the infinite-dimensional setting, condition (i) effectively forces us to work in the weak topology of the underlying space; however, this introduces certain problems with condition (ii), since few functions are actually continuous in the weak topology. Note that (weak) lower semicontinuity of θ_ν with respect to x^ν is not enough for existence: consider, for instance, the NEP given by

$$\theta_1(x, y) := \begin{cases} x^2 & \text{if } y < 1, \\ -x & \text{if } y = 1, \end{cases} \quad \theta_2(x, y) := \frac{1}{2}y^2 + (x-1)y, \quad \mathcal{F} := [0, 1]^2,$$

where x, y are the respective player variables. Then both objective functions are continuous and convex with respect to the corresponding variable, but θ_1 is not

continuous with respect to y . It is easily verified that this problem does not admit a Nash equilibrium.

On the other hand, the compactness of \mathcal{F} is also hard to relax, even in seemingly “good” cases. For instance, the unconstrained NEP [15, Ex. 4.5] given by

$$\theta_1(x, y) := \frac{1}{2}x^2 - xy, \quad \theta_2(x, y) := \frac{1}{2}y^2 - (x+1)y,$$

where x, y are again the respective player variables, does not admit a Nash equilibrium, even though both objective functions are strongly convex, and uniformly so with respect to the rival’s variable.

It follows from the above observations that a careful approach to the existence of Nash equilibria is necessary in our setting. To this end, we consider the existence of normalized equilibria, and define the NI function [41] by

$$(5) \quad \Psi(x, y) := \sum_{\nu=1}^N [\theta_{\nu}(x^{\nu}, x^{-\nu}) - \theta_{\nu}(y^{\nu}, x^{-\nu})].$$

It is evident that a point $\bar{x} \in \mathcal{F}$ is a normalized equilibrium if and only if

$$(6) \quad \Psi(\bar{x}, y) \leq 0 \quad \forall y \in \mathcal{F},$$

which is equivalent to \bar{x} being a solution of the maximization problem

$$(7) \quad \max_y \Psi(\bar{x}, y) \quad \text{s.t.} \quad y \in \mathcal{F}.$$

Problems of the type in (6) are usually referred to as equilibrium problems [29, 30]. Taking into account the existence theory for equilibrium problems, we are prompted to make the following assumption.

Assumption 2.2. The NI function (5) is weakly sequentially lower semicontinuous (l.s.c.) with respect to x .

Before we discuss this assumption, let us first give a direct consequence, which is an existence theorem for normalized equilibria.

THEOREM 2.3. *Let Assumption 2.2 hold and assume that \mathcal{F} is nonempty and weakly compact. Then the GNEP admits a normalized equilibrium.*

Proof. This follows from the Ky–Fan theorem (cf. [17, Thm. 1], [30, Thm. 1.1]). Note that the level sets of $\Psi(\cdot, y)$ are weakly sequentially closed subsets of C , and hence weakly sequentially compact and therefore weakly closed by the Eberlein–Šmulian theorem. This implies that $\Psi(\cdot, y)$ is not only weakly *sequentially* l.s.c. but weakly (topologically) l.s.c. \square

The assumption that Ψ is weakly sequentially l.s.c. with respect to x arises naturally from the Ky–Fan theorem. However, unless X is finite-dimensional (in which case Assumption 2.2 is implied by ordinary continuity), this is a nontrivial requirement due to the minus sign in (5). Clearly, a sufficient condition is the weak sequential lower semicontinuity of the functions

$$x \mapsto \theta_{\nu}(x^{\nu}, x^{-\nu}) - \theta_{\nu}(y^{\nu}, x^{-\nu})$$

for all ν and fixed y^{ν} , which can be expected to hold in certain applications. In fact, a rather common situation is $\theta_{\nu}(x) = \theta_{\nu}^1(x) + \theta_{\nu}^2(x^{\nu})$, where θ_{ν}^1 is weakly sequentially

continuous (e.g., if it involves operators that are compact or completely continuous), and θ_ν^2 is weakly sequentially l.s.c. in x^ν . This setting encompasses various potential-type games as well as the optimal control framework from [24, 25]. See section 6 for more details.

2.2. Cones, convexity, and KKT conditions. For a nonempty convex set S and a point x in some Banach space Z , we denote by

$$\mathcal{R}_S(x) := \{\alpha(s - x) \mid \alpha \geq 0, s \in S\}, \quad \mathcal{T}_S(x) := \overline{\mathcal{R}_S(x)}$$

the *radial cone* (or cone of feasible directions) and the (*Bouligand*) *tangent cone* of S at x , respectively. Furthermore, we write

$$S^+ := \{\varphi \in Z^* : \langle \varphi, s \rangle \geq 0 \ \forall s \in S\}, \quad S^\circ := -S^+$$

for the dual and polar cones of S . If the underlying space is a Hilbert space, we treat S^+ and S° as subsets of Z rather than Z^* .

We now consider a special case of the GNEP (1) where the constraint set \mathcal{F} is given by

$$(8) \quad \mathcal{F} := \{x \in C : g(x) \in K\}.$$

Here, the function $g: X \rightarrow Y$ represents the joint constraints (i.e., the constraint that couples the players' individual strategies), Y is assumed to be a Banach space, and $K \subseteq Y$ is a nonempty closed convex cone. The set C denotes the players' individual constraints, which are given by

$$C = C_1 \times \cdots \times C_N.$$

To make the feasible set \mathcal{F} convex and the GNEP a *jointly convex* one, recall that the cone K induces the order relation

$$(9) \quad y \leq_K z : \Longleftrightarrow z - y \in K,$$

which allows us to extend various familiar concepts from finite-dimensional optimization theory to our setting. For instance, we say that g is concave if

$$g(\alpha x + (1 - \alpha)y) \geq_K \alpha g(x) + (1 - \alpha)g(y)$$

holds for all $x, y \in X$ and $\alpha \in [0, 1]$. Other notions that involve an order, such as increasing, decreasing, or convex functions, are also defined in a straightforward way. For instance, the distance function $d_K: Y \rightarrow \mathbb{R}$ is decreasing since $z \geq_K y$ implies $z = y + k$, $k \in K$, and

$$(10) \quad d_K(z) = d_K(y + k) \leq \|y + k - (P_K(y) + k)\| = \|y - P_K(y)\| = d_K(y),$$

where the inequality uses the convexity of K . It turns out that concavity of g with respect to K is the appropriate condition to ensure the convexity of the set \mathcal{F} . This result, along with some other useful observations, is formulated in the following lemma.

LEMMA 2.4. Assume that $g: X \rightarrow Y$ is concave. If $m: Y \rightarrow \mathbb{R}$ is convex and decreasing, then $m \circ g$ is convex. In particular,

- (a) the function $d_K \circ g: X \rightarrow \mathbb{R}$ is convex,
- (b) if $\lambda \in K^\circ$, then $\langle \lambda, g \rangle: X \rightarrow \mathbb{R}$ is convex, where $\langle \lambda, g \rangle(x) := \langle \lambda, g(x) \rangle$,
- (c) the set $\mathcal{F} = \{x \in C: g(x) \in K\}$ is convex.

Proof. Let $x, y \in X$ and $x_\alpha = \alpha x + (1 - \alpha)y$, $\alpha \in (0, 1)$. Then $g(x_\alpha) \geq_K \alpha g(x) + (1 - \alpha)g(y)$ by the concavity of g . Applying m on both sides yields

$$m(g(x_\alpha)) \leq m(\alpha g(x) + (1 - \alpha)g(y)) \leq \alpha m(g(x)) + (1 - \alpha)m(g(y)),$$

where we have used the monotonicity and the convexity of m . Hence, the real-valued mapping $m \circ g$ is convex in the usual sense. Assertion (a) now follows because d_K is decreasing (see above) and convex [4, Cor. 12.12]. Similarly, for (b), the function $y \mapsto \langle \lambda, y \rangle$ with $\lambda \in K^\circ$ is obviously a convex function, and it is decreasing because $\langle \lambda, k \rangle \leq 0$ for all $k \in K$. Finally, for (c), note that

$$\mathcal{F} = \{x \in C: g(x) \in K\} = C \cap \{x \in X: d_K(g(x)) \leq 0\}.$$

The second set is a lower level set of the convex function $d_K \circ g$. Hence, \mathcal{F} is convex. \square

We now discuss the KKT conditions of the GNEP.

DEFINITION 2.5. A tuple $(\bar{x}, \bar{\lambda}^1, \dots, \bar{\lambda}^N) \in X \times (Y^*)^N$ is a KKT point of the GNEP if

$$\begin{aligned} D_{x^\nu} \theta_\nu(\bar{x}) + (D_{x^\nu} g(\bar{x}))^* \bar{\lambda}^\nu &\in \mathcal{T}_{C_\nu}(\bar{x}^\nu)^+, \\ \bar{x} \in \mathcal{F}, \quad \bar{\lambda}^\nu &\in K^\circ, \quad \text{and} \quad \langle \bar{\lambda}^\nu, g(\bar{x}) \rangle = 0 \end{aligned}$$

hold for all ν .

The connection between the GNEP and its KKT conditions is well known and essentially follows from the fact that the KKT system of the GNEP is just the concatenation of the KKT systems of each player. Since the player problems are convex, it follows that KKT points are always solutions of the GNEP. Moreover, if \bar{x} is a solution of the GNEP and an appropriate constraint qualification (like that of Robinson, Zowe, and Kurcyusz; see section 5.2) is satisfied, then there are multipliers $\bar{\lambda}^1, \dots, \bar{\lambda}^N$ such that $(\bar{x}, \bar{\lambda}^1, \dots, \bar{\lambda}^N)$ is a KKT point of the GNEP. In this case, the joint constraint yields separate multipliers $\bar{\lambda}^\nu$ for each player.

For normalized equilibria (cf. Definition 2.1), it is possible to obtain a stronger notion of KKT points, cf. [11]. To this end, recall that \bar{x} is a normalized equilibrium if and only if $\Psi(\bar{x}, y) \leq 0$ for all $y \in \mathcal{F}$ or, equivalently, if \bar{x} solves the concave maximization problem (7). Assuming a suitable constraint qualification, which will be given later, for the set \mathcal{F} , this problem is equivalent to its KKT conditions, which are given by

$$\begin{aligned} -D_y \Psi(\bar{x}, \bar{x}) + g'(\bar{x})^* \bar{\lambda} &\in \mathcal{T}_C(\bar{x})^+, \\ \bar{x} \in \mathcal{F}, \quad \bar{\lambda} &\in K^\circ, \quad \text{and} \quad \langle \bar{\lambda}, g(\bar{x}) \rangle = 0, \end{aligned}$$

where D_y is the derivative with respect to y . Recalling the definition of the NI function and the product form of the set C , we get $\mathcal{T}_C(\bar{x})^+ = \mathcal{T}_{C_1}(\bar{x}^1)^+ \times \dots \times \mathcal{T}_{C_N}(\bar{x}^N)^+$ (see [39, Prop. 1.2]), and the first inclusion can be reformulated as

$$D_{x^\nu} \theta_\nu(\bar{x}) + D_{x^\nu} g(\bar{x})^* \bar{\lambda}^\nu \in \mathcal{T}_{C_\nu}(\bar{x}^\nu)^+$$

for all ν . In other words, \bar{x} satisfies the KKT conditions from Definition 2.5 with $\bar{\lambda}^\nu := \bar{\lambda}$ for each ν , i.e., the multiplier is the same for every player.

3. The multiplier-penalty method. The method we present in this section computes normalized equilibria of GNEPs whose constraint set has the form (8) with $g: X \rightarrow Y$ a continuously differentiable concave operator. For the construction of the method, we assume that there is a (linear and continuous) dense embedding $e: Y \rightarrow H$ for some Hilbert space H , and that $K_H \subseteq H$ is a closed convex cone with preimage $e^{-1}(K_H) = K$. Hence, we have

$$\mathcal{F} = \{x \in C: g(x) \in K\} = \{x \in C: e(g(x)) \in K_H\}.$$

Moreover, we use the Moreau decomposition theorem, which can be stated as follows.

