SIAM J. OPTIM.
Vol. 25, No. 3, pp. 1826–1856
© 2015 Society for Industrial and Applied Mathematics

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

GENERALIZED NASH EQUILIBRIUM PROBLEMS IN BANACH SPACES: THEORY, NIKAIDO–ISODA-BASED PATH-FOLLOWING METHODS, AND APPLICATIONS*

M. HINTERMÜLLER†, T. SUROWIEC‡, AND A. KÄMMLER‡

Abstract. Building upon the results in [M. Hintermüller and T. Surowiec, *Pac. J. Optim.*, 9 (2013), pp. 251–273], a class of noncooperative Nash equilibrium problems is presented, in which the feasible set of each player is perturbed by the decisions of their competitors via a convex constraint. In addition, for every vector of decisions, a common "state" variable is given by the solution of an affine linear equation. The resulting problem is therefore a generalized Nash equilibrium problem (GNEP). The existence of an equilibrium for this problem is demonstrated, and first-order optimality conditions are derived under a constraint qualification. An approximation scheme is proposed, which involves the solution of a parameter-dependent sequence of standard Nash equilibrium problems. An associated path-following strategy based on the Nikaido–Isoda function is then proposed. Function-space-based numerics for parabolic GNEPs and a spot-market model are developed.

Key words. generalized Nash equilibrium problem, GNEP, jointly convex, variational equilibrium, PDE-constrained optimization, nonsmooth Newton methods, multiobjective PDE-constrained optimization, state constraints, spot markets, path-following

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

DOI. 10.1137/14096829X

1. Introduction.

In addition to the many forms of microeconomic models, a variety of design problems in engineering can be modeled by multiobjective optimization problems; see, e.g., [3] (and the references therein) for problems in electromagnetics and [28] for applications in aerodynamics. Due to the difficulty of solving multiobjective problems, some choose to forgo the multiobjective approach completely, whereas others seek to find Pareto optima by using scalarization methods, e.g., by considering a weighted sum of the objectives. This leads to a single optimization problem/cooperative game; see, e.g., [2, pp. 183–184]. In the latter case, one is tasked with solving a large number of potentially complex problems with the intent of finding the Pareto front. This, however, may not be related to the minimization of the individual objective functions. Searching instead for a solution vector such that no individual can deviate from their decision without worsening the objective of at least one of their competitors leads to the concept of a Nash equilibrium.

By extending some recent advances in the optimal control of partial differential equations (PDEs) (PDE-con-strained optimization), see, e.g., [16, 29] and the references therein, the first two authors of this paper developed a method for calculating Nash equilibria for classes of PDE-constrained multiobjective optimization problems

*Received by the editors May 8, 2014; accepted for publication (in revised form) May 7, 2015; published electronically September 17, 2015. The authors would like to acknowledge the financial support from the DFG Research Center MATHEON Project C28, the SPP 1253 Optimization with Partial Differential Equations, the START Project Y 305 Interfaces and Free Boundaries funded by the Austrian Ministry of Science and Education and administered by the Austrian Science Fund FWF, and the Isaac Newton Institute for Mathematical Sciences in Cambridge, UK.
http://www.siam.org/journals/siopt/25-3/96829.html

†Department of Mathematics, Humboldt-Universität zu Berlin, Germany, and Department of Mathematics and Scientific Computing, Karl-Franzens-University of Graz, Austria (hint@math.hu-berlin.de).

‡Department of Mathematics, Humboldt-Universität zu Berlin, Germany (surowiec@math.hu-berlin.de, kaemmler@math.hu-berlin.de).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

with control and state-constraints in [15], where the PDE took the form of a linear elliptic equation. In this paper, we use weaker assumptions and generalize the results in [15] to include a much broader class of infinite-dimensional GNEPs, e.g., parabolic GNEPs. We extend the usage of variational equilibria to our problem class, and we devise a path-following scheme by defining a Nikaido–Isoda-based merit function. In addition, we introduce a type of spot-market model with PDE-constraints.

In the literature, one finds several works concerning PDE-constrained multiobjective problems; see, e.g., [4, 21, 22, 23, 25, 26, 28]. However, none of these papers considers the inclusion of a state constraint. Without such constraints, the models generally reduce to classical noncooperative games in which the feasible sets of the individual players are independent of their competitors' decisions.

Besides the multigrid solver for the discretized problem developed in [4], the only convergent function-space-based method for PDE-constrained multiobjective problems is currently the one developed in [15]. We extend this method to include a class of parabolic GNEPs, and we propose a generalized projected gradient method for GNEPs to solve the spot market problem.

Our path-following approach differs from the more recently developed techniques based on either Moreau–Yosida-type regularization, see, e.g., [12, 13], or Lavrentiev-type regularization, see, e.g., [18], or using interior point methods, see, e.g., [14, 27, 30]. Indeed, these concepts were developed for classical PDE-constrained optimal control problems (with a single objective). Due to the definition of a solution (Nash equilibrium) and the presence of competing objectives, a direct extension to multiobjective Nash equilibrium problems with PDE-constraints is therefore not possible. Since the well-known Nikaido–Isoda function for a Nash equilibrium problem couples these objectives, it is a good candidate for defining a merit function; as such, it serves as the basis for our path-following concept.

The paper is structured as follows. In section 2, we introduce the general framework, a few characteristic examples, and we provide necessary and sufficient first-order optimality conditions in the form of classical KKT conditions as well as with the help of the Nikaido–Isoda function. In section 3, we highlight the concept of variational equilibria, which allows us to focus our efforts toward the calculation of a specific kind of equilibrium. Since the multiplier associated with the state constraint is typically of low regularity, we present in section 4 an approximation method, in which the original problem is approximated by a sequence of (directly solvable) parameter-dependent Nash equilibrium problems. This leads to the concept of paths of equilibria. Under a constraint qualification of Slater-type, boundedness of a path of equilibria is demonstrated and the convergence of (weak) accumulation points to variational equilibria for the original problem is shown. We then build upon these results by developing a path-following strategy. This strategy employs a nonsmooth value function for which a sensitivity result is derived in order to obtain an analytical update formula for the path-parameter. In section 5.1, we extend the method from [15] to the solution of a class of parabolic GNEPs. In section 5.2, inspired by the behavior in natural gas spot markets, we consider a model GNEP, which describes noncooperative economic behavior coupled by complex processes including a viscosity-regularized transport equation. We derive first-order optimality conditions and propose a new numerical method. The algorithmic approaches are illustrated by numerical tests.

2. Preliminary results. In this section, we set the general framework for our investigations. This is followed by a discussion of several problem instances in sec-

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

tion 2.2. Finally, we derive optimality conditions and characterizing relations for equilibria in section 2.3.

2.1. General problem setting. Throughout this paper, we consider an N-player noncooperative game in which the ith player considers the following problem:

$$
(\mathrm{P}_i) \qquad \begin{array}{l} \text{minimize } J_i^1(y) + J_i^2(u_i) \text{ over } (u_i, y) \in U_i \times Y \\ \text{subject to (s.t.)} \\ \qquad Ay = B(u_i, u_{-i}), \\ \qquad u_i \in U_{\mathrm{ad}}^i, \\ \qquad y \in K \end{array}
$$

for $i = 1, \ldots, N$. Note that we use the typical convention $(u_i, u_{-i}) = u \in U$ to emphasize the ith component and denote the components of u with the exception of u_i by the subscript $-i$. For some nonempty subset $C \subset X$, $I_C : X \to \mathbb{R} \cup \{+\infty\}$ represents the standard indicator function. Strong convergence of a sequence is denoted by $\to$, weak-convergence by $\rightharpoonup$, and weak-$*$-convergence by $\overset{*}{\rightharpoonup}$. The closed ε-ball with center x is denoted $\mathbb{B}_\varepsilon(x)$. The superscript $*$ is used to denote the adjoint operator or dual space. For a closed, nonempty, and convex set M in a Banach space Z, we define the classical convex normal cone by

$$
\mathcal{N}_M(z) := \begin{cases} \{z^* \in Z^* \,|\, \langle z^*, z' - z\rangle \le 0 \ \forall z' \in M\}, & z \in M, \\ \emptyset, & \text{else.} \end{cases}
$$

We invoke the following data assumptions throughout.

Assumption 2.1.

1. U_i $(i = 1, \ldots, N)$ are reflexive separable Banach spaces, Y, W are reflexive Banach spaces, and X is a Banach space, each endowed with the norm topology.
2. The embedding $Y \hookrightarrow X$ is continuous.
3. If some arbitrary set $M \subset X^*$ is bounded, then M is weak-$*$ relatively compact in X^* (e.g., X is separable or reflexive cf. [5]).
4. $U := \Pi_{i=1}^N U_i$ is endowed with the usual box topology.
5. A is a linear isomorphism from Y to W.
6. B is a bounded linear operator from U to W, i.e., $B \in \mathcal{L}(U, W)$. We can write $Bu = \sum_{i=1}^m B_i u_i$ with $B_i = B(\cdot, 0_{-i})$ bounded linear operators from U_i to W, i.e., $B_i \in \mathcal{L}(U_i, W)$.
7. $A^{-1}B : U \to X$ is compact.
8. $K \subset X$ is a nonempty, closed, and convex set.
9. The norm topology on X allows for the existence of $x \in K$ and $\varepsilon > 0$ such that $\mathbb{B}_\varepsilon(x) \subset K$.
10. $U_{\mathrm{ad}}^i \subset U_i$ is nonempty, bounded, closed, and convex; and $U_{\mathrm{ad}} := \Pi_{i=1}^N U_{\mathrm{ad}}^i$.
11. There exists a $u \in U_{\mathrm{ad}}$ with $A^{-1}Bu \in K$.
12. $J_i^1 : Y \to \mathbb{R}$ is convex and completely continuous, and $J_i^2 : U_i \to \mathbb{R}$ is strictly convex and continuous. In particular, if $v_k \overset{Y}{\rightharpoonup} v$, then $J_i^1(v_k) \to J_i^1(v)$.

Note that the separability assumption for the objective is not essential for our analysis and that a larger class of objectives could be used provided they satisfy suitable assumptions. Since our examples, which already cover a broad range of applications, are of this form, we prefer to forgo the additional assumptions and abstractions.

The overall noncooperative game will be referred to as (P). Moreover, we often use the solution operator $S : U \to Y$ defined by $Su := A^{-1}(Bu)$. This allows us to

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

also write (P_i) in reduced form:

$$\begin{aligned}&\text{minimize } \mathcal{J}_i(u_i, u_{-i}) := J_i^1(S(u_i, u_{-i})) + J_i^2(u_i) \text{ over } u_i \in U_i\\ &\text{s.t.}\\ &\qquad u_i \in U_{\mathrm{ad}}^i, \quad S(u_i, u_{-i}) \in K.\end{aligned}$$

Due to the dependence on the decision vector u_{-i} in the feasible set of (P_i), this type of game is often referred to as a *generalized Nash equilibrium problem* (GNEP). The current setting allows us to consider a wide array of problems, including GNEPs in which the feasible set is governed by a PDE.

Note that for $u \in U$, the strategy u_i is feasible for (P_i), given u_{-i}, for all $i = 1, \ldots, N$ if and only if $u \in C$, where

$$C := \{u \in U_{\mathrm{ad}} \mid Su \in K\}.$$

Since C is convex, our problem has the structure of a so-called *jointly convex* GNEP. We now state our desired solution concept.

DEFINITION 2.2. *A point $\bar{u} \in C$ is referred to as a* Nash equilibrium *provided the following inequality holds:*

$$\mathcal{J}_i(\bar{u}_i, \bar{u}_{-i}) \leq \mathcal{J}_i(v_i, \bar{u}_{-i}) \ \forall v_i \in U_i : (v_i, \bar{u}_{-i}) \in C \ \forall i = 1, \ldots, N. \tag{1}$$

2.2. Examples. As stated above, our rather general framework allows for applications of GNEPs with PDEs. For illustration purposes, we first highlight a GNEP governed by an elliptic state equation and then one with a parabolic state equation. Note that a.e. Ω means "almost everywhere on Ω with respect to the Lebesgue measure."

Example 2.3. Consider a GNEP in which (P_i) is defined by

$$\begin{aligned}&\text{minimize } \tfrac{1}{2}||K^i y - y_d^i||^2_{L^2(\Omega)} + \tfrac{\alpha_i}{2}||u_i||^2_{L^2(\Omega)} \text{ over } (u_i, y) \in L^2(\Omega) \times W_0^{1,r}(\Omega)\\ &\text{s.t.}\\ &\qquad -\Delta y = \textstyle\sum_{i=1}^N \chi_{\Omega_i} u_i \text{ in } W^{-1,r}(\Omega),\\ &\qquad a_i \leq u_i \leq b_i, \text{ a.e. } \Omega,\\ &\qquad y \leq \psi \text{ in } \Omega.\end{aligned}$$

Here, we have the following:

- $\Omega \subset \mathbb{R}^d, d \in \{1, 2, 3\}$, is an open bounded subset, and $\partial\Omega$ is sufficiently regular.
- $U_i = L^2(\Omega)$, $Y = W_0^{1,r}(\Omega)$, $W = W^{-1,r}(\Omega)$, $r > \max(d, 2)$, $X = C(\overline{\Omega})$.
- $A = -\Delta$, $Bu = \sum_{i=1}^N \chi_{\Omega_i} u_i$, where $\Omega_i \subset \Omega$ for each i, and $\chi_{\Omega_i} : \mathbb{R}^d \to \{0, 1\}$ is the standard characteristic function.
- $a_i, b_i \in L^2(\Omega)$, with $a_i \leq b_i$, a.e. Ω.
- $y_d^i \in L^2(\Omega)$, $\alpha_i > 0$, $\psi \in W_0^{1,r}(\Omega)$, $\psi|_{\partial\Omega} > 0$.
- $U_{\mathrm{ad}}^i := \{v \in L^2(\Omega) \,|\, a_i \leq v \leq b_i, \text{ a.e. } \Omega\}$.
- $K := \psi + C(\overline{\Omega})_-$, where $C(\overline{\Omega})_-$ is the cone of nonpositive continuous functions.
- $J_i^1(y) := \frac{1}{2}||K^i y - y_d^i||^2_{L^2(\Omega)}$ with $K^i \in \mathcal{L}(Y, L^2(\Omega))$ for each i, and $J_i^2(u_i) := \frac{\alpha_i}{2}||u_i||^2_{L^2(\Omega)}$.