LEMMA 3.1. *Every $y \in H$ can be uniquely written as $y = y_+ + y_-$ with $y_+ \in K_H$, $y_- \in K_H^\circ$, and $y_+ \perp y_-$. Moreover, we have $y_+ = P_{K_H}(y)$ and $y_- = P_{K_H^\circ}(y)$.*

In the following, $(\cdot)_-$ and $(\cdot)_+$ will always denote the projections from the Moreau decomposition. Let us note that $Y \hookrightarrow H \hookrightarrow Y^*$ is a Gel'fand triple, and that $e^*K_H^\circ \subseteq K^\circ$. For the remainder of this paper, we will usually omit the embeddings e and e^* , and we will also suppress the Riesz isomorphism between H and H^* .

Let us now turn to the multiplier-penalty method for the GNEP (1). The main idea of the method is to replace the (supposedly difficult) GNEP by a sequence of standard NEPs that include the constraint g within a penalty term. For the formal description of the method, we define the Lagrangian of player ν as

$$(11) \quad L^\nu: X \times H \rightarrow \mathbb{R}, \quad L^\nu(x, \lambda) := \theta_\nu(x) + (\lambda, g(x)),$$

and the corresponding augmented Lagrangian as

$$(12) \quad L_\rho^\nu: X \times H \rightarrow \mathbb{R}, \quad L_\rho^\nu(x, \lambda) := \theta_\nu(x) + \frac{\rho}{2} \left\| \left(g(x) + \frac{\lambda}{\rho} \right)_- \right\|_H^2.$$

Recall that $\|(\cdot)_-\|$ is the distance to K_H by the Moreau decomposition. Moreover, we note that there are other variants of L_ρ in the literature. However, these differ from (12) only by an additive constant (with respect to x).

Note that the augmented Lagrangian is a generalization of the corresponding concept for finite-dimensional optimization (see, e.g., [8, 9, 31, 47]). The only difference is that some authors use an augmented Lagrangian with an additional term (which arises from the standard slack variable approach). However, this term is independent of x , so its omission does not influence the minimization of L_ρ^ν with respect to x .

For the definition of our penalty updating scheme, we also define the auxiliary function

$$(13) \quad V(x, \lambda, \rho) := \left\| g(x) - \left(g(x) + \frac{\lambda}{\rho} \right)_+ \right\|_H.$$

This enables us to formulate our algorithm as follows.

Algorithm 3.2 (multiplier-penalty method).

- (S.0) Choose $(x_0, \lambda_0) \in X \times H$, a nonempty bounded set $B \subseteq H$, parameters $\rho_0 > 0$, $\gamma > 1$, $\tau \in (0, 1)$, and set $k := 0$.
- (S.1) If (x_k, λ_k) satisfies a suitable stopping criterion, STOP.

(S.2) Choose $w_k \in B$ and compute an approximate KKT point (see Assumption 3.3) x_{k+1} of the NEP consisting of the minimization problems

$$(14) \quad \min_{x^\nu} L_{\rho_k}^\nu(x^\nu, x^{-\nu}, w_k) \quad \text{s.t.} \quad x^\nu \in C_\nu.$$

(S.3) Update the multiplier estimate to

$$\lambda_{k+1} := (w_k + \rho_k g(x_{k+1}))_-.$$

(S.4) If $k = 0$ or

$$(15) \quad V(x_{k+1}, w_k, \rho_k) \leq \tau V(x_k, w_{k-1}, \rho_{k-1})$$

holds, set $\rho_{k+1} := \rho_k$. Otherwise, set $\rho_{k+1} := \gamma \rho_k$.

(S.5) Set $k \leftarrow k + 1$ and go to (S.1).

Some comments are due. First, note that we consider the case $k = 0$ separately in step (S.4), since w_{k-1} and ρ_{k-1} are not defined for $k = 0$. This treatment has no influence on our convergence theory.

Second, let us emphasize the importance of the sequence (w_k) in the above method. It is best to think of w_k as a safeguarded analogue of λ_k whose boundedness is enforced by requiring that $w_k \in B$ for all k . This simple fact will be crucial for our convergence analysis. Note that similar bounding schemes have been used, e.g., for augmented Lagrangian methods in nonlinear optimization [9], and that the resulting algorithm possesses strictly stronger convergence properties than the classical augmented Lagrangian method (which uses the possibly unbounded sequence $w_k := \lambda_k$); see also the example in [32].

In practice, the algorithm is usually most efficient if w_k is kept as close as possible to λ_k , e.g., by choosing it as a projection of λ_k onto the set B (see section 6).

We now consider the subproblems occurring in Algorithm 3.2, which we refer to as the *augmented NEPs*. Note that we have not specified what constitutes an “approximate KKT point” in step (S.2). Before we make this more precise, let us introduce the notation

$$L_k^\nu(x^\nu, x^{-\nu}) := L_{\rho_k}^\nu(x^\nu, x^{-\nu}, w_k)$$

for the utility function of player ν in the augmented NEP at iteration k . Therefore, x_{k+1}^ν should be an approximate KKT point of $L_k^\nu(\cdot, x_{k+1}^{-\nu})$. This is reflected in the following assumption.

Assumption 3.3. We have $x_{k+1} \in C$ for all k , and there is a null sequence $(\varepsilon_k) \subseteq X^* = X_1^* \times \cdots \times X_N^*$ such that

$$(16) \quad D_{x^\nu} L_k^\nu(x_{k+1}) \in \mathcal{T}_{C_\nu}(x_{k+1}^\nu)^+ + \varepsilon_k^\nu$$

for all ν and k .

The above is a fairly natural assumption which basically asserts that x_{k+1} is an approximate stationary point of the subproblem, and that the degree of inexactness vanishes as $k \rightarrow \infty$. Note that we assumed that each iterate x_{k+1} satisfies the additional constraint $x \in C$ exactly. This assumption is not strictly necessary for our analysis, but it is nevertheless convenient and usually satisfied in practice since the set C is assumed to consist of “simple” constraints.

4. Convergence to Nash equilibria. We now analyze the convergence properties of Algorithm 3.2. In finite-dimensional optimization, a standard way of stating convergence theorems is to assert optimality for any accumulation point of the sequence of iterates. Since we are dealing with possibly infinite-dimensional spaces, it is more natural to consider the case of *weak* limit points. The resulting convergence theorems obviously cover the case where the sequence (x_k) has a strong limit point, since any such point is also a weak limit point.

Throughout this section, we will make extensive use of Assumption 2.2, which asserts the weak sequential lower semicontinuity of the NI function with respect to x . As we will see, this condition not only is useful for the existence of equilibria (cf. Theorem 2.3), but also implies certain convergence properties for our Algorithm 3.2.

Before we proceed, recall that $g: X \rightarrow Y$ is assumed to be concave with respect to K . In the context of our multiplier-penalty method, we used the embedding $e: Y \rightarrow H$ into the Hilbert space H and the closed convex cone $K_H \subseteq H$ with $e^{-1}(K_H) = K$. In this setting, it is quite easy to see that concavity of g with respect to K is equivalent to concavity of $e \circ g$ with respect to K_H . This should be kept in mind when applying results related to convexity, such as Lemma 2.4.

Before we proceed, we give an auxiliary result.

LEMMA 4.1. *The functions $h_k(x) := \|(g(x) + w_k/\rho_k) -\|_H^2$ are convex, continuously differentiable, and weakly sequentially lower semicontinuous.*

Proof. By Lemma 2.4 it is not difficult to see that h_k is continuous and convex, and hence weakly (sequentially) lower semicontinuous [4, Thm. 9.1]. The continuous differentiability follows from [4, Cor. 12.30]. \square

4.1. Existence of subproblem solutions. In every iteration of Algorithm 3.2, we have to solve the augmented NEP for the given values w_k and ρ_k . Since the existence of Nash equilibria is not trivial in general (see the example in section 2.1), we want to state a result regarding this issue.

For the result to hold, we need some form of compactness of the set C . Recall that we made no standing assumptions on the compactness of C (we could even have $C = X$). However, if C is weakly compact, then we get the following result.

LEMMA 4.2. *Let Assumption 2.2 be satisfied. If C is weakly compact, then the augmented NEPs (14) admit solutions for all k .*

Proof. Let $k \in \mathbb{N}$ and define $h_k(x)$ as in Lemma 4.1. Consider now the function

$$\Psi_k(x, y) := \Psi(x, y) + \frac{\rho_k}{2} [h_k(x) - h_k(y)],$$

where Ψ is the usual NI function. Then Ψ_k is weakly l.s.c. with respect to x in view of Assumption 2.2 and Lemma 4.1. Hence, as in Theorem 2.3, there is a point $\hat{x} \in C$ with $\Psi_k(\hat{x}, y) \leq 0$ for all $y \in C$. We claim that \hat{x} is a solution of the penalized NEP (14). To this end, let μ be an arbitrary player index and let $y^\mu \in C_\mu$. With $y := (y^\mu, \hat{x}^{-\mu}) \in C$ we obtain

$$\begin{aligned} 0 \geq \Psi_k(\hat{x}, y) &= \sum_{\nu=1}^N [\theta_\nu(\hat{x}^\nu, \hat{x}^{-\nu}) - \theta_\nu(y^\nu, \hat{x}^{-\nu})] + \frac{\rho_k}{2} [h_k(\hat{x}) - h_k(y)] \\ &= \theta_\mu(\hat{x}^\mu, \hat{x}^{-\mu}) - \theta_\mu(y^\mu, \hat{x}^{-\mu}) + \frac{\rho_k}{2} [h_k(\hat{x}) - h_k(y)] \\ &= L_k^\mu(\hat{x}) - L_k^\mu(y). \end{aligned}$$

This completes the proof. \square

If C is not weakly compact (e.g., unbounded), then the existence of penalized Nash equilibria becomes more complicated. In theory, an appropriate form of coercivity should yield the existence of solutions of the subproblems, but this is a rather delicate topic due to the involved nature of NEPs and GNEPs (see the discussion in section 2 and in [15]).

4.2. Convergence to Nash equilibria. The aim of this section is to show feasibility and optimality of every weak limit point of the sequence (x_k) generated by Algorithm 3.2. We start by addressing feasibility.

LEMMA 4.3. *Every weak limit point of (x_k) is feasible.*

Proof. Recall that $\|g_-\|_H$ is the distance of $g(x)$ to K_H . Exploiting the properties of distance functions, we get that $\|g_-\|_H$ is convex and continuous, and hence weakly sequentially lower semicontinuous. Let us first consider the case where (ρ_k) remains bounded. The penalty updating scheme (15), together with $\tau \in (0, 1)$, yields

$$\left\| g(x_{k+1}) - \left(g(x_{k+1}) + \frac{w_k}{\rho_k} \right)_+ \right\|_H \rightarrow 0.$$

But $(g(x_{k+1}) + w_k/\rho_k)_+ \in K_H$. Therefore, $\|g_-(x_{k+1})\|_H = d_K(g(x_{k+1})) \rightarrow 0$, and the result follows. We now assume that $\rho_k \rightarrow \infty$, and define $h_k(x)$ as in Lemma 4.1. Let $x_{k+1} \rightarrow_{\mathcal{K}} \bar{x}$ for some $\mathcal{K} \subseteq \mathbb{N}$ and assume that \bar{x} is infeasible, i.e., $\|g_-(\bar{x})\|_H > 0$. Since \mathcal{F} is nonempty, there is a point $y \in \mathcal{F}$, and we get

$$\liminf_{k \in \mathcal{K}} [h_k(x_{k+1}) - h_k(y)] = \liminf_{k \in \mathcal{K}} h_k(x_{k+1}) = \liminf_{k \in \mathcal{K}} \|g_-(x_{k+1})\|_H^2 \geq \|g_-(\bar{x})\|_H^2 > 0.$$