Here, $L^2(\Omega)$ and $W_0^{1,r}(\Omega)$ denote the standard Lebesgue and Sobolev spaces; see, e.g., [1]. Furthermore, $W^{-1,r}(\Omega)$ represents the topological dual space of $W_0^{1,s}(\Omega)$ with $r^{-1} + s^{-1} = 1$ for $r \in [1, +\infty]$. GNEPs of this type were considered recently in [15].

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Another interesting class of examples involves the optimal control of a linear parabolic PDE with control and state constraints, which we write in reduced form (as opposed to the form used in Example 2.3).

Example 2.4. Let $y = Su$ for $u \in L^2(0,T;L^2(\Omega))^N$ solve the initial boundary value problem

$$\begin{aligned} y_t - \Delta y + c_0 y &= Bu && \text{in } Q = \Omega \times (0,T),\\ y &= 0 && \text{on } \Sigma = \Gamma \times (0,T),\\ y(\cdot,0) &= 0 && \text{in } \Omega, \end{aligned}$$

where $\Omega \subset \mathbb{R}^d$, $d \in \{1,2,3\}$, is open and bounded with Lipschitz boundary Γ, $T > 0$ (a scalar), $Q := \Omega \times (0,T)$, $\Sigma := \Gamma \times (0,T)$, $c_0 \in L^\infty(Q)$ with $c_0 \geq 0$ a.e. Q. We consider a GNEP in which (P_i) is defined as

$$\begin{aligned} &\text{minimize } \tfrac{1}{2}||K^i S(u_i,u_{-i}) - y_d^i||^2_{L^2(Q)} + \tfrac{\alpha_i}{2}||u_i||^2_{L^2(Q)} \text{ over } u_i \in L^2(Q)\\ &\text{s.t.}\\ &\qquad a_i \leq u_i \leq b_i, \text{ a.e. } Q, \quad S(u_i,u_{-i}) \leq \psi \text{ in } Q. \end{aligned} \tag{2}$$

Here, we consider the settings:

- $U_i = L^2(Q)$, $Y = \{y \in L^2(0,T;H_0^1(\Omega) \cap H^2(\Omega)) \,|\, y_t \in L^2(Q)\}$, $W = L^2(Q) \times H_0^1(\Omega)$, $X = C(\overline{Q})$.
- $Ay = (y_t - \Delta y + c_0 y, y(0))^\top$, $B \in \mathcal{L}(L^2(Q))$.
- $\psi \in Y$, $\psi|_{\partial\Omega} > 0$.
- $a_i, b_i \in L^2(Q)$, $a_i \leq b_i$, a.e. Q.
- $y_d^i \in L^2(Q)$, $\alpha_i > 0$.
- $U_{\mathrm{ad}}^i := \{v \in L^2(Q) \,|\, a_i \leq v \leq b_i, \text{ a.e. } Q\}$.
- $K = \psi + C(Q)_-$, where $C(Q)_-$ is the cone of nonpositive continuous functions on Q.
- $\mathcal{J}_i(u) = \frac{1}{2}||K^i S(u_i,u_{-i}) - y_d^i||^2_{L^2(L^2)} + \frac{\alpha_i}{2}||u_i||^2_{L^2(L^2)}$ with $K^i \in \mathcal{L}(Y, L^2(Q))$ for each i.

Here, $L^2(0,T;X)$ denotes the standard Bochner space of square integrable functions with values in a Banach space X; see, e.g., [9].

2.3. Optimality conditions for generalized Nash equilibria. The question of existence of a Nash equilibrium is postponed to section 3, where we concentrate on a specific class of equilibria that can be computed numerically. Nevertheless, in the interest of contrasting the difference between variational equilibria and all other equilibria, we provide the following results.

THEOREM 2.5 (first-order conditions for Nash equilibria). *If a Nash equilibrium $\bar{u} \in U$ of (P) satisfies*

$$\forall i = 1,\dots,N, \exists u_i \in U_{\mathrm{ad}}^i : \mathbb{B}_\varepsilon(0) \subset S(u_i,\bar{u}_{-i}) - K \tag{3}$$

for some $\varepsilon > 0$, then there exists $\bar{y} \in Y$, $\bar{p} \in (W^*)^N$, $\bar{\lambda} \in U^*$, and $\bar{\mu} \in (X^*)^N$ such that the first-order system

$$(\mathrm{OS}_i) \quad \begin{cases} \bar{y} = S\bar{u},\\ -\bar{p}_i \in A^{-*}\big(\partial J_i^1(\bar{y}) + \bar{\mu}_i\big),\\ \bar{\lambda}_i \in \partial I_{U_{\mathrm{ad}}^i}(\bar{u}_i),\\ \bar{\mu}_i \in \partial I_K(\bar{y}),\\ 0 \in \partial J_i^2(\bar{u}_i) - B_i^* \bar{p}_i + \bar{\lambda}_i \end{cases}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

is fulfilled for $i = 1, \dots, N$. The coupled system is denoted by (OS). *Conversely, if the tuple $(\bar{u}, \bar{y}, \bar{p}, \bar{\lambda}, \bar{\mu}) \in U \times Y \times (W^*)^N \times U^* \times (X^*)^N$ satisfies the coupled system* (OS), *then $\bar{u}$ is a Nash equilibrium.*

The constraint qualification (3) requires the existence of a Slater point for each individual problem (P_i) at the solution. Later, we will work only with variational equilibria, where only a Slater point for the entire problem is required (see section 3).

Proof. By definition, $\bar{u}$ is a Nash equilibrium if and only if $\bar{u}_i$ solves

$$\text{minimize}\, \mathcal{J}_i(u_i, \bar{u}_{-i}) + I_{U^i_{\mathrm{ad}}}(u_i) + I_K(S(u_i, \bar{u}_{-i})) \text{ over } u_i \in U_i.$$

By convexity, this is equivalent to

$$0 \in \partial \left(\mathcal{J}_i(\cdot, \bar{u}_{-i}) + I_{U^i_{\mathrm{ad}}}(\cdot) + I_K(S(\cdot, \bar{u}_{-i})) \right)(\bar{u}_i), \tag{4}$$

where ∂ denotes the standard subdifferential from convex analysis; cf. [17]. Note that by (3), $u_i \mapsto I_K(S(u_i, \bar{u}_{-i}))$ has a continuity point. Since $\mathcal{J}_i$ is continuous, we can apply the sum rule for convex subdifferentials, cf., e.g., [17, Chapter 4, section 4.2, Theorem 1], and obtain

$$\begin{aligned} &\partial \left(\mathcal{J}_i(\cdot, \bar{u}_{-i}) + I_{U^i_{\mathrm{ad}}}(\cdot) + I_K(S(\cdot, \bar{u}_{-i})) \right)(\bar{u}_i) \\ &\qquad = \partial \mathcal{J}_i(\cdot, \bar{u}_{-i})(\bar{u}_i) + \mathcal{N}_{U^i_{\mathrm{ad}}}(\bar{u}_i) + \partial I_K(S(\cdot, \bar{u}_{-i}))(\bar{u}_i). \end{aligned}$$

By continuity of J_i^1, we can apply the chain rule for convex subdifferentials, cf., e.g., [17, Chapter 4, section 4.2, Theorem 2] and obtain

$$\begin{aligned} \partial \mathcal{J}_i(\cdot, \bar{u}_{-i})(\bar{u}_i) &= B_i^* A^{-*} \partial J_i^1(S\bar{u}) + \partial J_i^2(\bar{u}_i), \\ \partial I_K(S(\cdot, \bar{u}_{-i}))(\bar{u}_i) &= B_i^* A^{-*} \mathcal{N}_K(S\bar{u}). \end{aligned}$$

Thus, (4), can be written as

$$0 \in B_i^* A^{-*} \partial J_i^1(S\bar{u}) + \partial J_i^2(\bar{u}_i) + \mathcal{N}_{U^i_{\mathrm{ad}}}(\bar{u}_i) + B_i^* A^{-*} \mathcal{N}_K(S\bar{u}). \qquad \square$$

Next, we introduce the so-called Nikaido–Isoda function $\Psi : U \times U \to \mathbb{R}$ defined by

$$\Psi(u, v) := \sum_{i=1}^{N} \left[\mathcal{J}_i(u_i, u_{-i}) - \mathcal{J}_i(v_i, u_{-i}) \right].$$

In addition, define $V : C \to \mathbb{R}$ by

$$V(u) = \max_v \{ \Psi(u, v) \mid v \in U : (v_i, u_{-i}) \in C \text{ for } i = 1, \dots, N \}.$$

Choosing $v = u$, we infer that $V(u) \geq \Psi(u, u) = 0$ for $u \in C$. The function V plays an essential role throughout the text. Moreover, it leads to the following convenient characterization of Nash equilibria.

LEMMA 2.6. *A point $\bar{u} \in U$ is a Nash equilibrium of (P) if and only if $\bar{u} \in C$ and $V(\bar{u}) = 0$.*

Proof. This is essentially Lemma 3.1 in [20]. $\square$

3. Variational equilibria. As mentioned above, we focus our attention on a specific class of equilibria that can be more readily computed as compared to the computation of Nash equilibria via the solution of the coupled system (OS). In essence, this class involves those equilibria for which $\mu_i = \mu_j$ for all $i, j \in \{1, \dots, N\}$ in Theorem 2.5.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

3.1. Essential results and observations. Since (P) is a jointly convex GNEP, we can use the more restrictive solution concept of variational equilibria, which was first introduced in [24] (as normalized equilibria in a slightly different form) and is nowadays widely used for (finite-dimensional) games; see, e.g., the survey paper [6]. We note that the proofs of Theorems 3.2 and 3.9, and Lemma 3.3 can be directly adapted from their finite-dimensional analogues; see in particular, Theorems 3.2 and 3.5 in [6]. As such, we present them without proof here.

Our definition of variational equilibria is based on the Nikaido–Isoda function Ψ defined above and the functions $\widehat{\mathcal{R}} : C \to C$ defined by

$$\widehat{\mathcal{R}}(u) := \operatorname*{argmax}_v \{\Psi(u,v) \mid v \in C\} = \operatorname*{argmin}_v \left\{ \sum_{i=1}^N \mathcal{J}_i(v_i, u_{-i}) \mid v \in C \right\}$$

and $\widehat{V} : C \to \mathbb{R}$ defined by

$$\widehat{V}(u) := \Psi(u, \widehat{\mathcal{R}}(u)) = \max_v \{\Psi(u,v) \mid v \in C\}.$$

Note that $\widehat{\mathcal{R}}$ and thus $\widehat{V}$ are well defined since C is nonempty, bounded, closed, and convex, and $v \mapsto \Psi(u,v)$ is continuous and strictly concave (thus upper semicontinuous) for fixed $u \in U$. Note that the parameter u in the definition of V perturbs both the objective and feasible set; in $\widehat{V}$, u only perturbs the objective. Obviously, we have $\widehat{V}(u) \geq \Psi(u,u) = 0$ for all $u \in C$.

DEFINITION 3.1. *A point $\bar{u} \in U$ is called a* variational equilibrium *of (P) if $\bar{u} \in C$ and $\widehat{V}(\bar{u}) = 0$.*

Note that variational equilibria and Nash equilibria are the same for an NEP (e.g., in our setting $K = Y$) by Lemma 2.6. As in these cases, there is no difference between $\widehat{V}$ and V. For our general jointly convex GNEP, we have the following relation.

THEOREM 3.2 (variational equilibria are Nash equilibria). *Every variational equilibrium of (P) is also a Nash equilibrium of (P).*

The following lemma characterizes variational equilibria as fixed points of $\widehat{\mathcal{R}}$.

LEMMA 3.3. *A point $\bar{u} \in C$ is a variational equilibrium if and only if $\bar{u} = \widehat{\mathcal{R}}(\bar{u})$.*

THEOREM 3.4 (existence of a variational equilibrium). *The GNEP (P) admits a variational equilibrium $\bar{u} \in U$.*

Proof. By Lemma 3.3, it suffices to show that $\widehat{\mathcal{R}} : C \to C$ has a fixed point. To apply the fixed point theorem of Kakutani, we take $\widehat{\mathcal{R}}$ as a set-valued mapping. Let $Z := U$ be endowed with the weak topology. Then $C \subseteq Z$ is nonempty, convex, and compact. Clearly, $\widehat{\mathcal{R}}(u)$ is nonempty, convex, and closed for all $u \in C$. To show sequential upper semicontinuity, suppose that $(u^k)_{k\in\mathbb{N}} \subseteq C$ such that $u^k \xrightarrow{Z} u^*$, and let $v^k \in \widehat{\mathcal{R}}(u^k)$, such that $v^k \xrightarrow{Z} v^*$. Since J_i^1, J_i^2 are convex and continuous, they are lower semicontinuous on Z. Moreover, $z^k \in U$ with $z^k \xrightarrow{Z} z$ implies $S(z^k) \xrightarrow{Y} S(z)$. It follows then by the assumptions on J_i^1 that for all $v \in C$,

$$\begin{aligned}\sum_{i=1}^N \mathcal{J}_i(v_i^*, u_{-i}^*) &\leq \liminf_{k\to\infty} \sum_{i=1}^N J_i^1(S(v_i^k, u_{-i}^k)) + J_i^2(v_i^k) \\ &\leq \liminf_{k\to\infty} \sum_{i=1}^N J_i^1(S(v_i, u_{-i}^k)) + J_i^2(v_i) = \sum_{i=1}^N \mathcal{J}_i(v_i, u_{-i}^*).\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

This implies $v^* \in \widehat{\mathcal{R}}(u^*)$. By the separability of Z, the weak topology on C is metrizable, and thus we have shown the sequential upper semicontinuity of $\widehat{\mathcal{R}}$. By Kakutani's fixed point theorem, there exists a fixed point of $\widehat{\mathcal{R}}$, as was to be shown. ∎

Remark 3.5. Note that the strict convexity of J_i^2 is not used in the previous proof. This allows us to argue in section 5.2 that the spot market model has a variational equilibrium.

From Theorems 3.2 and 3.4, one directly infers the following.

COROLLARY 3.6 (existence of Nash equilibria). *The GNEP (P) admits a Nash equilibrium $\bar{u} \in U$.*

By studying the first-order system for variational equilibria, we can get a better understanding of the relation to Nash equilibria. In what follows, we will need some constraint qualification of Slater type: We will say that the GNEP (P) satisfies the *Slater condition* if

$$0 \in \operatorname{int}\left(S(U_{\mathrm{ad}}) - K\right),$$

where the interior is taken in the space X. Note that this constraint qualification is much weaker than the one used for the characterization of Nash equilibria in Theorem 2.5.

THEOREM 3.7 (first-order conditions for variational equilibria). *Suppose that the Slater condition is satisfied. Then a point $\bar{u} \in U$ is a variational equilibrium of (P) if and only if there exists $\bar{y} \in Y$, $\bar{p} \in (W^*)^N$, $\bar{\lambda} \in U^*$, and $\bar{\mu} \in X^*$ such that the first-order system*

$$(\widehat{\mathrm{OS}}_i) \quad \begin{cases} \bar{y} = S\bar{u}, \\ -\bar{p}_i \in A^{-*}\big(\partial J_i^1(\bar{y}) + \bar{\mu}\big), \\ \bar{\lambda}_i \in \partial I_{U_{\mathrm{ad}}^i}(\bar{u}_i), \\ \bar{\mu} \in \partial I_K(\bar{y}), \\ 0 \in \partial J_i^2(\bar{u}_i) - B_i^* \bar{p}_i + \bar{\lambda}_i \end{cases}$$

is fulfilled for each $i = 1, \dots, N$. The coupled system is referred to by $(\widehat{\mathrm{OS}})$.

Remark 3.8. A comparison with Theorem 2.5 shows that (under a constraint qualification) a Nash equilibrium $\bar{u}$ is a variational equilibrium if and only if for the associated multipliers $\bar{\mu}_i \in X^*$, we have $\bar{\mu}_1 = \cdots = \bar{\mu}_N$. This observation links variational equilibria to the normalized equilibria as defined by Rosen [24], where the $\bar{\mu}_i$ are only allowed to differ by a scalar factor $r_i > 0$.

Proof. By Lemma 3.3, $\bar{u}$ is a variational equilibrium if and only if $\bar{u} = \widehat{\mathcal{R}}(\bar{u})$, i.e., $\bar{u}$ solves

$$\text{minimize } \sum_{i=1}^N \mathcal{J}_i(v_i, \bar{u}_{-i}) + I_C(v) \text{ over } v \in U.$$

Due to convexity, this is equivalent to $0 \in \partial(\sum_{i=1}^N \mathcal{J}_i((\cdot)_i, \bar{u}_{-i}) + I_C(\cdot))(\bar{u})$. Since the $\mathcal{J}_i$, $i = 1, \dots, N$, are continuous, we can apply the sum rule for convex subdifferentials, cf., e.g., [17, Chapter 4, section 4.2, Theorem 1], and obtain

$$\begin{aligned} \partial\left(\sum_{i=1}^N \mathcal{J}_i((\cdot)_i, \bar{u}_{-i}) + I_C(\cdot)\right)(\bar{u}) &= \sum_{i=1}^N \partial\left(\mathcal{J}_i((\cdot)_i, \bar{u}_{-i})\right)(\bar{u}) + \mathcal{N}_C(\bar{u}) \\ &= \prod_{i=1}^N (\partial_i \mathcal{J}_i(\cdot, \bar{u}_{-i}))(\bar{u}_i) + \mathcal{N}_C(\bar{u}), \end{aligned} \tag{5}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where ∂_i denotes the convex subdifferential with respect to the ith variable. Indeed, $u^* \in \partial(\mathcal{J}_i((\cdot)_i, \bar{u}_{-i}))(\bar{u})$ implies

$$\mathcal{J}_i(u_i', \bar{u}_{-i}) \geq \mathcal{J}_i(\bar{u}) + \langle u^*, u' - \bar{u}\rangle_{U^*,U} \ \forall u' \in U. \tag{6}$$

Then for $u' = (\bar{u}_i, \bar{u}_{-i} + v) \in U$ with $v \in U_{-i}$ arbitrary, we have

$$0 \geq \langle u^*, (0_i, v)\rangle_{U^*,U} = \langle u^*_{-i}, v\rangle_{U^*_{-i},U_{-i}} \ \forall v \in U_{-i}.$$

Hence, $u^*_{-i} = 0_{-i}$. But then (6) yields

$$\mathcal{J}_i(u_i', \bar{u}_{-i}) \geq \mathcal{J}_i(\bar{u}) + \langle u_i^*, u_i' - \bar{u}_i\rangle_{U_i^*,U_i} \ \forall u_i' \in U_i.$$

Therefore, $u_i^* \in \partial_i \mathcal{J}_i(\cdot, \bar{u}_{-i})(\bar{u}_i)$. This yields $u^* = (u_i^*, 0_{-i})$. The converse follows by substitution, in which case $u^* \in \partial\left(\mathcal{J}_i((\cdot)_i, \bar{u}_{-i})\right)(\bar{u})$ if and only if $u^* = (u_i^*, 0_{-i})$ with $u_i^* \in \partial_i \mathcal{J}_i(\cdot, \bar{u}_{-i})(\bar{u}_i)$. In order to see that the (Minkowski) sum of the subdifferentials equals the product in (5), take $u^* \in \sum_{i=1}^N \partial\left(\mathcal{J}_i((\cdot)_i, \bar{u}_{-i})\right)(\bar{u})$. By the previous argument, there exists for each $i = 1, \ldots, N$ some $u_i^* \in \partial_i \mathcal{J}_i(\cdot, \bar{u}_{-i})(\bar{u}_i)$ such that $u^* = \sum_{i=1}^N (u_i^*, 0_{-i}) = (u_1^*, \ldots, u_N^*) \in \prod_{i=1}^N (\partial_i \mathcal{J}_i(\cdot, \bar{u}_{-i}))(\bar{u}_i)$. By simply reversing this argument we obtain equality.

By the continuity of J_i^1, we can apply the chain rule for convex subdifferentials, cf., e.g., [17, Chapter 4, section 4.2, Theorem 2], and obtain

$$(\partial_i \mathcal{J}_i(\cdot, \bar{u}_{-i}))(\bar{u}_i) = B_i^* A^{-*} \partial J_i^1(S\bar{u}) + \partial J_i^2(\bar{u}_i).$$

Given $C = U_{\mathrm{ad}} \cap S^{-1}(K) := \{u \in U_{\mathrm{ad}} \mid S(u) \in K\}$, the constraint qualification $0 \in \operatorname{int}\big(S(U_{\mathrm{ad}}) - K\big)$ allows us to apply the chain rule for convex subdifferentials to obtain

$$\mathcal{N}_C(u) = \mathcal{N}_{U_{\mathrm{ad}}}(u) + S^* \mathcal{N}_K(Su) \quad \forall u \in U,$$

where we recall that $S := A^{-1}B$ is a bounded linear operator, and hence S^* is well defined. Moreover, cf., e.g., [2, section 4.6],

$$\mathcal{N}_{U_{\mathrm{ad}}}(u) = \mathcal{N}_{\prod_{i=1}^N U_{\mathrm{ad}}^i}(u) = \prod_{i=1}^N \mathcal{N}_{U_{\mathrm{ad}}^i}(u_i) \quad \forall u \in U.$$

Hence, we may write every $\nu \in \mathcal{N}_C(u)$ as $\nu = \lambda + S^*\mu$ with $\lambda = (\lambda_i)_{i=1,\ldots,N}$, $\lambda_i \in \mathcal{N}_{U_{\mathrm{ad}}^i}(u_i)$, and $\mu \in \mathcal{N}_K(Su)$. Inserting this into (5) and reading it componentwise, we get the equivalent formulation

$$0 \in B_i^* A^{-*} \partial J_i^1(S\bar{u}) + \partial J_i^2(\bar{u}_i) + \bar{\lambda}_i + B_i^* A^{-*} \bar{\mu} \quad \forall i = 1, \ldots, N$$

with $\bar{\lambda}_i \in \mathcal{N}_{U_{\mathrm{ad}}^i}(u_i)$ and $\bar{\mu} \in \mathcal{N}_K(S\bar{u})$. This yields the assertion. ∎

If the objectives are continuously Gâteaux-differentiable, i.e., the directional derivatives are continuous and linear in the direction, then variational equilibria can be characterized as solutions of an associated variational inequality. Frequently, this characterization is even taken as the definition of variational equilibria. For the remainder of this section, we assume that J_i^1, J_i^2, $i = 1, \ldots, N$, are continuously Gâteaux-differentiable. Define $F : U \to U^*$ via

$$F_i(u) := B_i^* A^{-*} (J_i^1)'(Su) + (J_i^2)'(u_i), \quad i = 1, \ldots, N,$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

and consider the variational inequality VI(C,F), i.e., the problem of finding $\bar{u} \in C$ with

$$\langle F(\bar{u}), u - \bar{u}\rangle_{U^*,U} \geq 0 \quad \forall u \in C. \tag{7}$$

We call VI(C,F) the *associated variational inequality* to (P).

THEOREM 3.9 (variational equilibria as solutions of the associated VI). *A point $\bar{u} \in U$ is a variational equilibrium of (P) if and only if it is a solution of the associated variational inequality.*

3.2. A special case: Reduction to a single control problem. We consider a situation in which the state-dependent part of the players' objectives differ only by a linear-affine functional. In such a case, the task of calculating a variational equilibrium for the GNEP can be reduced to a more tractable optimization problem with a single objective.

For the remainder of this section, we replace the assumptions on J_i^1 by

- $J_i^1 = J_0^1 + \tilde{J}_i^1$, where J_0^1 is convex and continuously Gâteaux differentiable, and $\tilde{J}_i^1$ is linear-affine; without loss of generality, we assume that $\tilde{J}_i^1 \in Y^*$.

This especially includes typical tracking-type functionals of the form

$$J_i^1(y) = \frac{1}{2}\|y - y_i^d\|_Y^2,$$

since

$$\frac{1}{2}\|y - y_i^d\|_{L^2}^2 = \frac{1}{2}\|y\|_{L^2}^2 - (y, y_d^i)_{L^2} + \frac{1}{2}||y_d^i||_{L^2}^2.$$

PROPOSITION 3.10. *Under the above assumption, there exists a unique variational equilibrium $\bar{u}$ of (P), which is the unique solution of the convex optimization problem*

$$\textit{minimize } \widehat{J}(u) := J_0^1(Su) + \sum_{i=1}^{N} \left(J_i^2(u_i) + \langle S_i^* \tilde{J}_i^1, u_i\rangle_{U_i^*,U_i}\right) \textit{ over } u \in U.$$

$$\textit{s.t. } u \in C.$$

Proof. Since C is convex and $\widehat{J}$ is convex and continuously Gâteaux differentiable, $\tilde{u} \in C$ is a global minimizer of $\widehat{J}$ if and only if

$$\langle \widehat{J}'(\tilde{u}), u - \tilde{u}\rangle_{U^*,U} \geq 0 \qquad \forall u \in C.$$

Moreover, since the components of $\widehat{J}'$ coincide with those of F, we see that $\tilde{u}$ is a global minimizer if and only if $\tilde{u}$ solves (7). It follows then from Theorem 3.9 that $\tilde{u}$ is a variational equilibrium. The uniqueness follows from the strict convexity of $\widehat{J}$. ∎

Thus, we have reduced the problem of finding a variational equilibrium to the (theoretically and numerically) better understood solution of an optimization problem with a single objective.

Finally, we remark that since a continuously Gâteaux differentiable monotone operator $\Phi : X \to X^*$ on a Banach space X is a potential operator if and only if $\Phi'(x)$ is self-adjoint for all $x \in X$, our reduction procedure is valid if and only if it has the special structure considered in this subsection and J_i^1, J_i^2 are twice continuously Gâteaux differentiable.

Remark 3.11 (relation to potential games). Given some $(u_i, u_{-i}), (v, u_{-i}) \in C$, we see that $\widehat{\mathcal{J}}(u) - \widehat{\mathcal{J}}(v, u_{-i}) = J_0^1(Su) + \langle S_i^* \tilde{J}_i^1, u_i\rangle_{U_i^*,U_i} + J_i^2(u_i) - (J_0^1(S(v, u_{-i})) +$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$\langle S_i^* \tilde{J}_i^1, v\rangle_{U_i^*,U_i} + J_i^2(v))$. Hence, $\widehat{\mathcal{J}}(u) - \widehat{\mathcal{J}}(v, u_{-i}) > 0$ if and only if $J_0^1(Su) + \langle S_i^* \tilde{J}_i^1, u_i\rangle_{U_i^*,U_i} + J_i^2(u_i) - (J_0^1(S(v, u_{-i})) + \langle S_i^* \tilde{J}_i^1, v\rangle_{U_i^*,U_i} + J_i^2(v)) > 0$. Therefore, $\widehat{\mathcal{J}}$ satisfies the conditions of an exact potential; cf. [19] or see [7, Definition 2.1(b)]. Moreover, the construction of C here fits into the generalization of potential games found in [7, Definition 2.1(a)]. This implies that the variational equilibrium problem is equivalent to a generalized potential game.

4. Calculating variational equilibria. In most applications, both J_i^1 and J_i^2 will be at least continuously Gâteaux differentiable, whereas U_{ad}^i is typically composed of box constraints on the various controls. However, as the state constraint $Su \in K$ must be understood in X, often an infinite-dimensional function space whose topology allows pointwise constraint systems to have nonempty interiors, the associated multiplier is often only a type of measure, e.g., as in Examples 2.3 and 2.4 above. As a consequence, the complementarity condition associated with $\bar{\mu} \in \partial I_K(\bar{y})$ cannot be rewritten in a way amenable to efficient numerical solvers. Thus, although we have both existence of an equilibrium and optimality conditions, we still need to address the issue of developing a robust (mesh-independent) numerical solution method.