Hence, there is a constant $c_1 > 0$ such that $h_k(x_{k+1}) - h_k(y) \geq c_1$ for all $k \in \mathcal{K}$ sufficiently large. Since h_k is convex and continuously differentiable by Lemma 4.1, it follows that

$$(17) \quad \langle h'_k(x_{k+1}), y - x_{k+1} \rangle \leq h_k(y) - h_k(x_{k+1}) \leq -c_1$$

for all $k \in \mathcal{K}$ sufficiently large. Now, let (ε_k) be the sequence from Assumption 3.3. Then

$$\begin{aligned} \langle \varepsilon_k, y - x_{k+1} \rangle &\leq \sum_{\nu=1}^N \langle D_{x^\nu} L_k^\nu(x_{k+1}), y^\nu - x_{k+1}^\nu \rangle \\ &= \sum_{\nu=1}^N \left[D_{x^\nu} \theta_\nu(x_{k+1})(y^\nu - x_{k+1}^\nu) \right] + \frac{\rho_k}{2} \langle h'_k(x_{k+1}), y - x_{k+1} \rangle \\ &\leq \sum_{\nu=1}^N \left[\theta_\nu(y^\nu, x_{k+1}^{-\nu}) - \theta_\nu(x_{k+1}) \right] + \frac{\rho_k}{2} \langle h'_k(x_{k+1}), y - x_{k+1} \rangle \\ &= \frac{\rho_k}{2} \langle h'_k(x_{k+1}), y - x_{k+1} \rangle - \Psi(x_{k+1}, y), \end{aligned}$$

where Ψ is the NI function from (5). By Assumption 2.2, Ψ is weakly l.s.c. with respect to the first argument; hence, there is a constant $c_2 \in \mathbb{R}$ such that $\Psi(x_{k+1}, y) \geq c_2$ for all $k \in \mathcal{K}$. This, together with (17), implies

$$\langle \varepsilon_k, y - x_{k+1} \rangle \leq -\frac{\rho_k c_1}{2} - c_2 \rightarrow -\infty$$

and therefore contradicts $\varepsilon_k \rightarrow 0$. \square

The feasibility of the iterates or their (weak) limit points is obviously a crucial issue for the success of Algorithm 3.2, since the method is essentially a penalty-type method. Before proving the optimality of weak limit points, we first need a technical lemma which will also be of use later on.

LEMMA 4.4. *We have $\liminf_{k \rightarrow \infty} \langle \lambda_{k+1}, g(x_{k+1}) \rangle \geq 0$.*

Proof. Using the Moreau decomposition, we obtain

$$\begin{aligned} (\lambda_{k+1}, g(x_{k+1})) &= \frac{1}{\rho_k} (\lambda_{k+1}, w_k + \rho_k g(x_{k+1})) - \frac{1}{\rho_k} (\lambda_{k+1}, w_k) \\ (18) \quad &= \frac{1}{\rho_k} \left[\|\lambda_{k+1}\|_H^2 - (\lambda_{k+1}, w_k) \right]. \end{aligned}$$

Now, if (ρ_k) is bounded, then (15) implies

$$\left\| \left(g(x_{k+1}) + \frac{w_k}{\rho_k} \right)_- - \frac{w_k}{\rho_k} \right\|_H = \left\| g(x_{k+1}) - \left(g(x_{k+1}) + \frac{w_k}{\rho_k} \right)_+ \right\|_H \rightarrow 0.$$

But the latter is obviously equal to $\|\lambda_{k+1} - w_k\|_H / \rho_k$. Therefore, $\|\lambda_{k+1} - w_k\|_H \rightarrow 0$, which implies the boundedness of (λ_{k+1}) in H as well as $\|\lambda_{k+1}\|_H^2 - (\lambda_{k+1}, w_k) = (\lambda_{k+1}, \lambda_{k+1} - w_k) \rightarrow 0$. Hence, the desired result follows from (18). We now assume that $\rho_k \rightarrow \infty$. Note that (18) is a quadratic function in λ . A simple calculation therefore shows that

$$(\lambda_{k+1}, g(x_{k+1})) \geq -\frac{1}{4\rho_k} \|w_k\|_H^2.$$

This completes the proof. \square

Exploiting the feasibility from Lemma 4.3 and the result from Lemma 4.4, we are now able to prove that every weak limit point of (x_k) is a normalized equilibrium of the GNEP.

THEOREM 4.5. *Every weak limit point of (x_k) is a normalized equilibrium of the GNEP.*

Proof. Let $x_{k+1} \rightharpoonup_{\mathcal{K}} \bar{x}$ for some $\mathcal{K} \subseteq \mathbb{N}$ (recall that \bar{x} is feasible by Lemma 4.3), and let $y \in \mathcal{F}$ be any point. An easy calculation shows that $D_{x^\nu} L_k^\nu(x_{k+1}) = D_{x^\nu} L^\nu(x_{k+1}, \lambda_{k+1})$ for all k . Since $y^\nu \in C_\nu$ for all ν , Assumption 3.3 implies

$$\begin{aligned} \langle \varepsilon_k^\nu, y^\nu - x_{k+1}^\nu \rangle &\leq \langle D_{x^\nu} L_k^\nu(x_{k+1}), y^\nu - x_{k+1}^\nu \rangle \\ &\leq \theta_\nu(y^\nu, x_{k+1}^{-\nu}) - \theta_\nu(x_{k+1}) + \langle \lambda_{k+1}, D_{x^\nu} g(x_{k+1})(y^\nu - x_{k+1}^\nu) \rangle, \end{aligned}$$

where we have used the convexity of θ_ν with respect to x^ν in the last estimate. We now sum this inequality over all ν , use the convexity of $x \mapsto \langle \lambda_{k+1}, g(x) \rangle$ (see Lemma 2.4), and the fact that $\langle \lambda_{k+1}, g(y) \rangle \leq 0$ since $\lambda_{k+1} \in K^\circ$ and $g(y) \in K$. This yields

$$\begin{aligned} \langle \varepsilon_k, y - x_{k+1} \rangle &\leq -\Psi(x_{k+1}, y) + \langle \lambda_{k+1}, g'(x_{k+1})(y - x_{k+1}) \rangle \\ &\leq -\Psi(x_{k+1}, y) + \langle \lambda_{k+1}, g(y) - g(x_{k+1}) \rangle \\ &\leq -\Psi(x_{k+1}, y) - \langle \lambda_{k+1}, g(x_{k+1}) \rangle. \end{aligned}$$

Taking the limit $k \rightarrow_{\mathcal{K}} \infty$ on both sides and using Lemma 4.4, $\varepsilon_k \rightarrow 0$ as well as the weak sequential lower semicontinuity of Ψ with respect to the first argument, we obtain $\Psi(\bar{x}, y) \leq 0$. Since $y \in \mathcal{F}$ was arbitrary, \bar{x} is a normalized equilibrium. \square

5. Further convergence results. In this section we deal with further convergence results under stronger assumptions than those used in section 4. We prove two central results: (i) strong convergence of the primal iterates, and (ii) weak-* convergence of the multiplier sequence.

For the remainder of this section, let $F: X \rightarrow X^*$ be given by

$$F(x) := (D_{x^1}\theta_1(x) \quad \cdots \quad D_{x^N}\theta_N(x)).$$

It is well known and easy to verify that the normalized equilibria of the GNEP can be characterized by means of the variational inequality

$$(19) \quad x \in \mathcal{F}, \quad \langle F(x), y - x \rangle \geq 0 \quad \forall y \in \mathcal{F}.$$

We refer the reader to [14] for a proof of this relationship in finite dimensions that directly extends to the infinite-dimensional case. Alternatively, one may simply observe that (19) is the first-order necessary condition of the concave maximization problem (7).

5.1. Strong convergence of the primal iterates.

THEOREM 5.1. *Assume that X is reflexive and that F is strongly monotone on C , i.e., there is a $c > 0$ such that*

$$(20) \quad \langle F(x) - F(y), x - y \rangle \geq c\|x - y\|_X^2 \quad \forall x, y \in C.$$

Then there is a unique normalized equilibrium \bar{x} of the GNEP. Moreover, if Assumption 3.3 holds, then $x_k \rightarrow \bar{x}$.

Proof. Existence and uniqueness of \bar{x} for the equivalent variational inequality (19) follow from standard arguments (see, e.g., [34]). For the proof of convergence, we first show that (x_k) is bounded. By Assumption 3.3, we have $D_{x^\nu}L_k^\nu(x_{k+1}) \in \mathcal{T}_{C_\nu}(x_{k+1})^+ + \varepsilon_k^\nu$ with a null sequence $(\varepsilon_k) \subseteq X^*$. Concatenating these relations for all ν and using the formula, (12), of the augmented Lagrangian yields

$$(21) \quad F(x_{k+1}) + g'(x_{k+1})^*(w_k + \rho_k g(x_{k+1}))_- \in \mathcal{T}_C(x_{k+1})^+ + \varepsilon_k,$$

where $\mathcal{T}_C(x_{k+1})^+ = \mathcal{T}_{C_1}(x_{k+1}^1)^+ \times \cdots \times \mathcal{T}_{C_N}(x_{k+1}^N)^+$ (see [39, Prop. 1.2]). Writing

$$F_k(x) := F(x) + g'(x)^*(w_k + \rho_k g(x))_-,$$

we see that F_k is the sum of the strongly monotone function F and the gradient of the convex function $x \mapsto (\rho_k/2)\|(g(x) + w_k/\rho_k)_-\|_H^2$. Hence, F_k is strongly monotone for all k with the same modulus c as in (20). This yields

$$c\|x_{k+1} - \bar{x}\|_X^2 \leq \langle F_k(\bar{x}) - F_k(x_{k+1}), \bar{x} - x_{k+1} \rangle \leq \langle F_k(\bar{x}) - \varepsilon_k, \bar{x} - x_{k+1} \rangle.$$

Recall that d_{K_H} is monotonically decreasing by (10). This implies $\|(w_k + \rho_k g(\bar{x}))_-\|_H \leq \|(w_k)_-\|_H = \|w_k\|_H$ for all k . Hence, $(F_k(\bar{x}))$ is bounded, and we obtain the existence of a $c_1 > 0$ with $c\|x_{k+1} - \bar{x}\|_X^2 \leq c_1\|x_{k+1} - \bar{x}\|_X$. This yields the boundedness of (x_k) .

We now prove the strong convergence of (x_k) to \bar{x} . Since (x_k) is bounded and X is reflexive, it follows from Theorem 4.5 that $x_k \rightharpoonup \bar{x}$. Now, using (20), it follows that

$$c\|x_{k+1} - \bar{x}\|_X^2 \leq \langle F(x_{k+1}) - F(\bar{x}), x_{k+1} - \bar{x} \rangle.$$

Since $x_{k+1} \rightharpoonup \bar{x}$, we see that $\langle F(\bar{x}), x_{k+1} - \bar{x} \rangle \rightarrow 0$. Hence, to conclude the proof, it suffices to show that $\limsup_{k \rightarrow \infty} \langle F(x_{k+1}), x_{k+1} - \bar{x} \rangle \leq 0$. Using (21) and the definition of λ_{k+1} , we see that

$$\langle F(x_{k+1}) + g'(x_{k+1})^* \lambda_{k+1}, \bar{x} - x_{k+1} \rangle \geq -\|\varepsilon_k\|_{X^*} \cdot \|\bar{x} - x_{k+1}\|_X.$$

Thus, it suffices to show $\limsup_{k \rightarrow \infty} r_k \leq 0$, where $r_k := \langle g'(x_{k+1})^* \lambda_{k+1}, \bar{x} - x_{k+1} \rangle$. By Lemma 2.4, $x \mapsto \langle \lambda_{k+1}, g(x) \rangle$ is convex. This yields $r_k \leq \langle \lambda_{k+1}, g(\bar{x}) - g(x_{k+1}) \rangle$ and, hence, $r_k \leq -\langle \lambda_{k+1}, g(x_{k+1}) \rangle$. Therefore, the result follows from Lemma 4.4. \square

It should be noted that the strong monotonicity of F is of course a rather restrictive assumption. There are certain applications where it is satisfied (see section 6), but this is not always the case (see Remark 6.3).