4.1. Approximating the GNEP. Since the state constraint is the source of many difficulties, we replace it by a parameter-dependent penalty term in the individual objectives. In our setting, this leads coincidentally to a classical NEP. Solving the NEP requires the solution of a variational inequality of the type

$$\langle F_\gamma(\tilde{u}), u - \tilde{u}\rangle \geq 0 \ \forall u \in U_{\text{ad}}.$$

Though U_{ad} may be a "nice" set with concrete structure, $F_\gamma(\tilde{u})$ is not necessarily strongly monotone. In fact, there is no reason why F_γ should even be monotone. This greatly complicates numerical efforts. Throughout the coming discussion, we invoke the following additional assumptions.

Assumption 4.1.

1. J_i^1 and J_i^2, $i = 1, \ldots, N$, are continuously Gâteaux differentiable with bounded Gâteaux derivatives $(J_i^1)' : Y \to Y^*$ and $(J_i^2)' : U_i \to U_i^*$, i.e., they map bounded sets into bounded sets.
2. $(J_i^1)'$ is completely continuous, i.e., $(J_i^1)'(y^n) \overset{Y^*}{\to} (J_i^1)'(y)$ whenever $y^n \overset{Y}{\rightharpoonup} y$.
3. $(J_i^2)'$ is weakly continuous, i.e., $(J_i^2)'(u_i^n) \overset{U_i^*}{\rightharpoonup} (J_i^2)'(u_i)$ whenever $u_i^n \overset{U_i}{\rightharpoonup} u_i$.
4. $(J_i^2)'$ has the property $\langle (J_i^2)'(u_i^*), u_i^*\rangle_{U_i^*,U_i} \leq \limsup_{n\to\infty} \langle (J_i^2)'(u_i^n), u_i^n\rangle_{U_i^*,U_i}$ whenever $u_i^n \overset{U_i}{\rightharpoonup} u_i$.
5. $\beta : X \to \mathbb{R}_+$ is convex, continuous, and continuously Gâteaux-differentiable with $\ker \beta = K$, i.e., $\beta(y) = 0$ whenever $y \in K$, otherwise $\beta(y) > 0$.

These assumptions on J_i^1 and J_i^2 are satisfied for the objectives in Examples 2.3 and 2.4. Also note that the assumptions on β are satisfied when β (multiplied by a positive scalar γ) arises from the Moreau–Yosida regularization of the indicator function I_K.

For each $\gamma > 0$ and $i = 1, \ldots, N$, we approximate (P_i) by

$$(\mathrm{P}_{i,\gamma}) \qquad \begin{aligned} &\text{minimize } J_i^1(y) + J_i^2(u_i) + \gamma\beta(y) \text{ over } (u_i, y) \in U_i \times Y \\ &\text{s.t.} \\ &\qquad Ay = B(u_i, u_{-i}), \\ &\qquad u_i \in U_{\text{ad}}^i. \end{aligned}$$

We write (P_γ) to refer to the entire relaxed Nash equilibrium problem. Using almost

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

identical arguments to those which were applied for (P), one demonstrates the existence of an equilibrium for every $\gamma > 0$. Moreover, the first-order necessary and sufficient optimality conditions are of the following form: For all $i = 1, \dots, N$, u^γ is a Nash equilibrium if and only if there exist $y^\gamma \in Y$, $p^\gamma \in (W^*)^N$, $\lambda^\gamma \in U^*$, and $\mu^\gamma \in X^*$ such that

$$(\mathrm{OS}_{i,\gamma}) \quad \begin{cases} y^\gamma = Su^\gamma, \\ -p_i^\gamma = A^{-*}\big((J_i^1)'(y^\gamma) + \mu^\gamma\big), \\ \lambda_i^\gamma \in \partial I_{U_{\mathrm{ad}}^i}(u_i^\gamma), \\ \mu^\gamma = \gamma\beta'(y^\gamma), \\ 0 = (J_i^2)'(u_i^\gamma) - B_i^* p_i^\gamma + \lambda_i^\gamma. \end{cases}$$

The coupled system is denoted by (OS_γ). Note that no constraint qualification is required to derive the optimality conditions.

4.2. The primal-dual path and its properties. Our aim now is to show the convergence of stationary points of OS_γ to a solution of (P). Since (P_γ) does not have a unique solution in general, we have to consider boundedness of selections of paths of stationary points of (P_γ).

DEFINITION 4.2. *For $\gamma > 0$, denote by $\mathcal{S}_\gamma \subseteq U \times Y \times (W^*)^N \times U^* \times X^*$ the set of solutions of (OS_γ) and set*

$$\mathscr{C} := \Big\{ \big((u^\gamma, y^\gamma, p^\gamma, \lambda^\gamma, \mu^\gamma)\big)_{\gamma>0} \mid \forall \gamma > 0 : (u^\gamma, y^\gamma, p^\gamma, \lambda^\gamma, \mu^\gamma) \in \mathcal{S}_\gamma \Big\}.$$

We call every element $\mathcal{C} = ((u^\gamma, y^\gamma, p^\gamma, \lambda^\gamma, \mu^\gamma))_{\gamma>0} \in \mathscr{C}$ a primal-dual path.

The following lemma provides the boundedness of a primal-dual path.

LEMMA 4.3. *If (P) fulfills the Slater condition, then there exists a positive constant $\rho < \infty$ such that for all $\gamma > 0$, $(u^\gamma, y^\gamma, p^\gamma, \lambda^\gamma, \mu^\gamma) \in \mathcal{S}_\gamma$, satisfies*

$$\|u^\gamma\|_U + \|y^\gamma\|_Y + \|p^\gamma\|_{(W^*)^N} + \|\lambda^\gamma\|_{U^*} + \|\mu^\gamma\|_{X^*} \le \rho.$$

Proof. Since $u_i^\gamma \in U_{\mathrm{ad}}^i$, we have

$$\|u^\gamma\|_U \le C_u < \infty \tag{8}$$

for some $C_u > 0$ by the boundedness of U_{ad}^i. This immediately implies the boundedness of y^γ, since

$$\|y^\gamma\|_Y \le \|S\|_{U\to Y} C_u < \infty. \tag{9}$$

The assumptions on β and $\mu^\gamma = \gamma\beta'(y^\gamma)$ imply for $z \in K$,

$$0 = \gamma\beta(z) \ge \gamma\beta(y^\gamma) + \langle \mu^\gamma, z - y^\gamma\rangle_{X^*,X} \ge \langle \mu^\gamma, z - y^\gamma\rangle_{X^*,X}.$$

By the Slater condition, there exists an $\varepsilon > 0$ and $u^0 \in U_{\mathrm{ad}}$ such that

$$\forall z \in \mathbb{B}_\varepsilon(0) : Su^0 + z \in K.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Fixing an arbitrary $z \in \mathbb{B}_\varepsilon(0) \subseteq X$, we have

$$\begin{aligned}
\langle \mu^\gamma, z\rangle_{X^*,X} &= \langle \mu^\gamma, Su^0 + z\rangle_{X^*,X} - \langle \mu^\gamma, Su^0\rangle_{X^*,X} \\
&\le \langle \mu^\gamma, y^\gamma - Su^0\rangle_{X^*,X} \\
&= \sum_{i=1}^N \langle -B_i^* p_i^\gamma - B_i^* A^{-*}(J_i^1)'(y^\gamma), u_i^\gamma - u_i^0\rangle_{U_i^*,U_i} \\
&= \sum_{i=1}^N \langle (J_i^2)'(u_i^\gamma) + \lambda_i^\gamma + B_i^* A^{-*}(J_i^1)'(y^\gamma), u_i^0 - u_i^\gamma\rangle_{U_i^*,U_i} \\
&\le \sum_{i=1}^N \langle (J_i^2)'(u_i^\gamma) + B_i^* A^{-*}(J_i^1)'(y^\gamma), u_i^0 - u_i^\gamma\rangle_{U_i^*,U_i} \\
&\le C^0 < \infty
\end{aligned}$$

for some $C^0 > 0$. Here, we used the second equation in (OS_γ) together with $S = A^{-1}B$ along with (8), (9), $u_i^0 \in U_{\mathrm{ad}}^i$, $\lambda_i^\gamma \in \mathcal{N}_{U_{\mathrm{ad}}^i}(u^\gamma)$, and the boundedness assumptions on $(J_i^1)'$, $(J_i^2)'$. This implies

$$\|\mu^\gamma\|_{X^*} = \tfrac{1}{\varepsilon}\sup\{\langle \mu^\gamma, z\rangle_{X^*,X} \mid z \in \mathbb{B}_\varepsilon(0)\} \le \tfrac{1}{\varepsilon}C^0 =: C_\mu < \infty$$

The boundedness of p^γ now follows from

$$\|p_i^\gamma\|_{W^*} = \|A^{-*}((J_i^1)'(y^\gamma) + \mu^\gamma)\|_{W^*} \le C_p^i < \infty$$

for some positive C_p^i. Furthermore, we infer

$$\|\lambda_i\|_{U_i^*} = \|B_i^* p_i^\gamma - (J_i^2)'(u_i^\gamma)\|_{U_i^*} \le C_\lambda^i,$$

for some positive constant C_λ^i. This concludes the proof. ∎

The following main result of this section establishes convergence of u^γ to a variational equilibrium of (P).

THEOREM 4.4 (limits of primal-dual paths). *If (P) fulfills the Slater condition, then for every primal-dual path $\mathcal{C} \in \mathscr{C}$ there exists a sequence $\gamma_n \to \infty$ such that*

$$u^{\gamma_n} \overset{U}{\rightharpoonup} u^*,\ y^{\gamma_n} \overset{Y}{\rightharpoonup} y^*,\ p^{\gamma_n} \overset{(W^*)^N}{\rightharpoonup} p^*,\ \lambda^{\gamma_n} \overset{U^*}{\rightharpoonup} \lambda^*,\ \mu^{\gamma_n} \overset{X^*}{\rightharpoonup}{}^* \mu^*.$$

Moreover, the point $(u^*, y^*, p^*, \lambda^*, \mu^*)$ fulfills the first-order optimality conditions $(\widehat{\mathrm{OS}})$, in particular, u^* is a variational equilibrium of (P).

Proof. Let $\mathcal{C} = (u^\gamma, y^\gamma, p^\gamma, \lambda^\gamma, \mu^\gamma)_{\gamma>0} \in \mathscr{C}$ be a primal-dual path. Then $\mathcal{C}$ is uniformly bounded in γ by Lemma 4.3. The first assertion follows readily by assumption, since U, Y, W^* are reflexive Banach spaces and bounded sets in X^* are weakly* sequentially compact.

Note that every element of the primal-dual-path fulfills the optimality system system (OS_{γ_n}). This immediately yields $Ay^* = Bu^*$. Since $B_i^* A^{-*}(J_i^1)' : Y \to U^*$ and $B_i^* A^{-*} : X^* \to U^*$ are completely continuous, we have

$$B_i^* p_i^{\gamma_n} = -B_i^* A^{-*}\big((J_i^1)'(y^{\gamma_n}) + \mu^{\gamma_n}\big) \overset{U^*}{\to} -B_i^* A^{-*}\big((J_i^1)'(y^*) + \mu^*\big) = B_i^* p_i^*.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The assumption on $(J_i^2)'$ and (OS_γ) yield for all $u_i \in U_{\mathrm{ad}}^i$

$$\begin{aligned}\langle -B_i^* p_i^* + (J_i^2)'(u_i^*), u_i^* \rangle_{U_i^*,U_i} &\le \limsup_{n\to\infty} \langle -B_i^* p_i^{\gamma_n} + (J_i^2)'(u_i^{\gamma_n}), u_i^{\gamma_n} \rangle_{U_i^*,U_i} \\ &\le \limsup_{n\to\infty} \langle -B_i^* p_i^{\gamma_n} + (J_i^2)'(u_i^{\gamma_n}), u_i \rangle_{U_i^*,U_i} \\ &= \langle -B_i^* p_i^* + (J_i^2)'(u_i^*), u_i \rangle_{U_i^*,U_i},\end{aligned}$$

i.e., $B_i^* p_i^* - (J_i^2)'(u_i^*) \in \mathcal{N}_{U_{\mathrm{ad}}^i}(u_i^*)$. We then deduce the existence of a $\bar{\lambda}_i \in \mathcal{N}_{U_{\mathrm{ad}}^i}(u_i^*)$ for which it holds that

$$\lambda_i^{\gamma_n} = B_i^* p_i^{\gamma_n} - (J_i^2)'(u_i^{\gamma_n}) \overset{U_i^*}{\rightharpoonup} B_i^* p_i^* - (J_i^2)'(u_i^*) = \bar{\lambda}_i.$$

Since the weak limit is unique, we have $\bar{\lambda}_i = \lambda_i^*$ for all $i = 1, \dots, N$, and thus $\lambda^* \in \mathcal{N}_{U_{\mathrm{ad}}^i}(u_i^*)$. The Slater condition implies in particular that there exists a feasible point $u^0 \in C$. Thus, we have

$$0 \le \gamma_n \beta(y^{\gamma_n}) \le \gamma_n \beta(Su^0) - \gamma \langle \beta'(y^{\gamma_n}), Su^0 - y^{\gamma_n} \rangle_{X^*,X} = -\langle \mu^{\gamma_n}, Su^0 - y^{\gamma_n} \rangle_{X^*,X}.$$

The boundedness of $(\mu^{\gamma_n})_{n\in\mathbb{N}}$ and $(y^{\gamma_n})_{n\in\mathbb{N}}$ implies the boundedness of $\gamma_n \beta(y^{\gamma_n})$, and regarding $\gamma_n \to \infty$ we can conclude that $\beta(y^{\gamma_n}) \to 0$. Since $u^{\gamma_n} \overset{U}{\rightharpoonup} u^*$, S is compact from U to X, and β is continuous; this yields $\beta(y^*) = 0$, and hence $y^* \in K$.

Since β is convex and differentiable, its derivative is (maximal) monotone. Moreover, for all $z \in K$, $\beta(z) = 0$ is a minimizer, which implies $\beta'(z) = 0$. It follows then that

$$\langle \beta'(y^{\gamma_n}) - \beta'(z), y^{\gamma_n} - z \rangle_{X^*,X} = \langle \beta'(y^{\gamma_n}), y^{\gamma_n} - z \rangle_{X^*,X} \ge 0.$$

This yields $\langle \mu^{\gamma_n}, z - y^{\gamma_n} \rangle_{X^*,X} \le 0$ for all $z \in K$. Passing to the limit, in light of the compactness of S from U to X, yields $\langle \mu^*, z - y^* \rangle_{X^*,X} \le 0$ for all $z \in K$. It follows then that $\mu^* \in \mathcal{N}_K(y^*)$. ∎

4.3. Nikaido–Isoda-based path-following. The general path-following strategy of [12, 13] for PDE-constrained optimization problems with state-constraints uses the optimal value function $V(\gamma)$ of the reduced problem as a merit function. When no control constraints are present, one can often find or approximate first and second derivatives of $V(\gamma)$. These are then used to develop a model of the merit function, whose values dictate the update strategy for γ.

The current game-theoretic/multiobjective setting poses two immediate difficulties: (i) each player has their own optimal value function; (ii) the presence of control constraints prohibits the derivation of similar (continuous) differentiability results.

Minimizing a weighted sum of the reduced objectives subject to the individual control constraints would lead to a parameter tuning related to the Pareto-optimality of the current solution. As we are considering (generalized) Nash equilibria, another concept is necessary.

For any $\gamma > 0$, let $\Psi_\gamma : U \times U \to \mathbb{R}$ be the Nikaido–Isoda function for (P_γ), i.e.,

$$\Psi_\gamma(u, v) := \sum_{i=1}^N \left[\mathcal{J}_i^\gamma(u_i, u_{-i}) - \mathcal{J}_i^\gamma(v_i, u_{-i}) \right],$$

where $\mathcal{J}_i^\gamma(u) := J_i^1(Su) + J_i^2(u) + \gamma\beta(Su)$ represents the objective of $(\mathrm{P}_{i,\gamma})$, and consider $V : U \times \mathbb{R}_+ \to \mathbb{R}$ defined by

$$V(u, \gamma) := \max_{v \in U_{\mathrm{ad}}} \Psi_\gamma(u, v) = \sum_{i=1}^N \mathcal{J}_i^\gamma(u_i, u_{-i}) - \min_{v \in U_{\mathrm{ad}}} \sum_{i=1}^N \mathcal{J}_i^\gamma(v_i, u_{-i}).$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

One immediately observes that $V(u,\gamma) \geq 0$ for all $u \in U_{\mathrm{ad}}$ and that analogously to Lemma 2.6, we have

$$V(u,\gamma) = 0 \text{ if and only if } u \text{ is an equilibrium.}$$

The latter property leads to the development of an update strategy for γ. Indeed, suppose for $\gamma > 0$ that u^γ is a corresponding Nash equilibrium. Then, as observed above, $V(u^\gamma,\gamma) = 0$. Since we will be using a locally convergent numerical solution method in which $\gamma \to +\infty$ and u^γ serves as the starting point for the problem with $\gamma' > \gamma$, it is important that $V(u^\gamma,\gamma')$ is sufficiently small. This would require us to solve $\min_{v\in U_{\mathrm{ad}}} \sum_{i=1}^N \mathcal{J}_i^{\gamma'}(v_i, u_{-i}^\gamma)$, decreasing γ' if necessary, until $V(u^\gamma,\gamma')$ drops below a certain threshold. Such a strategy would be highly impractical as each minimization problem requires the solution of a system of N PDEs for the state, N PDEs for the adjoint state, and N variational inequalities for the controls.

Letting $\mathcal{V}(\gamma+\eta) := V(u^\gamma,\gamma+\eta)$, $\eta > 0$, and assuming that the directional derivative $\mathcal{V}'(\gamma,\eta)$ exists, we observe that $\mathcal{V}(\gamma+t\eta) = \mathcal{V}(\gamma) + \mathcal{V}'(\gamma;t\eta) + o(t) = \mathcal{V}'(\gamma;t\eta) + o(t)$. Therefore, we can base an alternative strategy either directly on $\mathcal{V}'(\gamma;\eta)$ or an efficient approximation thereof. In fact, we will only investigate the behavior of secants of $\mathcal{V}$. Using such an approach, one could easily extend this concept to settings in which $\mathcal{V}'(\gamma;\eta)$ does not necessarily exist or where v^γ is not unique. Indeed, the crucial upper-bound derived in (11) below only requires $v^{\gamma+t\eta} \in U_{\mathrm{ad}}$ to be feasible for the problem associated with $\mathcal{V}(\gamma)$ and that v^γ is any global minimizer of the objective.

Although an explicit form for $\mathcal{V}'$ may be derived, its computation might be expensive. We therefore provide upper and lower bounds on the difference quotients $t^{-1}(\mathcal{V}(\gamma+t\eta) - \mathcal{V}(\gamma))$, which only use information that is available at the time of updating γ.

THEOREM 4.5. *For any $\gamma > 0$, let u^γ be the corresponding equilibrium and define $\mathcal{V}(\gamma+\eta) := V(u^\gamma,\gamma+\eta)$, $\eta > 0$. It holds that*

$$\begin{aligned} \eta N \beta(S(u^\gamma)) &\geq \limsup_{t\downarrow 0} t^{-1}(\mathcal{V}(\gamma+t\eta) - \mathcal{V}(\gamma)) \\ &\geq \liminf_{t\downarrow 0} t^{-1}(\mathcal{V}(\gamma+t\eta) - \mathcal{V}(\gamma)) \geq 0 \ \forall \eta > 0. \end{aligned} \tag{10}$$

Proof. Suppose that $\mathcal{V}(\gamma+t\eta) < 0$. Since $u^\gamma \in U_{\mathrm{ad}}$, it follows from the definition of $\mathcal{V}$ that

$$\sum_{i=1}^N \mathcal{J}_i^{\gamma+t\eta}(u^\gamma) < \min_{v\in U_{\mathrm{ad}}} \sum_{i=1}^N \mathcal{J}_i^{\gamma+t\eta}(v_i, u_{-i}^\gamma) \leq \sum_{i=1}^N \mathcal{J}_i^{\gamma+t\eta}(u^\gamma),$$

a contradiction. Hence, we must have $\mathcal{V}(\gamma+t\eta) \geq 0$. Then, since $\mathcal{V}(\gamma) = 0$, it follows that

$$\liminf_{t\downarrow 0} t^{-1}(\mathcal{V}(\gamma+t\eta) - \mathcal{V}(\gamma)) \geq 0.$$

Next, we set

$$v^{\gamma+t\eta} := \operatorname{argmin}\left\{\sum_{i=1}^N \mathcal{J}_i^{\gamma+t\eta}(v_i, u_{-i}^\gamma) : v \in U_{\mathrm{ad}}\right\},$$

$$v^\gamma := \operatorname{argmin}\left\{\sum_{i=1}^N \mathcal{J}_i^{\gamma}(v_i, u_{-i}^\gamma) : v \in U_{\mathrm{ad}}\right\},$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

and we make the following observation:

$$
\begin{aligned}
(11)\quad & t^{-1}(\mathcal{V}(\gamma+t\eta)-\mathcal{V}(\gamma)) \\
&= t^{-1}\left(\sum_{i=1}^{N}\mathcal{J}_i^{\gamma+t\eta}(u^\gamma)-\sum_{i=1}^{N}\mathcal{J}_i^{\gamma}(u^\gamma)\right. \\
&\qquad \left.-\left(\min_{v\in U_{\mathrm{ad}}}\sum_{i=1}^{N}\mathcal{J}_i^{\gamma+t\eta}(v_i,u_{-i}^\gamma)-\min_{v\in U_{\mathrm{ad}}}\sum_{i=1}^{N}\mathcal{J}_i^{\gamma}(v_i,u_{-i}^\gamma)\right)\right) \\
&= t^{-1}\left(t\eta N\beta(S(u^\gamma))-\left(\sum_{i=1}^{N}\mathcal{J}_i^{\gamma+t\eta}(v_i^{\gamma+t\eta},u_{-i}^\gamma)-\sum_{i=1}^{N}\mathcal{J}_i^{\gamma}(v_i^{\gamma},u_{-i}^\gamma)\right)\right) \\
&\le \eta N\beta(S(u^\gamma))-t^{-1}\left(\left(\sum_{i=1}^{N}\mathcal{J}_i^{\gamma+t\eta}(v_i^{\gamma+t\eta},u_{-i}^\gamma)-\sum_{i=1}^{N}\mathcal{J}_i^{\gamma}(v_i^{\gamma+t\eta},u_{-i}^\gamma)\right)\right) \\
&= \eta N\beta(S(u^\gamma))-\sum_{i=1}^{N}\eta\beta(S(v_i^{\gamma+t\eta},u_{-i}^\gamma))\le \eta N\beta(S(u^\gamma)).
\end{aligned}
$$

It follows that

$$\eta N\beta(S(u^\gamma))\ge \limsup_{t\downarrow 0} t^{-1}(\mathcal{V}(\gamma+t\eta)-\mathcal{V}(\gamma)),$$

as was to be shown. ∎

Recalling the discussion preceding Theorem 4.5, we see that $\mathcal{V}(\gamma+\eta)$ can be approximated from above by the term $\eta N\beta(S(u^\gamma))$. Of course, if $\beta(S(u^\gamma))=0$, then there is no need to increase γ, as the current state y^γ is feasible. This in turn means that we have found a Nash equilibrium. If we assume that the state constraint is not redundant, then we can bound the secants by a fixed threshold $\pi_{path}>0$ and by choosing $\eta>0$ such that

$$\eta N\beta(S(u^\gamma))\le \pi_{path}.$$

For example, we may choose

$$\eta=\frac{\pi_{path}}{N\beta(S(u^\gamma))}$$

and then use the update $\gamma := \gamma+\eta$. In such a case, the γ update is inversely proportional to the violation of the state constraint. This leads to the following general update strategy.

Algorithm 1. Nikaido–Isoda-based path-following for GNEPs.

Input: γ, π_{path}, $\epsilon>0$.

1: Solve (OS_γ) to obtain u^γ.
2: **while** $\beta(S(u^\gamma))>0$ **do**
3: Set $\eta := \max(\pi_{path}/N\beta(S(u^\gamma)),\epsilon)$.
4: Set $\gamma := \gamma+\eta$.
5: Solve (OS_γ) to obtain u^γ.
6: **end while**

Algorithm 1 yields $\gamma\to+\infty$, which in turn implies convergence of the path-following scheme via Theorem 4.4. In our numerical experiments, we chose ϵ on the

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

order of machine precision. Solving (OS_γ) can be a challenge in its own right. In [15], a reduction procedure for a special class of problems was developed that allowed the application of a nonsmooth Newton method. This method was shown to converge locally superlinearly for each γ in function space.

5. Applications. In section 5.1, we extend the numerical results of [15] to the case of linear parabolic PDEs, and in section 5.2, we consider a spot market problem.

5.1. Solving a parabolic GNEP. A reduction step and a nonsmooth Newton iteration. Consider the problem class in Example 2.4 in which the linear operator K^i is merely the standard canonical embedding of the state space into $L^2(Q)$. For β, we use the standard Moreau–Yosida regularization of the indicator function for the state constraint, i.e., $\beta(y) := \int_Q (y-\psi)_+^2\,dxdt = \|(y-\psi)_+\|^2_{L^2(Q)}$.

In the following, we use $\tilde{\alpha}_i = \alpha_i^{-1}$ for readability and let $(\cdot)_+ := \max(0,\cdot)$ be understood pointwise almost everywhere. The setting of the problem under investigation leads to the following first-order system associated with (P_γ) (cf. [29]):

$$u_i = \tilde{\alpha}_i B_i^* p_i - (\tilde{\alpha}_i B_i^* p_i - b_i)_+ + (a_i - \tilde{\alpha}_i B_i^* p_i)_+, \tag{12}$$

$$\partial_t y - \Delta y = Bu, \qquad y|_{\partial\Omega} = 0, \quad y(0) = 0, \tag{13}$$

$$-\partial_t p_i - \Delta p_i = y_d^i - y - \gamma(y-\psi)_+, \quad p_i|_{\partial\Omega} = 0, \quad p_i(T) = 0. \tag{14}$$

Note that $y|_{\partial\Omega} = 0$ and $y(0) = 0$ are shorthand for the boundary and initial conditions in Example 2.4. The same shorthand is analogously used for $p_i|_{\partial\Omega} = 0$ and $p_i(T) = 0$. We assume that Ω is convex or $C^{1,1}$.

Due to the form of this problem, we can reduce the number of variables when solving this system. Indeed, we first observe that $p_i = p + q_i$, where

$$-\partial_t p - \Delta p = -(y + \gamma(y-\psi)_+), \quad p|_{\partial\Omega} = 0, \quad p(T) = 0, \tag{15}$$

$$-\partial_t q_i - \Delta q_i = y_d^i, \qquad q_i|_{\partial\Omega} = 0, \quad q_i(T) = 0. \tag{16}$$

Note that (16) depends solely on data; as such, it need only be calculated once. Hence, if we define $\tilde{q}_i := \tilde{\alpha}_i B_i^* q_i$ along with $\tilde{b}_i := b_i - \tilde{q}_i$, $\tilde{a}_i := a_i - \tilde{q}_i$, and then it follows that

$$u_i = \tilde{\alpha}_i B_i^* p + \tilde{q}_i - (\tilde{\alpha}_i B_i^* p - \tilde{b}_i)_+ + (\tilde{a}_i - \tilde{\alpha}_i B_i^* p)_+.$$

Let $\mathcal{B}(p)$ denote the nonsmooth operator (applied to p) obtained by substituting this formula for u_i into the vector u in the term Bu in (13). We then consider the reduced system,

$$F_1(y,p) := \partial_t y - \Delta y - \mathcal{B}(p) = 0, \quad y|_{\partial\Omega} = 0, \quad y(0) = 0, \tag{17}$$

$$F_2(y,p) := -\partial_t p - \Delta p - (y + \gamma(y-\psi)_+) = 0, \quad p|_{\partial\Omega} = 0, \quad p(T) = 0 \tag{18}$$

with $F_1, F_2 : Y \times Y \to L^2(Q)$. It can be argued, see, e.g., [12], that the system (17)–(18) can be solved by using a nonsmooth Newton step as defined in [12, 13]. This implies in particular that for each γ, we have local superlinear convergence in the space,

$$Y = \{y \in L^2(0,T; H_0^1(\Omega) \cap H^2(\Omega)) | y_t \in L^2(0,T; L^2(\Omega)\},$$

upon invoking the usual boundedness assumptions on the Newton/slant derivative. The controls u_i are computed by substitution following the solution of (17) and (18).