5.2. Convergence of the multipliers. Having proved the strong convergence of the primal iterates, in the following theorem we want to show convergence of the multiplier sequence. Recall that the radial cones \mathcal{R}_C , \mathcal{R}_K to C and K are defined in section 2.2.

THEOREM 5.2. *Let Assumption 3.3 be satisfied. If $x_k \rightarrow \bar{x}$ and the Robinson–Zowe–Kurcyusz regularity condition*

$$(22) \quad g'(\bar{x})\mathcal{R}_C(\bar{x}) - \mathcal{R}_K(g(\bar{x})) = Y$$

is satisfied, then (λ_k) is bounded in Y^ . Furthermore, every weak- $*$ limit point of (λ_k) is a Lagrange multiplier corresponding to \bar{x} .*

Proof. By Assumption 3.3 and Lemma 4.4, we have

$$(23) \quad \begin{aligned} F(x_{k+1}) + g'(x_{k+1})^* \lambda_{k+1} &\in \mathcal{T}_C(x_{k+1})^+ + \varepsilon_k \\ \text{and } \liminf_{k \rightarrow \infty} \langle \lambda_{k+1}, g(x_{k+1}) \rangle &\geq 0 \end{aligned}$$

with a null sequence $(\varepsilon_k) \subseteq X^*$ and $\lambda_{k+1} \in K^\circ$. Since \bar{x} is feasible (Lemma 4.3), this implies the second statement. We now show the boundedness of (λ_k) in Y^* . By [55, Thm. 2.1], there is an $r > 0$ such that

$$B_r^Y \subseteq g'(\bar{x})[(C - \bar{x}) \cap B_1^X] - (K - g(\bar{x})) \cap B_1^Y,$$

where B_r^X and B_r^Y are the closed r -balls around zero in X and Y , respectively. Since $x_k \rightarrow \bar{x}$, we can choose $k_0 \in \mathbb{N}$ such that

$$\|g(x_k) - g(\bar{x})\|_Y \leq \frac{r}{4} \quad \text{and} \quad \|g'(x_k) - g'(\bar{x})\|_{\mathcal{L}(X,Y)} \leq \frac{r}{4}$$

for all $k \geq k_0$. Now, let $u \in B_r^Y$. It follows that $-u = g'(\bar{x})w - z$ with $\|w\|_X, \|z\|_Y \leq 1$, and $w = w^1 - \bar{x}$, $z = z^1 - g(\bar{x})$ for some $w^1 \in C$, $z^1 \in K$. Furthermore,

$$(24) \quad \langle \lambda_k, z \rangle = \langle \lambda_k, z^1 - g(\bar{x}) \rangle \leq \langle \lambda_k, -g(\bar{x}) \rangle \leq \langle \lambda_k, -g(x_k) \rangle + \frac{r}{4} \|\lambda_k\|_{Y^*}.$$

Moreover, by (23), $\varphi_k := F(x_k) + g'(x_k)^* \lambda_k - \varepsilon_{k-1} \in \mathcal{R}_C(x_k)^+$ for all $k \geq 2$. Hence,

$$(25) \quad \begin{aligned} \langle \lambda_k, g'(\bar{x})w \rangle &= \langle g'(\bar{x})^* \lambda_k, w \rangle \geq \langle g'(x_k)^* \lambda_k, w \rangle - \frac{r}{4} \|\lambda_k\|_{Y^*} \\ &= \langle \varepsilon_{k-1} + \varphi_k - F(x_k), w \rangle - \frac{r}{4} \|\lambda_k\|_{Y^*}. \end{aligned}$$

We now use the fact that both $\langle \lambda_k, g(x_k) \rangle$ and $\langle \varepsilon_{k-1} + \varphi_k - F(x_k), w \rangle$ are bounded from below independently of w , which is an easy consequence of (23) and $\varepsilon_k \rightarrow 0$. Putting together (24) and (25), we obtain

$$\langle \lambda_k, u \rangle = \langle \lambda_k, z \rangle - \langle \lambda_k, g'(\bar{x})w \rangle \leq \frac{r}{2} \|\lambda_k\|_{Y^*} + c$$

for some constant $c > 0$. This implies

$$\|\lambda_k\|_{Y^*} = \sup_{\|u\| \leq r} \left\langle \lambda_k, \frac{1}{r} u \right\rangle \leq \frac{1}{r} \left(c + \frac{r}{2} \|\lambda_k\|_{Y^*} \right)$$

and, hence, $\|\lambda_k\|_{Y^*} \leq 2c/r$. \square

Note that the above result obviously remains true if we assume that $x_k \rightarrow_{\mathcal{K}} \bar{x}$ on some (infinite) subset $\mathcal{K} \subseteq \mathbb{N}$. In this case, we get the boundedness of $(\lambda_k)_{\mathcal{K}}$ in Y^* and the same assertion about weak-* limit points of this subsequence.

6. Applications. In the following we give some applications and numerical examples for Algorithm 3.2. Since optimal control problems are a suitable problem class for infinite-dimensional GNEPs, we have chosen two examples of this type from the literature. As a third example we have chosen an N -person differential game.

The first example is a state-constrained elliptic optimal control problem. First, we give a detailed overview of this example and analyze in detail why this type of problem is suitable for our convergence analysis. Then, we examine another example that is control-constrained only.

All implementations of the optimal control problems with PDE constraints were done in FEniCS [37] using the DOLFIN [38] Python interface and the domain $\Omega = (0, 1)^2$. In the examples, the spaces X , Y , and H are function spaces or Cartesian products thereof. Unless stated otherwise, we choose $\rho_0 := 1$, $\tau := 0.1$, and $\gamma := 10$ as the parameters for the algorithm. Moreover, the set B is chosen as the box $B := [-10^6, 10^6]$, and we define w_k as the projection of λ_k onto B . The initial values of (y, u, p, λ) are set equal to zero.

Let us introduce a slight change of notation for this section. In the optimal control setting, the players' strategies $x^\nu \in X_\nu$ are called the controls $u^\nu \in L^2(\Omega)$. The so-called state $y \in Y$ is in general the solution of a PDE constraint that is dependent on the players' controls $u = (u^\nu, u^{-\nu})$. Here Y depends on the kind of partial differential equation. Each player's cost functional in this context is denoted by $J_\nu(y, u^\nu)$.

6.1. State-constrained elliptic optimal control problems. Let us start with a multiobjective optimal control problem including tracking-type cost functionals and elliptic PDE constraints as well as state constraints. Problems of this type arise, for instance, in the optimization of aerodynamic designs if multiple (conflicting) objectives are taken into account [44, 53], or in spot-market models [24]. For further reading about this problem class, we refer the reader to [13, 24].

Let each player ν be given a cost functional $J_\nu(y, u^\nu)$, where the state y is dependent on the decisions $u^{-\nu} \in L^2(\Omega)^{N-1}$ of the other competitors. Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded Lipschitz domain. Then player ν wants to minimize

$$(26) \quad \frac{1}{2} \|y - y_d^\nu\|_{L^2(\Omega)}^2 + \frac{\alpha_\nu}{2} \|u^\nu\|_{L^2(\Omega)}^2$$

over all $(y, u^\nu) \in (H_0^1(\Omega) \cap C(\bar{\Omega})) \times L^2(\Omega)$ subject to the partial differential equation

and pointwise control and state constraints

$$(27) \quad Ay = \sum_{\nu=1}^N \chi_{\Omega_\nu} u^\nu + f, \quad u^\nu \in U_{\text{ad}}^\nu, \quad \text{and} \quad y \geq \psi \quad \text{a.e. in } \Omega,$$

where A is a suitable elliptic differential operator (e.g., $A = -\Delta$) and $f \in L^2(\Omega)$. Moreover, $U_{\text{ad}}^\nu \subseteq L^2(\Omega)$ is a nonempty closed convex set. The given data satisfy $y_d^\nu \in L^2(\Omega)$, $\alpha_\nu > 0$, and $\psi \in C(\bar{\Omega})$. Moreover, $\chi_{\Omega_\nu}: \mathbb{R}^d \rightarrow \{0, 1\}$ denotes the characteristic function of a suitable player-specific domain $\Omega_\nu \subseteq \Omega$. The sets U_{ad}^ν are given by

$$U_{\text{ad}}^\nu := \{u^\nu \in L^2(\Omega): u_a^\nu(x) \leq u^\nu(x) \leq u_b^\nu(x) \text{ a.e. in } \Omega\}$$

with $u_a^\nu, u_b^\nu \in L^2(\Omega)$ and $u_a^\nu \leq u_b^\nu$ for all ν , and we define $U_{\text{ad}} := U_{\text{ad}}^1 \times \cdots \times U_{\text{ad}}^N$. Obviously, U_{ad} and $U_{\text{ad}}^1, \dots, U_{\text{ad}}^N$ are closed, bounded, and convex. Now, let $A := -\Delta$ and let S denote the solution operator of the standard Poisson equation with homogeneous Dirichlet boundary conditions. Then we can define the control-to-state mapping via

$$y(u) := S \left(\sum_{\nu=1}^N \chi_{\Omega_\nu} u^\nu + f \right).$$

Hence, we get the reduced formulation of the optimal control problem that coincides with the definition of a jointly convex GNEP:

$$(28) \quad \begin{aligned} \min_{u^\nu} \quad & J_\nu(u) := \frac{1}{2} \|y(u) - y_d^\nu\|_{L^2(\Omega)}^2 + \frac{\alpha_\nu}{2} \|u^\nu\|_{L^2(\Omega)}^2, \\ \text{s.t.} \quad & u^\nu \in U_{\text{ad}}^\nu, \quad y(u) \geq \psi \text{ a.e. in } \Omega. \end{aligned}$$

In the notation of our abstract setting (1), (8), we have $C = U_{\text{ad}}$, $C_\nu = U_{\text{ad}}^\nu$ for every ν , $g(u) = y(u) - \psi$, and K is the nonnegative cone in $C(\bar{\Omega})$. The feasible set \mathcal{F} as defined in (8) takes on the form $\mathcal{F} = \{u \in U_{\text{ad}}: y(u) \geq \psi\}$ and is closed, bounded, and convex. Moreover, in the context of the augmented Lagrangian method, we use the additional space $H := L^2(\Omega)$ and define K_H as the nonnegative cone in H .

The NI function of the GNEP is given by

$$(29) \quad \Psi(u, v) = \sum_{\nu=1}^N \left[J_\nu(u^\nu, u^{-\nu}) - J_\nu(v^\nu, u^{-\nu}) \right],$$

where $v = (v^\nu, v^{-\nu})$ and $u, v \in L^2(\Omega)^N$. The next lemma gives us the weak sequential lower semicontinuity of the NI function, i.e., Assumption 2.2 is satisfied.

LEMMA 6.1. *The NI function (29) is weakly sequentially l.s.c. with respect to u .*

Proof. The result follows from the weak sequential lower semicontinuity of the norm and the compactness of the solution operator S , i.e., weakly convergent sequences $u_k \rightharpoonup u$ are mapped onto strongly convergent sequences $Su_k \rightarrow Su$. \square

Lemma 6.1 has a number of important consequences. First, the weak sequential lower semicontinuity of Ψ , together with the weak compactness of \mathcal{F} , implies the existence of a normalized Nash equilibrium by Theorem 2.3. Moreover, it follows that the augmented Lagrangian subproblems generated by Algorithm 3.2 always admit solutions (Lemma 4.2) and every weak limit point of the sequence of controls (u_k) is a normalized Nash equilibrium of the GNEP (Theorem 4.5).

Strong monotonicity of the mapping F . We now show that the mapping F induced by the GNEP (28) is strongly monotone in the sense of Theorem 5.1. This function is given by

$$F(u) = (D_{u^1} J_1(u) \quad \cdots \quad D_{u^N} J_N(u)).$$

Recall that strong monotonicity of F implies the strong convergence of the whole sequence (u_k) to the unique normalized Nash equilibrium \bar{u} .

LEMMA 6.2. *The operator F is strongly monotone.*

Proof. Splitting the cost functional J_ν into two parts $J_\nu(u) = J_\nu^1(u) + J_\nu^2(u)$ with

$$J_\nu^1(u) := \frac{1}{2} \|y(u) - y_d^\nu\|_{L^2(\Omega)}^2, \quad J_\nu^2(u) := \frac{\alpha_\nu}{2} \|u^\nu\|_{L^2(\Omega)}^2$$

yields $F(u) = F_1(u) + F_2(u)$, with $F_1(u) = (D_{u^1} J_1^1(u), \dots, D_{u^N} J_N^1(u))$ and F_2 defined similarly. For the first part, we use $y(u) = S(\sum_{\nu=1}^N \chi_{\Omega_\nu} u^\nu + f)$ and obtain

$$\begin{aligned} (F_1(u) - F_1(\tilde{u}), u - \tilde{u}) &= \sum_{\nu=1}^N (S^* \chi_{\Omega_\nu}^* (y(u) - y_d^\nu) - S^* \chi_{\Omega_\nu}^* (y(\tilde{u}) - y_d^\nu), u^\nu - \tilde{u}^\nu) \\ &= \sum_{\nu=1}^N (y(u) - y(\tilde{u}), \chi_{\Omega_\nu} S(u^\nu - \tilde{u}^\nu)) = \|y(u) - y(\tilde{u})\|_{L^2(\Omega)}^2 \geq 0. \end{aligned}$$

We now analyze F_2 and set $\alpha := \min\{\alpha_1, \dots, \alpha_N\} > 0$. Then

$$(F_2(u) - F_2(\tilde{u}), u - \tilde{u}) = \sum_{\nu=1}^N (\alpha_\nu (u^\nu - \tilde{u}^\nu), u^\nu - \tilde{u}^\nu) \geq \alpha \|u - \tilde{u}\|_{L^2(\Omega)}^2,$$

and the proof is complete. \square

Let us remark here that the Tikhonov terms in the objective functions are of great importance since they yield the strong monotonicity of the operator F . In particular, $F_1(u)$ is only monotone and *not* strongly monotone, since different controls u may yield the same state $y(u)$.

Remark 6.3. In practical applications, one often needs to restrict the observation of the state y . Thus, the first part of the cost functional takes on the form

$$J_\nu^1(u) := \|T_\nu y(u) - y_d^\nu\|_{L^2(\Omega)}^2$$

with bounded linear operators T_ν , $\nu = 1, \dots, N$. In this setting, the NI function still satisfies Assumption 2.2 (see Lemma 6.1). However, in general, we cannot expect the mapping F to be strongly monotone. It follows that we cannot rely on the supplemental convergence results from section 5, but this of course has no impact on the validity of the convergence results from section 4.

Existence and convergence of multipliers. Note that the GNEP (28) admits a normalized equilibrium by Lemma 6.1 and Theorem 2.3. Denoting such an equilibrium by \bar{u} , we obtain a corresponding optimal state $\bar{y}(\bar{u})$. Using a suitable constraint qualification, we can furthermore establish the existence of a Lagrange multiplier $\bar{\lambda}$. To this end, we use the following Slater condition.

Assumption 6.4. We assume that there are $\hat{u} \in U_{\text{ad}}$ and $\sigma > 0$ such that

$$(30) \quad y(\hat{u})(x) \geq \psi(x) + \sigma \quad \forall x \in \bar{\Omega}.$$

The above assumption essentially boils down to $y(\hat{u}) - \psi$ lying in the interior of the nonnegative cone of Y . Note that $Y = H_0^1(\Omega) \cap C(\bar{\Omega})$ and, therefore, the nonnegative cone has a nonempty interior (as opposed to spaces such as $L^2(\Omega)$). Furthermore, since S is linear, it is easy to see that Assumption 6.4 is equivalent to the *linearized Slater condition*

$$\exists u \in U_{\text{ad}}: y(\bar{u}) + y'(u)(u - \bar{u}) \geq \psi + \sigma,$$

e.g., by taking $u := \hat{u}$ with \hat{u} as in (30). The linearized Slater condition in turn implies the Robinson–Zowe–Kurcyusz regularity condition, which we used in section 5 (see [54, p. 332]).

The above discussion implies two things. First, the optimal control and state (\bar{u}, \bar{y}) admit a Lagrange multiplier $\bar{\lambda} \in C(\bar{\Omega})^*$ such that the first-order necessary conditions

$$(31a) \quad A\bar{y} = \sum_{\nu=1}^N \chi_{\Omega_\nu} \bar{u}^\nu + f,$$

$$(31b) \quad A^* \bar{p}^\nu = \bar{y} - y_d^\nu + \bar{\lambda},$$

$$(31c) \quad (\chi_{\Omega_\nu} \bar{p}^\nu + \alpha_\nu \bar{u}^\nu, z - \bar{u}^\nu) \geq 0 \quad \forall z \in U_{\text{ad}}^\nu,$$

$$(31d) \quad \langle \bar{\lambda}, \bar{y} - \psi \rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})} = 0, \quad \bar{y} \geq \psi, \quad \bar{\lambda} \leq 0,$$

are satisfied for all ν ; cf. [12] and the discussion in section 2.2. Here, $\bar{p}^\nu \in W_0^{1,s}(\Omega)$, $1 < s < N/(N-1)$, is the adjoint state of player ν , and the inequality $\bar{\lambda} \leq 0$ has to be understood as $\langle \bar{\lambda}, \varphi \rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})} \leq 0$ for all $\varphi \in C(\bar{\Omega})$ with $\varphi \geq 0$. In other words, $\bar{\lambda}$ lies in the polar of the nonnegative cone of $C(\bar{\Omega})$.

The second implication of the Slater condition (or, equivalently, of the Robinson–Zowe–Kurcyusz condition) is that the assertions of Theorem 5.2 hold, i.e., the multiplier sequence (λ_k) generated by Algorithm 3.2 is bounded in $C(\bar{\Omega})^*$ and each of its weak-* limit points is a Lagrange multiplier satisfying the optimality system (31).

Now, let us briefly consider the subproblems that occur in every iteration of the algorithm. We know that, by Lemmas 6.1 and 4.2, these problems always admit a Nash equilibrium \bar{u}_k . Since the problem is convex and control constrained only, first-order necessary optimality conditions can be established without any further regularity assumptions. Setting $\bar{y}_k := y(\bar{u}_k)$, there exist unique adjoint states $\bar{p}_k^\nu \in H_0^1(\Omega)$ which satisfy (compare with (31)) the system

$$(32a) \quad A\bar{y}_k = \sum_{\nu=1}^N \chi_{\Omega_\nu} \bar{u}_k^\nu + f,$$

$$(32b) \quad A^* \bar{p}_k^\nu = \bar{y}_k - y_d^\nu + \bar{\lambda}_k,$$

$$(32c) \quad (\chi_{\Omega_\nu} \bar{p}_k^\nu + \alpha_\nu \bar{u}_k^\nu, z - \bar{u}_k^\nu) \geq 0 \quad \forall z \in U_{\text{ad}}^\nu,$$

$$(32d) \quad \bar{\lambda}_k = (w_k + \rho_k(S\bar{u}_k - \psi))_-$$

for all ν .

Summary of the convergence properties. We can associate the adjoint states $p_k^\nu = S^*(y_k - y_d^\nu + \lambda_k)$ with each iterate u_k and its state $y_k = y(u_k)$. Using standard arguments (e.g., [35, Lem. 11]), it is easy to show that $p_k^\nu \rightharpoonup \bar{p}^\nu$ in $W_0^{1,s}(\Omega)$,

$1 < s < N/(N-1)$, for all ν . Concluding, as $k \rightarrow \infty$ we have, for the sequence $(y_k, u_k^\nu, u_k^{-\nu}, p_k^\nu, \lambda_k)$ generated by Algorithm 3.2, that

$$\begin{aligned} (y_k, u_k) &\rightarrow (\bar{y}, \bar{u}) && \text{in } (H_0^1(\Omega) \cap C(\bar{\Omega})) \times L^2(\Omega)^N, \\ p_k^\nu &\rightarrow \bar{p}^\nu && \text{in } W_0^{1,s}(\Omega), \\ \lambda_k &\overset{*}{\rightharpoonup} \bar{\lambda} && \text{in } C(\bar{\Omega})^*. \end{aligned}$$

Numerical results. In the following let us report our numerical results. As a test problem, we chose the four-player game presented in [24], which is a special instance of the problem presented above where $\Omega_\nu = \Omega$ for all ν and $f \equiv 1$. The vector of Tikhonov parameters is given by $\alpha = (2.8859, 4.3374, 2.5921, 3.9481)$, and the control constraints are defined by $u_a^\nu \equiv -12$, $u_b^\nu \equiv 12$ for all ν . The state has to fulfill the state constraint for

$$\psi(x_1, x_2) = \cos(5\sqrt{(x_1 - 0.5)^2 + (x_2 - 0.5)^2}) + 0.1.$$

Defining

$$\xi_\nu(x_1, x_2) := 10^3 \max(0, 4(0.25 - \max(|x_1 - z_\nu^1|, |x_2 - z_\nu^2|)))$$

with $z^1 := (0.25, 0.75, 0.25, 0.75)$ and $z^2 := (0.25, 0.25, 0.75, 0.75)$, we set

$$y_d^1 := \xi_1 - \xi_4, \quad y_d^2 := \xi_2 - \xi_3, \quad y_d^3 := \xi_3 - \xi_2, \quad y_d^4 := \xi_4 - \xi_1.$$

The subproblems arising within the computation are solved exactly by applying an active set method [6, 7] to the corresponding KKT conditions (32). The algorithm is stopped as soon as the quantities $\|(\psi - y_k)_+\|_{C(\bar{\Omega})}$ and $|(\lambda_k, y_k - \psi)|$ drop below 10^{-6} . Since the NEPs in the inner iterations are solved exactly, these values represent the residual of the first-order system (the stationarity part is always satisfied).

The table below lists some iteration numbers for different discretization levels, where n is the number of grid points per dimension, as well as the maximum penalty parameter ρ_{\max} reached during the given iterations. Note that the inner iterations are accumulated over the whole outer iterations.

n	16	32	64	128	256
outer it.	12	12	13	14	14
inner it.	24	28	34	43	50
ρ_{\max}	10^7	10^9	10^{10}	10^{12}	10^{12}

It is worth noting that the outer iteration numbers and final penalty parameters only grow moderately as n increases. This observation suggests that our algorithm works quite well for the given optimal control problem.

Further, we applied a nested grid strategy to this problem. Here, as in [24], we refined the mesh if for a given mesh size the inequality $1000\rho^{-1} \leq \frac{1}{n^2}$ was satisfied. We refined until $n = 512$ was reached. The next table gives the outer and inner iteration numbers with the corresponding maximum of the penalty parameter that has been reached in the given mesh size.

n	4	8	16	32	64	128	256	512	Σ
outer it.	6	1	1	1	1	1	1	4	16
inner it.	8	2	4	4	5	5	5	14	47
ρ_{\max}	10^5	10^6	10^7	10^8	10^9	10^{10}	10^{11}	10^{15}	

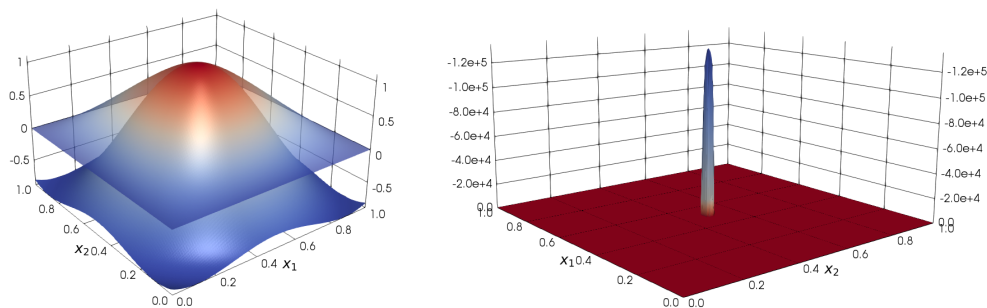


FIG. 1. (First example.) Left: computed discrete optimal state y_h (transparent, upper) and state constraint ψ . Right: computed Lagrange multiplier λ_h .