Given y and p, we require Newton derivatives of $\mathcal{B}$ and $(\cdot - \psi)_+$ in order to set up the linear system for the Newton iteration, cf. [11, 12, 13]. The assumed Newton

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

derivatives are as follows:

$$G_{\mathcal{B}}(p)\delta p = B(\tilde{\alpha}_{i1}\chi_{\mathcal{I}_1(p)}B_1^*\delta p, \dots, \tilde{\alpha}_{iN}\chi_{\mathcal{I}_N(p)}B_N^*\delta p), \quad G_{(\cdot-\psi)_+}\delta y = \chi_{\mathcal{A}(y)}\delta y,$$

where

$$\mathcal{I}_i(p) := \left\{ x \in \Omega \,\middle|\, \tilde{a}_i(x) < \tilde{\alpha}_i(B_i^* p)(x) < \tilde{b}_i(x) \right\}, \quad \mathcal{A}(y) := \{ x \in \Omega \,|\, y(x) > \psi(x) \},$$

and χ is the usual characteristic function, i.e., $\chi_S(x) = 1$, if $x \in S$, and $\chi_S(x) = 0$ else. Then, given an approximate solution pair (y, p) for (17) and (18), we solve at each iteration of the semismooth Newton scheme the linear system in $(\delta y, \delta p)$:

$$\partial_t \delta y - \Delta \delta y - G_{\mathcal{B}}(p)\delta p = -F_1(y,p), \quad \delta y|_{\partial\Omega} = 0, \quad \delta y(0) = 0, \tag{19}$$

$$-\partial_t \delta p - \Delta \delta p + \delta y + \gamma \chi_{A(y)} \delta y = -F_2(y,p), \quad \delta p|_{\partial\Omega} = 0, \quad \delta p(T) = 0. \tag{20}$$

Using an implicit Euler scheme to discretize (19)–(20), we arrive at the following system:

$$L_\tau \delta y_{j+1} - \tau G_{\mathcal{B}}(p_{j+1})\delta p_{j+1} = \delta y_j - \tau F_1(y_{j+1}, p_{j+1}), \quad \delta y_0 = 0, \tag{21}$$

$$L_\tau \delta p_j + \tau \delta y_j + \tau\gamma\chi_{\mathcal{A}(y_j)}\delta y_j = \delta p_{j+1} - \tau F_2(y_j, p_j), \quad \delta p_M = 0, \tag{22}$$

where $\tau = 1/M$ is the uniform time-step size, $j = 0, \dots, M$, and $L_\tau := I - \tau\Delta$. Due to the regularity of the domain Ω, the solutions δp_i, δy_i of (21)–(22) satisfy δp_i, $\delta y_i \in H^2(\Omega) \cap H_0^1(\Omega)$ for each $i = 1, \dots, N$. Hence, (21) and (22) can be understood in $L^2(\Omega)$. We then solve (21)–(22) using a forward-backward sweep until the residuals of the equations (calculated using the L^2-norm) are sufficiently small. We then set $y := y + \delta y$, $p := p + \delta p$ and check the residual of the time-discretized version of (17)–(18). Due to the data assumptions, we may use the L^2-norm to calculate these residuals. Upon reaching the stopping tolerance, we increase γ and repeat the outer iteration until γ satisfies a stopping criterion, which we specify later.

Numerical results. Throughout this subsection, we let $N = 4$ and $\Omega = (0,1) \times (0,1)$. The Laplacian $-\Delta$ is discretized by finite differences using a standard 5-point stencil and a uniform mesh with mesh size h. The penalty parameter γ was updated following Algorithm 1 with $\pi_{\text{path}} = 10^{-5}$. Larger values of π_{path}, e.g., 10^{-2}, may lead to nonconvergence of the Newton iteration as no globalization is used. The time interval $[0, T]$, $T = 1$, is partitioned into equidistant intervals using a time step $\tau := 1/250$.

We employ a nested-grid strategy for the space discretization by using a standard 9-point prolongation to pass from the coarse to the fine mesh; cf. [8]. Motivated by elliptic control problems [10, section 4, Example 1], the mesh is refined as soon as $C\gamma^{-1} > h^2$. We use $C = 100$ in our experiments.

The operator B is defined by $Bu := \sum_{i=1}^4 \chi_{S_i} u_i$, where χ_{S_i} is the standard characteristic function of the set $S_i \subset \Omega$ given by

$$S_1 =]0, 0.5[\times]0, 0.5[\,,\ S_2 =]0.5, 1[\times]0, 0.5[\,,\ S_3 =]0, 0.5[\times]0.5, 1[\,, S_4 =]0.5, 1[\times]0.5, 1[\,.$$

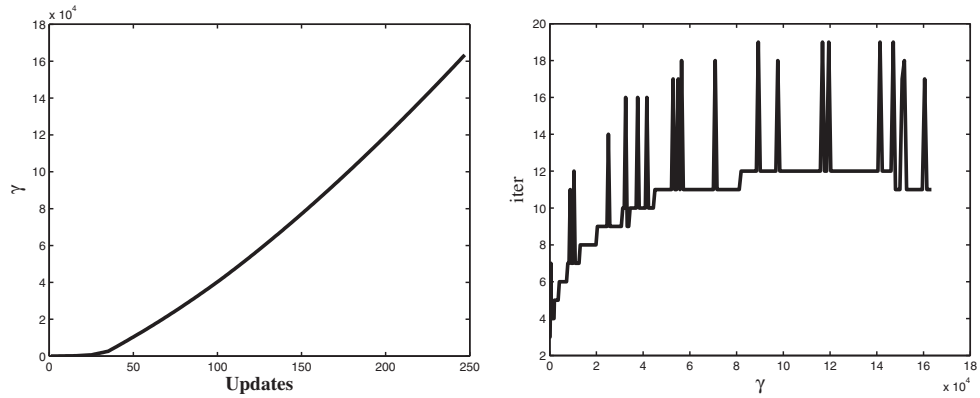
We let $a_i \equiv -.1$ and $b_i \equiv 1$ for all i and set

$$\xi(\mathbf{x}_1, \mathbf{x}_2) := .98 - \cos(5\sqrt{(\mathbf{x}_1 - 0.5)^2 + (\mathbf{x}_2 - 0.5)^2}).$$

The upper bounded ψ for the state constraint is given by

$$\psi(\mathbf{x}_1, \mathbf{x}_2, t) := \begin{cases} 2(\frac{T}{2} - t)/T + \xi(\mathbf{x}_1, \mathbf{x}_2), & t \in (0, T/2], \\ -(\frac{T}{2} - t)/T + \xi(\mathbf{x}_1, \mathbf{x}_2), & t \in (T/2, T]. \end{cases}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) Update behavior of γ (b) γ versus total sweeps of (21)–(22)

FIG. 1. *(Parabolic GNEP).*

By first defining $z_d \in H_0^1(\Omega)$ to be the solution of $-\Delta z_d = 1$, which implies $z_d > 0$ on Ω, the desired states are given by

$$y_d^1 = z_d, \quad y_d^2 = -z_d, \quad y_d^3 = -z_d, \quad y_d^4 = z_d.$$

Finally, we set $\alpha_i = 1$ for all i.

In our experiments, the nonsmooth Newton step converged for each γ in no more than two steps with a system residual (discrete L^2-norm in space-time) on the order of 10^{-10}. In Figure 1, we see the behavior of the γ-update as well as the total behavior of the fixed point iterations for each fixed γ. Notice that the update scheme is initially quite conservative but that the γ's increase more than linearly as we near feasibility in the state. The tolerance for the fixed point iteration, i.e., the forward-backward sweep of (21)–(22) needed to determine $(\delta y, \delta p)$, was set to 10^{-12}. In (b), we note that "iter" represents the total number of fixed point iterations over all the Newton steps for a fixed γ. The spikes in the number of iterations correspond to those cases in which two Newton steps were needed to solve the equilibrium problem. One also sees that, on average, 8–10 fixed point iterations were needed for each Newton iteration.

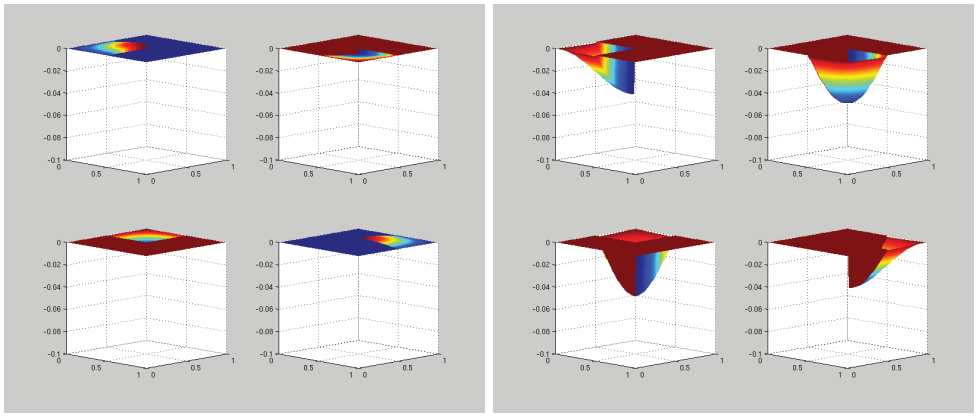
In Figure 2, we have chosen a few time frames in the evolution of the controls during which the lower bounds a_i become partially active for each of the players; see Figure 2, frames (c)–(e). The evolution of the state y^* at equilibrium can be seen in Figure 3, where activity happens around $t = 124$; see Figure 3(e).

According to the results in section 3.2, the search for a variational equilibrium in this case is equivalent to the solution of a single control problem in which the sum of the individual objective functions is minimized over $L^2(Q)^4 \times Y$ with the four controls acting on four separate regions of the domain Ω such that each control satisfies its own box constraints and both the PDE as well as the state constraint are fulfilled. This is most likely the reason for the encouraging numerical results. In the following section, we develop an alternative method that can be applied regardless of the reducibility of the search for a variational equilibrium; though convergence is not always guaranteed if the problem is not reducible.

5.2. A simplified spot market model inspired by natural gas markets. In this section, we consider a market model in which one vertically integrated producer operating a production facility transports a single (infinitely divisible) product via a

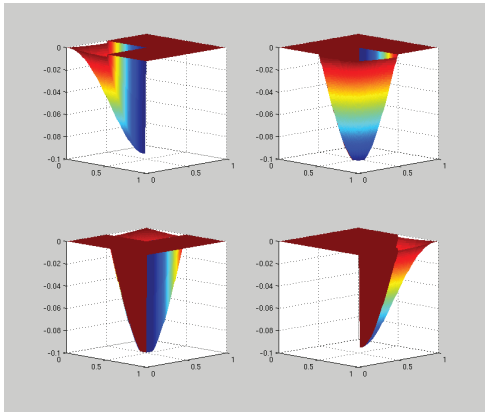
Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