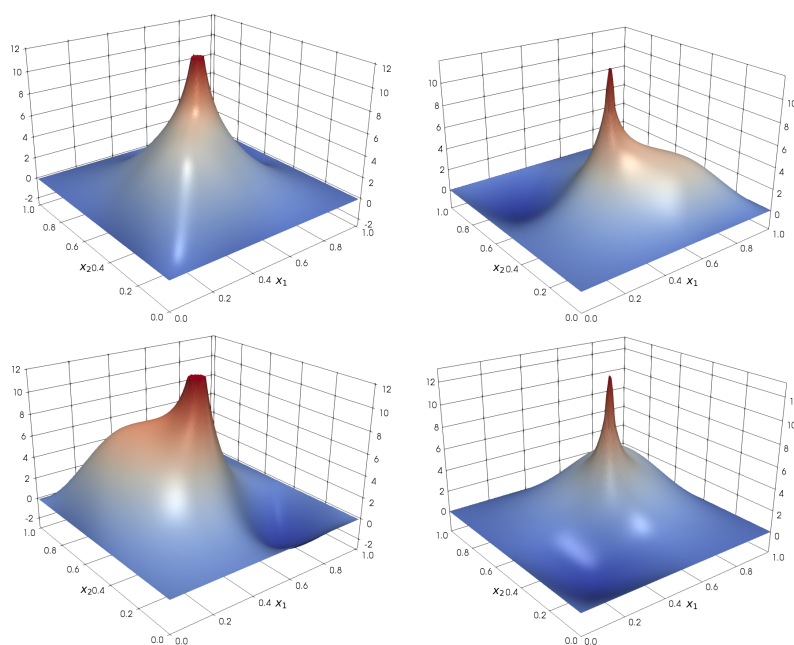


FIG. 2. (First example.) Computed optimal control $\bar{u}_h = (\bar{u}_h^1, \bar{u}_h^2, \bar{u}_h^3, \bar{u}_h^4)$.

Figures 1, 2, and 3 show the numerical solution of the first example. All figures depict results gained for a triangular mesh with $n = 128$ grid points per dimension.

6.2. Control-constrained optimal control problems. The following example does not include constraints on the state. However, this example is still of interest since it has a known analytic solution, allowing us to make error estimates on our computed solution. Let $N = 2$ be the number of players. Every player wants to minimize the tracking-type functional (26) over all $(y, u^\nu) \in (H_0^1(\Omega) \cap C(\bar{\Omega})) \times L^2(\Omega)$ subject to the PDE and control constraints

$$-\Delta y = \sum_{\nu=1}^N u^\nu + f \quad \text{and} \quad u^\nu \in U_{\text{ad}}^\nu,$$

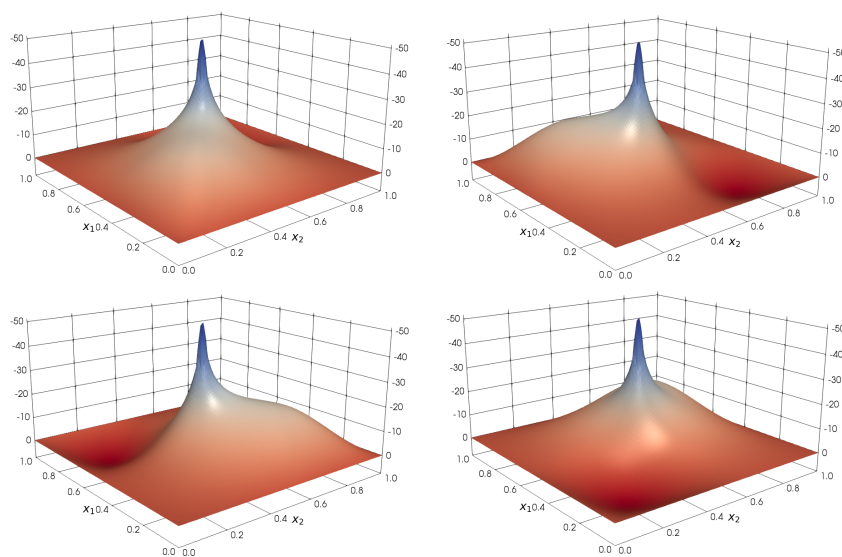


FIG. 3. (First example.) Computed adjoint state $\bar{p}_h = (\bar{p}_h^1, \bar{p}_h^2, \bar{p}_h^3, \bar{p}_h^4)$.

where $U_{\text{ad}}^\nu = \{u^\nu \in L^2(\Omega) : u_a^\nu \leq u^\nu \leq u_b^\nu \text{ a.e. in } \Omega\}$ as before and $f \in L^2(\Omega)$. Choosing $C = L^2(\Omega)^2$ and setting $g(u) = (g_1(u), g_2(u), \dots, g_N(u))$ with

$$g_\nu(u) := \begin{pmatrix} u^\nu - u_a^\nu \\ u_b^\nu - u^\nu \end{pmatrix}, \quad \nu = 1, \dots, N,$$

our set of constraints from (8) is given by $\mathcal{F} = \{u \in L^2(\Omega) : g(u) \geq 0\}$. Moreover, we have $Y := H := L^2(\Omega)^{2N}$, and $K := K_H$ is the nonnegative cone in Y .

In the present problem, each player's feasible set is independent of the rival players' strategies. Hence, $\mathcal{F}_\nu(u^{-\nu}) = U_{\text{ad}}^\nu$ for all ν , and the problem is a standard NEP. Let $\lambda_a^\nu, \lambda_b^\nu$ denote the multipliers corresponding to the lower and upper control constraints, respectively. Then the complete multiplier vector λ is given by $\lambda := (\lambda^1, \dots, \lambda^N)$ with $\lambda^\nu = (\lambda_a^\nu, \lambda_b^\nu)$ for all ν .

In this example we apply our algorithm by augmenting the given control constraints. Special care needs to be taken because the set C is not bounded and therefore not weakly compact as assumed in Lemma 4.2. However, because of the special structure of the cost functional, the augmented NEP can be reduced to a single control problem, yielding the existence of a unique normalized Nash equilibrium (cf. [25, Prop. 3.10]).

The test problem we chose was first presented in [10]. Here, we state a reformulated version from [13]. The Tikhonov parameters are given by $\alpha_1 = \alpha_2 = 1$. Moreover, we define the subsets $\Omega_\nu \subseteq \Omega$ by

$$\Omega_1 := (0, 1) \times (0, 0.5), \quad \Omega_2 := (0, 1) \times (0.5, 1)$$

and the control constraints $u_a^\nu := a_\nu \chi_{\Omega_\nu}$, $u_b^\nu := b_\nu \chi_{\Omega_\nu}$, where $a_\nu := -0.5$ and $b_\nu := 0.5$ for all ν . Finally, we set

$$y_d^1(x) := y(x) + 8\pi^2 y(2x), \quad y_d^2(x) := y(x) + 18\pi^2 y(3x),$$

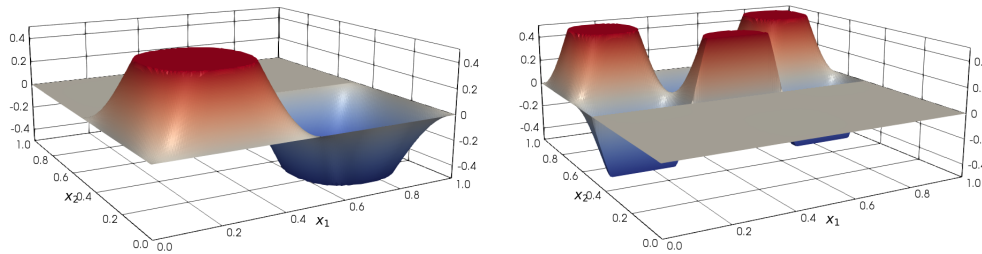


FIG. 4. (Example 2.) Computed discrete optimal control u_h^1 (left) and u_h^2 (right).

and $f := -\Delta y - u_1 - u_2$. The exact solution of the resulting problem is given by

$$y(x) := \sin(\pi x_1) \sin(\pi x_2), \quad u_1(x) := \chi_{\Omega_1} P_{[a_1, b_1]} y(2x), \quad u_2(x) := \chi_{\Omega_2} P_{[a_2, b_2]} y(3x).$$

We used $\tau := 0.6$ and stopped the algorithm as soon as

$$\sum_{\nu=1}^2 \left[\|(u_a^\nu - u^\nu)_+\|_{C(\bar{\Omega})} + |(\lambda_a^\nu, u^\nu - u_a^\nu)| + \|(u_b^\nu - u^\nu)_+\|_{C(\bar{\Omega})} + |(\lambda_b^\nu, u^\nu - u_b^\nu)| \right] \leq 10^{-7}$$

was satisfied. Once again, the subproblems occurring within the algorithm were solved exactly by applying an active set method. Thus, the stationarity term in the first-order optimality conditions is always (numerically) equal to zero, and it suffices to check feasibility and complementarity in the stopping criterion above.

Numerical results. Figure 4 depicts the obtained results for the optimal controls u_h^1 , u_h^2 of the two players using $n = 128$ grid points per dimension.

The iteration numbers of outer and inner iterations as well as the final value of the penalty parameter ρ_{\max} are shown in the table below, where n again denotes the number of grid points per dimension.

n	16	32	64	128	256	512
outer it.	10	10	10	10	9	10
inner it.	10	10	10	10	9	10
ρ_{\max}	10^1	10^1	10^1	10^1	10^1	10^2

Note that the outer iteration numbers and the final penalty parameter remain constant with increasing n . Finally, let us report about the behavior of the discretized errors with increasing dimension. The following table illustrates the corresponding L^2 -norms for different discretization levels.

n	16	32	64	128	256	512
$\ \bar{u}_1 - u_h^1\ _{L^2(\Omega)}$	1.72e-03	4.29e-04	1.16e-04	2.89e-05	7.24e-06	1.81e-06
$\ \bar{u}_2 - u_h^2\ _{L^2(\Omega)}$	3.84e-03	1.04e-03	2.59e-04	6.50e-05	1.63e-05	4.07e-06
$\ \bar{y} - y_h\ _{L^2(\Omega)}$	1.60e-03	4.01e-04	1.00e-04	2.50e-05	6.26e-06	1.57e-06

Considering the consistent number of outer and also inner iterations and the quite good approximation of the Nash equilibrium, we can conclude that our algorithm works rather well for this kind of problem.

6.3. Environmental differential games. We end with a problem from the class of N -person differential games. Problems of this type are rather popular in the literature [19, 20, 43]. They arise, for instance, in economic simulations if the underlying model is not limited to a fixed point in time but takes into account a whole time interval. In such cases, infinite-dimensional spaces arise naturally, and the problems can therefore be tackled by our algorithmic framework.

The particular example we present here is an environmental management problem based on the framework in [23]. The N players are given by N companies that compete on a common market. Let the strategy space of all players be given by $X := L^2([0, T])^N$. Let $u^\nu(t) \in \mathbb{R}$ denote the investment (control) of company ν at time t and $y^\nu(t) \in \mathbb{R}$ the production capacity. Then u^ν and y^ν are coupled through the differential equation

$$(33) \quad \dot{y}^\nu(t) + b_\nu y^\nu(t) = u^\nu(t), \quad y(0) = y_0,$$

where $b_\nu \in \mathbb{R}$. Further, at each time, the investments $u^\nu(t)$ are bounded in the sense that

$$u^\nu \in U_{\text{ad}}^\nu, \quad \text{where} \quad U_{\text{ad}}^\nu := \{u \in L^2([0, T]): 0 \leq u(t) \leq u_{\max}(t) \ \forall t \in [0, T]\}.$$

The production capacity y^ν evokes a certain environmental pollution. For the sake of pollution control, the companies have to comply with legal requirements. A global constraint of pollution is given by

$$E(y)(t) \leq \psi(t) \quad \text{for all } t \in [0, T],$$

where $\psi \in C([0, T])$, $y := (y^1, \dots, y^N)$, and E denotes some emission rate function that depends on the installed capacity. Each company's production and adjustment costs are given by the function $q_\nu(y^\nu(t), u^\nu(t))$. The market price $r_\nu(y(t))$ of the observed product is associated with the total supply $\sum_{\nu=1}^N y^\nu(t)$. Hence, the revenue of the ν th company can be modelled via $r_\nu(y(t))y^\nu(t)$. Since each company aims for maximizing its profit, each player attempts to solve the optimization problem

$$\begin{aligned} \min_{u^\nu, y^\nu} \quad & \theta_\nu(u^\nu, y) := \int_0^T q_\nu(y^\nu(t), u^\nu(t)) - r_\nu(y(t))y^\nu(t) \, dt \\ \text{s.t.} \quad & \dot{y}^\nu(t) + b_\nu y^\nu(t) = u^\nu(t), \quad y^\nu(0) = y_0^\nu, \\ & 0 \leq u^\nu(t) \leq u_{\max}^\nu(t), \quad E(y)(t) \leq \psi(t). \end{aligned}$$

By a well-known theorem of Carathéodory, the differential equation (33) admits a unique solution [22, p. 30], which can be written explicitly as

$$(34) \quad y^\nu(t) = e^t y_0^\nu + \int_0^t e^{t-s} u^\nu(s) \, ds$$

(see [52, p. 488]). Using this expression and the fact that $H^1([0, T])$ is compactly embedded in $C([0, T])$, one can easily see that $y^\nu \in H^1([0, T])$ and that the control-to-state mappings $y^\nu(\cdot): L^2([0, T]) \rightarrow C([0, T])$ are completely continuous (compact). Similarly to the previous examples, we can now pass to the reduced formulation by inserting $y^\nu := y^\nu(u^\nu)$ for all ν . This results in the GNEP where player ν attempts to solve

$$(35) \quad \min_{u^\nu} \theta_\nu(u^\nu, y(u)) \quad \text{s.t.} \quad u^\nu \in U_{\text{ad}}^\nu, \quad E(y(u)) \leq \psi.$$

Clearly, the sets U_{ad}^ν and $U_{\text{ad}} := U_{\text{ad}}^1 \times \cdots \times U_{\text{ad}}^N$ are closed, bounded, and convex. Further, the feasible set \mathcal{F} as defined in (8) takes on the form $\mathcal{F} = \{u \in U_{\text{ad}} : E(y(u)) \leq \psi\}$ and is also closed, bounded, and convex.

The following assumption guarantees the convexity of the reduced objective functions (see [23]).

Assumption 6.5. For each player ν , we assume $q_\nu(y^\nu(u^\nu), u^\nu)$ and $-r_\nu(y(u))y^\nu(u^\nu)$ are convex and continuously Fréchet differentiable for each given $u^{-\nu}$.

The NI function of the reduced GNEP is weakly sequentially l.s.c. by the arguments of Lemma 6.1. Due to Theorem 2.3, this now yields the existence of a normalized equilibrium. Despite this, the problem is more complex than those in the previous two sections since it cannot be reduced to a single control problem. Moreover, the operator $F := (D_{u^1}\theta_1, \dots, D_{u^N}\theta_N)$ may not be strongly monotone (depending on the functions q_ν and r_ν). Hence, the convergence results from section 5 cannot be applied to this example, but this does not affect our convergence results from section 4.

Similarly to the theory in section 6.1, we make the following Slater-type assumption.

Assumption 6.6. We assume that there are $\hat{u} \in U_{\text{ad}}$ and $\sigma > 0$ such that

$$E(y(\hat{u}))(t) \leq \psi(t) - \sigma \quad \forall t \in [0, T].$$

Due to this Slater condition, we get the existence of a Lagrange multiplier $\bar{\lambda} \in C([0, T])^*$ that satisfies the following optimality system (see [51, p. 61]):

$$(36a) \quad \dot{\bar{y}}^\nu + b_\nu \bar{y}^\nu = \bar{u}^\nu, \quad \bar{y}^\nu(0) = y_0^\nu,$$

$$(36b) \quad \dot{\bar{p}}^\nu - b_\nu \bar{p}^\nu = -D_{y^\nu} \theta_\nu(\bar{u}^\nu, \bar{y}) - E_\nu^* \bar{\lambda}, \quad \bar{p}^\nu(T) = 0,$$

$$(36c) \quad \bar{u}^\nu = P_{U_{\text{ad}}^\nu}(\bar{u}^\nu - (\bar{p}^\nu + D_{u^\nu} \theta_\nu(\bar{u}^\nu, \bar{y}))),$$

$$(36d) \quad \bar{\lambda} \geq 0, \quad E(\bar{y}) \leq \psi, \quad \langle \bar{\lambda}, E(\bar{y}) - \psi \rangle = 0.$$

We now describe how the problem can be tackled by the augmented Lagrangian approach. To this end, observe that the reduced formulation (35) fits precisely into our general framework (8) with $C_\nu := U_{\text{ad}}^\nu$ for all ν , $g(u) := E(y^1(u^1), \dots, y^N(u^N)) - \psi$, $Y := H^1([0, T])$, and K the negative cone in Y . For the augmentation, we set $H := L^2([0, T])$ and choose K_H as the negative cone in H .

The augmented subproblems arising in Algorithm 3.2 are convex and control-constrained only. Moreover, they yield an optimality system similar to (36). To this end, let k be the current iteration index and let $\bar{u}_k = (\bar{u}_k^\nu)_{\nu=1}^N$ be a solution of the k th subproblem. Setting $\bar{y}_k^\nu := y^\nu(\bar{u}_k^\nu)$, there exists a unique adjoint state $\bar{p}_k^\nu \in H^1([0, T])$ that satisfies the following first-order necessary optimality conditions (see [36, Thm. 14, p. 235] and [51, p. 33]):

$$(37a) \quad \dot{\bar{y}}_k^\nu + b_\nu \bar{y}_k^\nu = \bar{u}_k^\nu, \quad \bar{y}_k^\nu(0) = y_0^\nu,$$

$$(37b) \quad \dot{\bar{p}}_k^\nu - b_\nu \bar{p}_k^\nu = -D_{y^\nu} \theta_\nu(\bar{u}_k^\nu, \bar{y}_k) - E_\nu^* \bar{\lambda}_k, \quad \bar{p}_k^\nu(T) = 0,$$

$$(37c) \quad \bar{u}_k^\nu = P_{U_{\text{ad}}^\nu}(\bar{u}_k^\nu - (\bar{p}_k^\nu + D_{u^\nu} \theta_\nu(\bar{u}_k^\nu, \bar{y}_k))),$$

$$(37d) \quad \bar{\lambda}_k = (w_k + \rho_k(E(\bar{y}_k) - \psi))_+.$$

Numerical results. Let us now report our numerical results. We set $N := 2$, $T := 3$, and use the functions

$$q_\nu(y^\nu(t), u^\nu(t)) := \frac{a_1}{2} y^\nu(t)^2 + \frac{a_2}{2} u^\nu(t)^2, \quad r_\nu(y(t)) := \frac{c}{\sum_{\nu=1}^N y^\nu(t) + \epsilon},$$

$$E(y(t)) := e_1 y^1(t) + e_2 y^2(t),$$

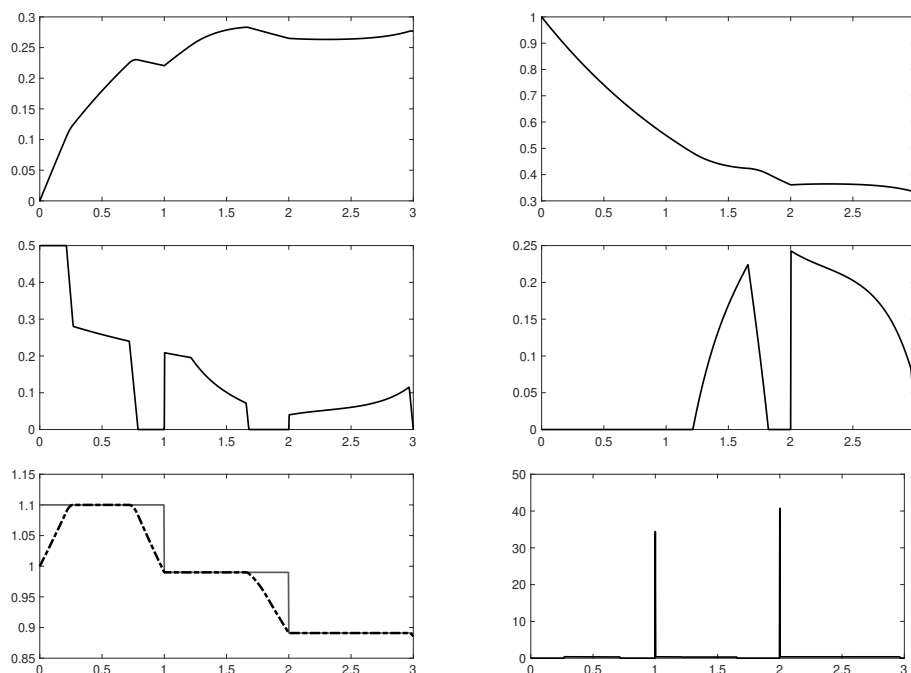


FIG. 5. (Example 3.) From top to bottom, from left to right: computed states y_h^1 , y_h^2 ; controls u_h^1 , u_h^2 ; state constraint ψ (solid gray) and $E(y_h)$ (dotted black); multiplier λ_h .

where $\epsilon > 0$. The implementation for this example was done in MATLAB. We discretize the time derivative by finite differences and solve, in each iteration, the first-order system (37) by applying a semi-smooth Newton method. Here, the subproblems arising are solved with an accuracy of 10^{-6} . We stop the algorithm as soon as the residual of the first-order system (36) drops below 10^{-6} . The initial values are given by $y_0 := u_0 := p_0 := 0.5$ and $\lambda_0 := 0$. We used the parameters $\rho_0 := 1.0$, $\tau := 0.1$, and $\gamma := 10$. The set B is chosen as the box $B := [-100, 100]$. The parameters of the model are given by $u_{\max}^v := 0.5$, $a_1 := 0.7$, $a_2 := 0.2$, $b_1 = 0.2$, $b_2 := 0.6$, $c := 1$, $e_1 := 2$, $e_2 := 1$, and $\epsilon = 10^{-9}$. The initial values of the state are given by $y^1(0) := 0$, $y^2(0) := 1$. The state constraint is given by

$$\psi(t) := \begin{cases} 1.1 & \text{for } t \in [0, 1], \\ 0.99 & \text{for } t \in (1, 2], \\ 0.891 & \text{for } t \in (2, 3], \end{cases}$$

which is a decrease of ten percent after every third of the time interval. The table below shows some iteration numbers for the given parameters, and Figure 5 depicts the calculated results for $n = 1024$ grid points.

n	32	64	128	256	512	1024	2048
outer it.	8	9	8	10	10	10	10
inner it.	74	117	90	90	94	105	96
ρ_{\max}	10^3	10^3	10^4	10^4	10^5	10^6	10^6

7. Final remarks. In this paper, we have introduced an augmented Lagrangian method for jointly convex GNEPs in Banach spaces. Under relatively weak assumptions, we obtain feasibility and optimality of every weak limit point of the generated sequence, and under additional regularity assumptions, we show strong convergence of the primal sequence and weak-* convergence of the multiplier sequence. The fact that we stated the problem in a quite general setting allows us to consider a broad range of applications, including multiobjective optimal control problems as well as N -player differential games. For both these problem classes, our numerical tests indicate that the method works quite well since it possesses good convergence properties and yields a relatively high accuracy.

Acknowledgment. The authors would like to thank three anonymous referees for their helpful suggestions, which contributed to improving the paper.