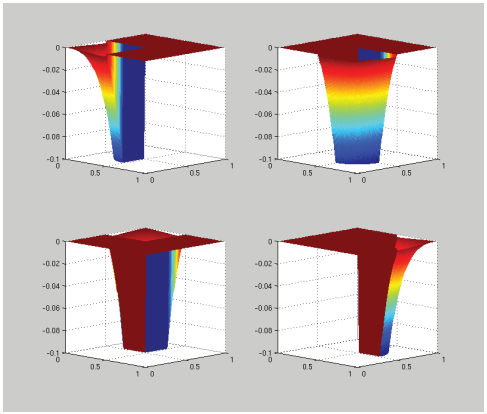


(a) $t = 0$

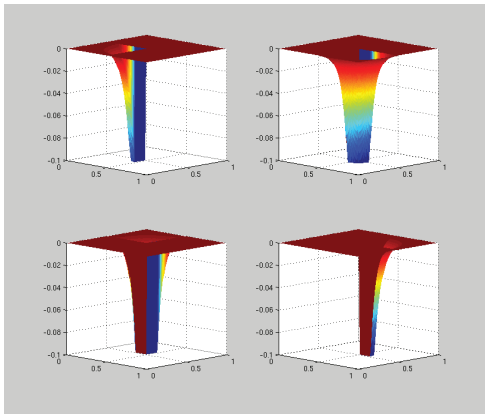
(b) $t = 100$



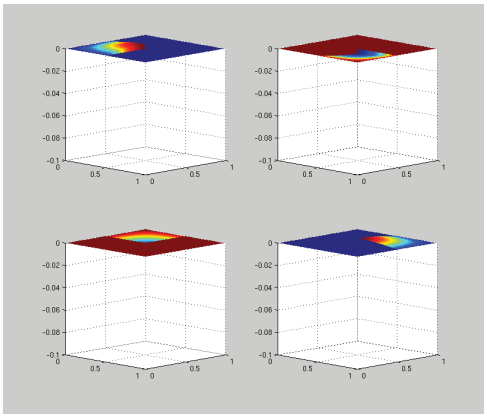
(c) $t = 110$



(d) $t = 120$



(e) $t = 125$



(f) $t = 130$

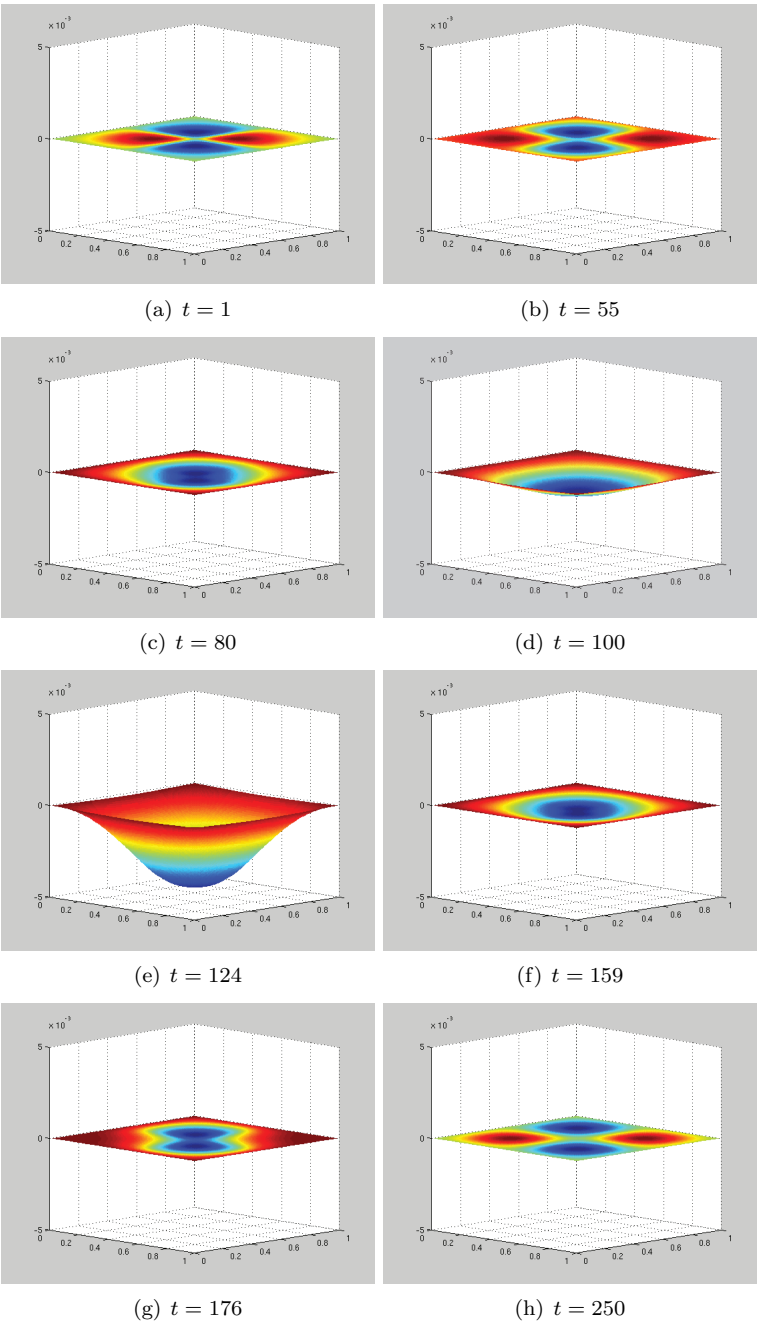
FIG. 2. *Evolution of equilibrium controls u^* (parabolic GNEP).*

pipeline. The product can be bought by wholesalers from the producer. Both the producer and the wholesalers are assumed to have customer demands to satisfy over a certain time period.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php



(a) $t = 1$ (b) $t = 55$

(c) $t = 80$ (d) $t = 100$

(e) $t = 124$ (f) $t = 159$

(g) $t = 176$ (h) $t = 250$

FIG. 3. *Evolution of equilibrium state y^* (parabolic GNEP).*

5.2.1. Derivation of the model and formulation as a GNEP. Transport and market structure. The pipeline is modeled by $\Omega := (0, 1)$. We assume a finite time horizon $[0, T]$, $T > 0$, and set $Q = (0, 1) \times (0, T)$. As before, we assume that there are at least two agents. The vertically integrated producer is denoted by the subscript 0. The superscripts 0 and 1 are used to distinguish between the production control

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

"0" and purchases on the market "1." The dynamics on the pipeline are modeled by the following parabolic system, which may be interpreted as a viscosity regularized transport equation:

$$y_t(x,t) + vy_x(x,t) - \varepsilon y_{xx}(x,t) = 0, \qquad \text{a.e. } Q, \tag{23a}$$

$$y(0,t) = u^0(t), \qquad \text{a.e. } t \in (0,T), \tag{23b}$$

$$y(1,t) = \sum_{i=0}^{N} u_i^1(t), \text{ a.e. } t \in (0,T), \tag{23c}$$

$$y(x,0) = y_0(x), \qquad \text{a.e. } x \in (0,1). \tag{23d}$$

Here, $u^0 : [0,T] \to \mathbb{R}$ is the amount of product introduced into the system at the production facility at time t. The functions $u_i^1 : [0,T] \to \mathbb{R}$, $i = 0,\dots,N$, represent the amount purchased on the market at time t by agent i. The decision variables are (u^0, u_0^1) for the producer and u_j^1, $j = 1,\dots,N$, for the wholesalers. We assume no collusion and look for a Nash equilibrium. Note that v is a velocity constant which we will henceforth take to be equal to 1 and $\varepsilon > 0$ is the viscosity parameter. Without loss of generality, we set the initial distribution in the structure $y_0 \equiv 0$. Moreover, we note that the boundary condition $y(1,t) = \sum_{i=0}^{N} u_i^1(t)$ ensures market clearing in the sense that the amount of product arriving at time t equals the amount bought by the wholesalers plus the amount sold by the producer. This might not mean, however, that demand is met exactly.

Control constraints. Since u^0 is a production quantity it is natural to impose the following lower and upper bounds:

$$0 \le u^0(t) \le \bar{u}_p, \text{ a.e. } t \in (0,1), \tag{24}$$

where $\bar{u}_p > 0$ is a scalar. Similarly we impose the purchasing capacity constraints

$$0 \le u_i^1(t) \le \bar{u}_{c,i}, \text{ a.e. } t \in (0,1), \ i = 0,\dots,N. \tag{25}$$

where $\bar{u}_{c,i} > 0$ is a scalar.

State constraints. We are interested in situations in which the total demand is satisfied at each point in time. This leads to the state constraint:

$$y(1,t) \ge \sum_{i=0}^{N} u_{d,i}(t), \text{ a.e. } t \in (0,1), \tag{26}$$

where $u_{d,i} : [0,T] \to \mathbb{R}$ is the aggregate demand of the ith agent's customers over the time period $[0,T]$. The structure of the governing system ensures that the amount taken out of the pipeline at point 1 does not exceed what is put into it at point 0:

$$u^0(t) \ge y(1,t), \text{ a.e. } t \in (0,1). \tag{27}$$

In addition, it is necessary that the amount of product in point x at time t is nonnegative and that it does not exceed a certain capacity $\bar{y}_c > 0$, a positive scalar . This yields

$$\bar{y}_c \ge y(x,t) \ge 0, \text{ a.e. } Q. \tag{28}$$

Due to the maximum principle and the nonnegative controls, the constraint $y(x,t) \ge 0$ is redundant.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(Dis)utility functions. We define $c_p(t) \geq 0$, $c_p \in L^\infty(0,T)$, to be the cost of production at time t and $c_m(t) \geq 0$, $c_m \in L^\infty(0,T)$, to be the price of product on the market at time t. In addition, we let $\mu_i > 0$, $i = 0, \ldots, N$, be a penalty that each agent must pay if their demand $u_{d,i}$ is not met exactly. For simplicity, we assume that the penalties for underdelivery and overpurchasing are equal. We let $u^1 := (u_0^1, \ldots, u_N^1)$ for readability.

Given these data, we model the objective of the producer by

$$J_0(u^0, u^1) := \underbrace{\frac{\mu_0}{2}\int_0^T |u_0^1(t) - u_{d,0}(t)|^2 dt}_{\text{Demand Misfit}} + \underbrace{\int_0^T c_p(t)u^0(t)dt}_{\text{Total (Production) Cost}} - \underbrace{\sum_{i=0}^N \int_0^T c_m(t)u_i^1(t)dt}_{\text{Total Revenue}}.$$

The wholesalers' objectives are then given by

$$J_j(u_j^1) := \underbrace{\frac{\mu_i}{2}\int_0^T |u_j^1(t) - u_{d,j}(t)|^2 dt}_{\text{Demand Misfit}} + \underbrace{\int_0^T c_m(t)u_i^1(t)dt}_{\text{Total Cost}}, \ j = 1, \ldots, N.$$

The development of the prices c_p, c_m on the markets will also be added in future works.

The GNEP. For $j = 1, \ldots, N$, using (23c) we have

$$\begin{aligned}
&\frac{\mu_j}{2}\int_0^T |u_j^1(t) - u_{d,j}(t)|^2 dt + \int_0^T c_m(t)u_j^1(t)dt \\
&= \frac{\mu_j}{2}\int_0^T |y(1,t) - u_{d,j}(t)|^2 - 2y(1,t)\overbrace{\sum_{\substack{i=0\\ i\neq j}}^N u_i^1(t)}^{=:f_j(t,u_{-j}^1)} + \overbrace{2u_{d,j}(t)\sum_{\substack{i=0\\ i\neq j}}^N u_i^1(t) + \left|\sum_{\substack{i=0\\ i\neq j}}^N u_i^1(t)\right|^2}^{=:g_j(t,u_{-j}^1)} dt \\
&\quad + \int_0^T c_m(t)u_j^1(t)dt =: J_j(u^0, u^1, y),
\end{aligned}$$

and, by analogously defining f_0 and g_0,

$$\begin{aligned}
&\frac{\mu_0}{2}\int_0^T |u_0^1(t) - u_{d,0}(t)|^2 dt + \int_0^T c_p(t)u^0(t)dt - \sum_{i=0}^N \int_0^T c_m(t)u_i^1(t)dt \\
&\quad = \frac{\mu_0}{2}\int_0^T |y(1,t) - u_{d,0}(t)|^2 - 2y(1,t)f_0(t, u_{-0}^1) + g_0(t, u_{-0}^1)dt \\
&\qquad + \int_0^T c_p(t)u^0(t)dt - \sum_{i=0}^N \int_0^T c_m(t)u_i^1(t)dt =: J_0(u^0, u^1, y).
\end{aligned}$$

The producer's problem is given by

$$\begin{aligned}
&\text{minimize } J_0(u^0, u^1, y) \text{ over } (u^0, u_0^1, y) \in L^2(0,T) \times L^2(0,T) \times Y \\
&\text{s.t.} \\
&\quad (u^0, u_0^1) \in U_{\text{ad}}^0, \\
&\qquad \bar{y}_c \geq y(x,t), \text{ in } (0,1) \times (0,T), \\
&\quad u^0(t) \geq y(1,t) \geq \textstyle\sum_{i=0}^N u_{d,j}(t), \text{ a.e. } (0,T), \\
&\qquad y \text{ solves (23)}
\end{aligned} \tag{29}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

with

$$U^0_{\text{ad}} := \big\{(v_0, v_1) \in L^2(0,T) \times L^2(0,T) \big| \, 0 \le v_0 \le \bar{u}_p, \\ \text{a.e. } (0,T) \quad 0 \le v_1 \le \bar{u}_{c,0}, \text{ a.e. } (0,T)\big\}.$$

Analogously, the wholesalers consider

$$\begin{aligned} &\text{minimize } J_j(u^0, u^1, y) \text{ over } (u^1_j, y) \in L^2(0,T) \times Y \\ &\text{s.t.} \\ &\quad u^1_j \in U^j_{\text{ad}}, \\ &\qquad \bar{y}_c \ge y(x,t), \text{ in } (0,1) \times (0,T), \\ &\quad u^0(t) \ge y(1,t) \ge \textstyle\sum_{i=0}^N u_{d,j}(t), \text{ a.e. } (0,T), \\ &\qquad y \text{ solves (23)} \end{aligned} \tag{30}$$

with

$$U^j_{\text{ad}} := \left\{ w \in L^2(0,T) \,|\, 0 \le w \le \bar{u}_{c,j}, \text{ a.e. } (0,T) \right\}.$$

Together, (29) and (30) represent a GNEP that fits into the form in section 2. The only missing aspect is the strict convexity in the decision variables u^0 and u^1. However, this aspect is not needed in the proof of Theorem 3.4. Therefore, there exists a Nash equilibrium for the GNEP defined via (29) and (30). Note also that the inclusion of additional producers and pipelines only marginally changes the form of the current problem.

Reduced form of the GNEP. In the following, we let $E_1, E_0 : Y \to L^2(0,T)$ be the evaluation operators defined by $E_1 w = w(1,\cdot) \in L^2(0,T)$ and $E_0 w = w(0,\cdot) \in L^2(0,T)$ with $w \in Y$. Since the state space requires $y(\cdot,t)$ to be at least $H^1(0,1)$ for almost every $t \in (0,T)$, it follows from the Sobolev embedding theorem that E_1, E_0 are well defined. Due to the superposition principle, y can be written as

$$y = S(u^0, u^1) = S_{\Gamma_0} u^0 + \sum_{i=0}^N S_{\Gamma_1} u^1_i.$$

Indeed, given

$$\begin{aligned} y_t(x,t) + v y_x(x,t) - \varepsilon y_{xx}(x,t) &= a(x,t), && \text{a.e. } Q, \\ y(0,t) &= b(t), && \text{a.e. } t \in (0,T), \\ y(1,t) &= c(t), && \text{a.e. } t \in (0,T), \\ y(x,0) &= d(x), && \text{a.e. } x \in (0,1) \end{aligned} \tag{31}$$

with $a \in L^2((0,1) \times (0,T))$, $b, c \in L^2(0,T)$, and $d \in H^1(0,1)$ we have the following:

1. $S_{\Gamma_0} u^0$ is the solution of (31) with $a \equiv 0$, $b = u^0$, $c \equiv 0$, and $d \equiv 0$.
2. $S_{\Gamma_1} u^1_i$ is the solution of (31) with $a \equiv 0$, $b = 0$, $c = u^1_i$, and $d \equiv 0$.