REFERENCES

- [1] K. J. ARROW AND G. DEBREU, *Existence of an equilibrium for a competitive economy*, *Econometrica*, 22 (1954), pp. 265–290, <https://doi.org/10.2307/1907353>.
- [2] P. D. BARBA, *Multiobjective Shape Design in Electricity and Magnetism*, *Lect. Notes Electr. Eng.* 47, Springer, Berlin, 2010.
- [3] T. BAŞAR AND G. J. OLSDER, *Dynamic Noncooperative Game Theory*, *Classics Appl. Math.* 23, SIAM, Philadelphia, 1999.
- [4] H. H. BAUSCHKE AND P. L. COMBETTES, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011, <https://doi.org/10.1007/978-1-4419-9467-7>.
- [5] M. R. BAYE, G. Q. TIAN, AND J. ZHOU, *Characterizations of the existence of equilibria in games with discontinuous and nonquasiconcave payoffs*, *Rev. Econom. Stud.*, 60 (1993), pp. 935–948, <https://doi.org/10.2307/2298107>.
- [6] M. BERGOUNIOUX, K. ITO, AND K. KUNISCH, *Primal-dual strategy for constrained optimal control problems*, *SIAM J. Control Optim.*, 37 (1999), pp. 1176–1194, <https://doi.org/10.1137/S0363012997328609>.
- [7] M. BERGOUNIOUX AND K. KUNISCH, *Primal-dual strategy for state-constrained optimal control problems*, *Comput. Optim. Appl.*, 22 (2002), pp. 193–224, <https://doi.org/10.1023/A:1015489608037>.
- [8] D. P. BERTSEKAS, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, 1982.
- [9] E. G. BIRGIN AND J. M. MARTÍNEZ, *Practical Augmented Lagrangian Methods for Constrained Optimization*, SIAM, Philadelphia, 2014, <https://doi.org/10.1137/1.9781611973365>.
- [10] A. BORZÌ AND C. KANZOW, *Formulation and numerical solution of Nash equilibrium multiobjective elliptic control problems*, *SIAM J. Control Optim.*, 51 (2013), pp. 718–744, <https://doi.org/10.1137/120864921>.
- [11] D. A. CARLSON, *Uniqueness of normalized Nash equilibrium for a class of games with strategies in Banach spaces*, in *Decision & Control in Management Science: Essays in Honor of Alain Haurie*, G. Zaccour, ed., Springer, Boston, MA, 2002, pp. 333–348, https://doi.org/10.1007/978-1-4757-3561-1_18.
- [12] E. CASAS, *Control of an elliptic problem with pointwise state constraints*, *SIAM J. Control Optim.*, 24 (1986), pp. 1309–1318, <https://doi.org/10.1137/0324078>.
- [13] A. DREVES AND J. GWINNER, *Jointly convex generalized Nash equilibria and elliptic multiobjective optimal control*, *J. Optim. Theory Appl.*, 168 (2016), pp. 1065–1086, <https://doi.org/10.1007/s10957-015-0788-7>.
- [14] F. FACCHINEI, A. FISCHER, AND V. PICCIALLI, *On generalized Nash games and variational inequalities*, *Oper. Res. Lett.*, 35 (2007), pp. 159–164, <https://doi.org/10.1016/j.orl.2006.03.004>.
- [15] F. FACCHINEI AND C. KANZOW, *Generalized Nash equilibrium problems*, *Ann. Oper. Res.*, 175 (2010), pp. 177–211, <https://doi.org/10.1007/s10479-009-0653-x>.
- [16] F. FACCHINEI AND C. KANZOW, *Penalty methods for the solution of generalized Nash equilibrium problems*, *SIAM J. Optim.*, 20 (2010), pp. 2228–2253, <https://doi.org/10.1137/090749499>.
- [17] K. FAN, *A minimax inequality and applications*, in *Inequalities III: Proceedings of the Third Symposium on Inequalities*, Academic Press, New York, 1972, pp. 103–113.
- [18] A. FISCHER, M. HERRICH, AND K. SCHÖNEFELD, *Generalized Nash equilibrium problems: Recent*

- advances and challenges*, Pesq. Oper., 34 (2014), pp. 521–558.
- [19] A. FRIEDMAN, *Differential Games*, Pure Appl. Math. 25, Wiley-Interscience, New York, 1971.
 - [20] T. L. FRIESZ, *Dynamic Optimization and Differential Games*, Internat. Ser. Oper. Res. Management Sci. 135, Springer, Cham, Switzerland, 2010.
 - [21] M. FUKUSHIMA, *Restricted generalized Nash equilibria and controlled penalty algorithm*, Comput. Manag. Sci., 8 (2011), pp. 201–218, <https://doi.org/10.1007/s10287-009-0097-4>.
 - [22] J. K. HALE, *Ordinary Differential Equations*, 2nd ed., Krieger, Huntington, NY, 1980.
 - [23] A. HAURIE, *Environmental coordination in dynamic oligopolistic markets*, Group Decis. Negot., 4 (1995), pp. 39–57, <https://doi.org/10.1007/BF01384292>.
 - [24] M. HINTERMÜLLER AND T. SUROWIEC, *A PDE-constrained generalized Nash equilibrium problem with pointwise control and state constraints*, Pac. J. Optim., 9 (2013), pp. 251–273.
 - [25] 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 (2015), pp. 1826–1856, <https://doi.org/10.1137/14096829X>.
 - [26] T. ICHISHI, *Game Theory for Economic Analysis*, Econ. Theory Econom. Math. Econ., Academic Press, New York, 1983.
 - [27] K. ITO AND K. KUNISCH, *An augmented Lagrangian technique for variational inequalities*, Appl. Math. Optim., 21 (1990), pp. 223–241, <https://doi.org/10.1007/BF01445164>.
 - [28] K. ITO AND K. KUNISCH, *Lagrange Multiplier Approach to Variational Problems and Applications*, SIAM, Philadelphia, 2008, <https://doi.org/10.1137/1.9780898718614>.
 - [29] A. N. IUSEM, G. KASSAY, AND W. SOSA, *On certain conditions for the existence of solutions of equilibrium problems*, Math. Program., 116 (2009), pp. 259–273, <https://doi.org/10.1007/s10107-007-0125-5>.
 - [30] A. N. IUSEM AND W. SOSA, *New existence results for equilibrium problems*, Nonlinear Anal., 52 (2003), pp. 621–635, [https://doi.org/10.1016/S0362-546X\(02\)00154-2](https://doi.org/10.1016/S0362-546X(02)00154-2).
 - [31] C. KANZOW AND D. STECK, *Augmented Lagrangian methods for the solution of generalized Nash equilibrium problems*, SIAM J. Optim., 26 (2016), pp. 2034–2058, <https://doi.org/10.1137/16M1068256>.
 - [32] C. KANZOW AND D. STECK, *An example comparing the standard and safeguarded augmented Lagrangian methods*, Oper. Res. Lett., 45 (2017), pp. 598–603, <https://doi.org/10.1016/j.orl.2017.09.005>.
 - [33] C. KANZOW, D. STECK, AND D. WACHSMUTH, *An augmented Lagrangian method for optimization problems in Banach spaces*, SIAM J. Control Optim., 56 (2018), pp. 272–291, <https://doi.org/10.1137/16M1107103>.
 - [34] D. KINDERLEHRER AND G. STAMPACCHIA, *An Introduction to Variational Inequalities and Their Applications*, Classics Appl. Math. 31, SIAM, Philadelphia, 2000, <https://doi.org/10.1137/1.9780898719451>.
 - [35] K. KRUMBIEGEL, I. NEITZEL, AND A. RÖSCH, *Regularization for semilinear elliptic optimal control problems with pointwise state and control constraints*, Comput. Optim. Appl., 52 (2012), pp. 181–207, <https://doi.org/10.1007/s10589-010-9357-z>.
 - [36] E. B. LEE AND L. MARKUS, *Foundations of Optimal Control Theory*, Wiley, New York, 1967.
 - [37] A. LOGG, K.-A. MARDAL, AND G. N. WELLS, EDS., *Automated Solution of Differential Equations by the Finite Element Method*, Lect. Notes Comput. Sci. Eng. 84, Springer, 2012, <https://doi.org/10.1007/978-3-642-23099-8>.
 - [38] A. LOGG AND G. N. WELLS, *DOLFIN: Automated Finite Element Computing*, ACM Trans. Math. Software 37, ACM, New York, 2010, <https://doi.org/10.1145/1731022.1731030>.
 - [39] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation. I: Basic Theory*, Grundlehren Math. Wiss. 330, Springer, Cham, Switzerland, 2006.
 - [40] J. MORGAN AND V. SCALZO, *Existence of equilibria in discontinuous abstract economies*, Preprint 53-2004, Dipartimento di Matematica e Applicazioni R. Caccioppoli, Napoli, 2004.
 - [41] H. NIKAIIDO AND K. ISODA, *Note on non-cooperative convex games*, Pacific J. Math., 5 (1955), pp. 807–815, <http://projecteuclid.org/euclid.pjm/1171984836>.
 - [42] J.-S. PANG AND M. FUKUSHIMA, *Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games*, Comput. Manag. Sci., 2 (2005), pp. 21–56, <https://doi.org/10.1007/s10287-004-0010-0>.
 - [43] K. M. RAMACHANDRAN AND C. P. TSOKOS, *Stochastic Differential Games: Theory and Applications*, Atlantis Stud. Probab. Stat. 2, Atlantis Press, Paris, 2012, <https://doi.org/10.2991/978-94-91216-47-3>.
 - [44] A. M. RAMOS, R. GLOWINSKI, AND J. PERIAUX, *Nash equilibria for the multiobjective control of linear partial differential equations*, J. Optim. Theory Appl., 112 (2002), pp. 457–498, <https://doi.org/10.1023/A:1017981514093>.
 - [45] A. M. RAMOS, R. GLOWINSKI, AND J. PERIAUX, *Pointwise control of the Burgers equation and*

- related Nash equilibrium problems: Computational approach*, J. Optim. Theory Appl., 112 (2002), pp. 499–516, <https://doi.org/10.1023/A:1017907930931>.
- [46] A. M. RAMOS AND T. ROUBÍČEK, *Nash equilibria in noncooperative predator-prey games*, Appl. Math. Optim., 56 (2007), pp. 211–241, <https://doi.org/10.1007/s00245-007-0894-5>.
 - [47] R. T. ROCKAFELLAR, *A dual approach to solving nonlinear programming problems by unconstrained optimization*, Math. Program., 5 (1973), pp. 354–373, <https://doi.org/10.1007/BF01580138>.
 - [48] J. B. ROSEN, *Existence and uniqueness of equilibrium points for concave n -person games*, Econometrica, 33 (1965), pp. 520–534, <https://doi.org/10.2307/1911749>.
 - [49] T. ROUBÍČEK, *Noncooperative games with elliptic systems*, in Optimal Control of Partial Differential Equations, Internat. Ser. Numer. Math. 133, Birkhäuser, Basel, 1999, pp. 245–255.
 - [50] T. ROUBÍČEK, *On noncooperative nonlinear differential games*, Kybernetika (Prague), 35 (1999), pp. 487–498.
 - [51] S. P. SETHI AND G. L. THOMPSON, *Optimal Control Theory: Applications to Management Science and Economics*, 2nd ed., Springer, New York, 2000.
 - [52] E. D. SONTAG, *Mathematical Control Theory: Deterministic Finite-Dimensional Systems*, 2nd ed., Texts Appl. Math. 6, Springer, New York, 1998, <https://doi.org/10.1007/978-1-4612-0577-7>.
 - [53] Z. TANG, J.-A. DÉSIDÉRI, AND J. PÉRIAUX, *Multicriterion aerodynamic shape design optimization and inverse problems using control theory and Nash games*, J. Optim. Theory Appl., 135 (2007), pp. 599–622, <https://doi.org/10.1007/s10957-007-9255-4>.
 - [54] F. TRÖLTZSCH, *Optimal Control of Partial Differential Equations*, American Mathematical Society, Providence, RI, 2010, <https://doi.org/10.1090/gsm/112>.
 - [55] J. ZOWE AND S. KURCYUSZ, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim., 5 (1979), pp. 49–62, <https://doi.org/10.1007/BF01442543>.