This allows us to rewrite (29) as

$$\begin{aligned} &\text{minimize } \mathcal{J}_0(u^0, u^1) := J_0(u^0, u^1, E_1 S(u^0, u^1)) \text{ over } (u^0, u^1_1) \in L^2(0,T) \times L^2(0,T) \\ &\text{s.t.} \\ &\quad (u^0, u^1_0) \in U^0_{\text{ad}}, \qquad u^0(t) \ge E_1 S(u^0, u^1) \ge \textstyle\sum_{i=0}^N u_{d,j}(t), \quad \text{a.e. } t \in (0,T), \\ &\qquad\qquad\qquad S(u^0, u^1) \le \bar{y}_c, \qquad\qquad\qquad\qquad\qquad\qquad \text{a.e. } Q, \end{aligned} \tag{32}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

and (30) as

$$
\begin{aligned}
&\text{minimize}\, \mathcal{J}_j(u^0,u^1) := J_j(u^0,u^1,E_1S(u^0,u^1)) \text{ over } u_j^1 \in L^2(0,T)\\
&\text{s.t.}\\
&\quad u_j^1 \in U_{\text{ad}}^j, \qquad u^0(t) \geq E_1S(u^0,u^1) \geq \textstyle\sum_{i=0}^N u_{d,j}(t), \qquad \text{a.e. } t \in (0,T),\\
&\qquad\qquad\qquad\quad S(u^0,u^1) \leq \bar{y}_c, \qquad\qquad\qquad\qquad\qquad\qquad \text{a.e. } Q.
\end{aligned}
\tag{33}
$$

Since $E_1S(u^0,u^1) = y(1,t) = \sum_{i=0}^N u_j^1(t)$, the associated bilateral constraint in (33) could be understood as a further control constraint. This would then amount to the form of GNEP used throughout the paper.

5.2.2. Solving the GNEP. We now apply the method developed in section 4 for the computation of variational equilibria.

Formulation of the NEP$_\gamma$. Using

$$
\begin{aligned}
\beta_1(v,z) &:= \frac{1}{2}\left\| \left(\sum_{i=0}^N u_{d,j} - v \right)_+ \right\|_{L^2(0,T)}^2 + \frac{1}{2}\|(v-z)_+\|_{L^2(0,T)}^2, \ v,z \in L^2(0,T),\\
\beta_2(w) &:= \frac{1}{2}\|(w-\bar{y}_c)\|_{L^2(Q)}^2, \ w \in L^2(0,T;L^2(0,T)),
\end{aligned}
$$

we obtain the γ-dependent NEP,

$$
\begin{aligned}
&\text{minimize}\, \mathcal{J}_0(u^0,u^1) + \gamma(\beta_1(E_1S(u^0,u^1),E_0S(u^0,u^1)) + \beta_2(S(u^0,u^1))) \text{ over } (u^0,u_0^1) \in L^2(0,T)^2\\
&\text{s.t.} \qquad (u^0,u_0^1) \in U_{\text{ad}}^0,
\end{aligned}
\tag{34}
$$

and for $j = 1,\dots,N$,

$$
\begin{aligned}
&\text{minimize}\, \mathcal{J}_j(u^0,u^1) + \gamma(\beta_1(E_1S(u^0,u^1),E_0S(u^0,u^1)) + \beta_2(S(u^0,u^1))) \text{ over } u_j^1 \in L^2(0,T)\\
&\text{s.t.} \qquad u_j^1 \in U_{\text{ad}}^j.
\end{aligned}
\tag{35}
$$

Due to the convexity and continuity of the objectives (with respect to their associated decision variables) and the weak compactness of the feasible sets U_{ad}^0 and U_{ad}^j, $j = 1,\dots,N$, it can be easily argued that there exists a Nash equilibrium to (34)–(35).

Reduction and optimality conditions. Although the current GNEP is structurally more complex than the parabolic GNEP in the previous section, there is still a possibility to reduce the computational effort via a reduction step. We first introduce adjoint variables, which are then used to illustrate a major difficulty in the development of an algorithm. We then exploit the structure of the solution and evaluation operators in order to write the optimality conditions as a system of coupled nonsmooth equations of only the controls.

First, we define the adjoint state. Given some $a \in L^2(Q)$, let p be the solution of

$$
\begin{aligned}
-\psi_t(x,t) - \psi_x(x,t) - \varepsilon\psi_{xx}(x,t) &= -a(x,t),\\
\psi(0,t) &= 0, \qquad \text{a.e. } t \in (0,T),\\
\psi(1,t) &= 0, \qquad \text{a.e. } t \in (0,T),\\
\psi(x,T) &= 0, \qquad \text{a.e. } x \in (0,1).
\end{aligned}
\tag{36}
$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Next, consider that

$$\begin{aligned}
&\int_0^T \int_0^1 a(x,t)(S(v_0,(v_1,0_{-0}))(x,t) - S(u^0,(u_0^1,0_{-0}))(x,t))dxdt \\
&\quad = \int_0^T \int_0^1 (p_t(x,t) + p_x(x,t) + \varepsilon p_{xx}(x,t))(S(v_0,(v_1,0_{-0}))(x,t) \\
&\qquad - S(u^0,(u_0^1,0_{-0}))(x,t))dxdt \\
&\quad = \varepsilon \int_0^T p_x(1,t)(v_1(t) - u_0^1(t)) - p_x(0,t)(v_0(t) - u^0(t))dt.
\end{aligned}$$

Letting $a(x,t) = \gamma(y(x,t) - \bar{y}_c)_+$, a.e. Q, we can then write the first-order optimality conditions of (34),

$$\begin{aligned}
(37)\quad &\int_0^T (c_p(t) - \gamma\left(\sum_{i=1}^N u_i^1(t) - u^0(t)\right)_+ - \varepsilon p_x(0,t))(v_0(t) - u^0(t))dt \\
&\quad + \int_0^T (\mu_0(u_0^1(t) - u_{d,0}(t)) + \gamma\left(\sum_{i=1}^N u_i^1(t) - u^0(t)\right)_+ \\
&\quad - \gamma\left(\sum_{i=0}^N u_{d,i}(t) - \sum_{i=0}^N u_i^1(t)\right)_+ - c_m(t) + \varepsilon p_x(1,t))(v_1(t) - u_0^1(t))dt \geq 0 \\
&\quad \forall (v_0,v_1) \in U_{\text{ad}}^0,
\end{aligned}$$

and of (35),

(38)

$$\begin{aligned}
&\int_0^T (c_m(t) + \mu_j(u_j^1(t) - u_{d,j}(t)) + \gamma\left(\sum_{i=1}^N u_i^1(t) - u^0(t)\right)_+ \\
&\quad - \gamma\left(\sum_{i=0}^N u_{d,i}(t) - \sum_{i=0}^N u_i^1(t)\right)_+ + \varepsilon p_x(1,t))(v_1(t) - u_j^1(t))dt \geq 0 \ \forall v_1 \in U_{\text{ad}}^j.
\end{aligned}$$

Setting $v_1 = u_0^1$ in (37), we obtain

$$\int_0^T (c_p(t) - \gamma\left(\sum_{i=1}^N u_i^1(t) - u^0(t)\right)_+ - \varepsilon p_x(0,t))(v_0(t) - u^0(t))dt \geq 0$$
$$\forall v_0 \in L^2(0,T) : v_0 \in [0,\bar{u}_p], \text{ a.e. } t \in (0,T).$$

Therefore,

$$-c_p(t) + \gamma\left(\sum_{i=1}^N u_i^1(t) - u^0(t)\right)_+ + \varepsilon p_x(0,t) \in \mathcal{N}_{[0,\bar{u}_p]}(u^0(t)), \text{ a.e. } t \in (0,T),$$

or equivalently,

$$(39)\quad u^0(t) = \text{Proj}_{[0,\bar{u}_p]}\left(u^0(t) - \alpha\left(c_p(t) - \gamma\left(\sum_{i=1}^N u_i^1(t) - u^0(t)\right)_+ - \varepsilon p_x(0,t)\right)\right),$$
$$\text{a.e. } t \in (0,T) \ \forall \alpha > 0.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Hence, using $\alpha = 1$, we have

$$
\begin{aligned}
u^0(t) = u^0(t) - c_p(t) + \gamma \left(\sum_{i=1}^N u_i^1(t) - u^0(t) \right)_+ + \varepsilon p_x(0,t) \\
+ \left(c_p(t) - \gamma \left(\sum_{i=1}^N u_i^1(t) - u^0(t) \right)_+ - \varepsilon p_x(0,t) - u^0(t) \right)_+ \\
- \left(u^0(t) - c_p(t) + \gamma \left(\sum_{i=1}^N u_i^1(t) - u^0(t) \right)_+ + \varepsilon p_x(0,t) - \bar{u}_p \right)_+ .
\end{aligned} \tag{40}
$$

Analogously, setting $v_0(t) = u^0(t)$ in (37), we obtain a structure similar to (40). Similarly, we deduce from (38) that (for almost every $t \in (0,T)$)

$$
u_j^1(t) = \Phi_j(t, u^0(t), u^1(t)) + (-\Phi_j(t, u^0(t), u^1(t)))_+)_+ - (\Phi_j(t, u^0(t), u^1(t)) - \bar{u}_{c,j})_+, \tag{41}
$$

where

$$
\begin{aligned}
\Phi_j(t, u^0(t), u^1(t)) := u_j^1(t) - c_m(t) - \mu_j(u_j^1(t) - u_{d,j}(t)) - \gamma \left(\sum_{i=1}^N u_i^1(t) - u^0(t) \right)_+ \\
+ \gamma \left(\sum_{i=0}^N u_{d,i}(t) - \sum_{i=0}^N u_i^1(t) \right)_+ - \varepsilon p_x(1,t).
\end{aligned}
$$

We have thus demonstrated the following result.

THEOREM 5.1 (characterization of Nash equilibrium). *If* $(\bar{u}^0, \bar{u}^1)$ *solves the NEP* (34)–(35), *then* $\bar{u}^0$ *satisfies* (40) *and* $\bar{u}_j^1$ *satisfies* (41), *where* y *solves* (23) *with* $(u^0, u^1) = (\bar{u}^0, \bar{u}^1)$ *and* p *solves* (36) *with* $a(x,t) = \gamma(y(x,t) - \bar{y}_c)_+$.

We propose a type of projected gradient method, which we adapt to the GNEP setting. Given some feasible (u^0, u^1) and an arbitrary $\nu > 0$, we define $(u^0(\nu), u^1(\nu))$ as the projections onto the feasible sets for the NEP_γ.

ALGORITHM 2. A PROJECTED GRADIENT METHOD FOR GNEPS.

Input: (u^0, u^1), res $>$ tol > 0, $\nu = 1$, $\sigma, \kappa \in (0,1)$, $\gamma > 0$, problem data.

```
while res > tol do
    Calculate y by solving (23).
    Calculate p using y from step 2.
    Update (u^0, u^1) using (40) and (41).
    for i = 0, ..., N do
        while J_i^γ(u^0(ν), u^1(ν)) − J_i^γ(u^0, u^1) > (−σ/ν) ||(u^0(ν), u^1(ν)) − (u^0, u^1)||^2_{L^2(0,T)^{N+1}}
        do
            ν = κν
        end while
    end for
    Set (u^0, u^1) = (u^0(ν), u^1(ν)).
    Set res := residual of first-order system in Theorem 5.1.
end while
```

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

TABLE 1
Input data for Example 5.2.

h	τ	tol	π_{path}	γ_0	ε	N	$\bar{y}_c$	$\bar{u}_c$	$\bar{u}_p$
1/256	1/200	1e-06	1e-05	1e+02	1e+00	3	3	1	3

Player	0	1	2
μ	65.2883	88.4484	25.7334

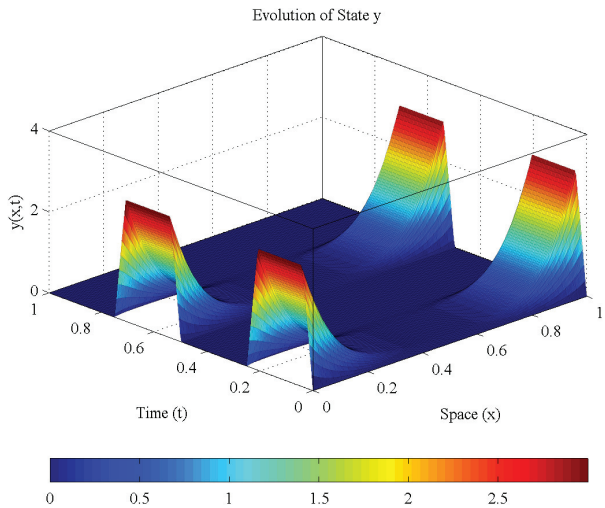


FIG. 4. *State y at equilibrium in Example 5.2.*

A full convergence analysis of this algorithm goes beyond the scope of this paper and will be investigated in future works. Nevertheless, the algorithm performs quite well in practice, i.e., similar to the projected gradient method for nonlinear programs, when used with the path-following method developed in section 4. In the following, we present and analyze the results of an implementation of Algorithm 2. We consider $c_p \equiv 1$ to be fixed in time. Given a discretization of the time interval, we assume that the discrete market price c_m is a random vector, which is uniformly distributed over $(0, 1.1)$. We generate samples of c_m in the examples using the MATLAB function rand().

Example 5.2. *A three-agent market with randomized misfit costs μ_j and market prices c_m under periodic demand.* In this example, we consider a three agent game (one producer, two wholesalers) in which the misfit costs are asymmetric, the market prices are randomly generated as described above, and the demand is periodic but symmetric. The input data can be seen in Table 1.

This example was chosen so that there are coincident active sets for the control constraints $0 \le u_j^1 \le \bar{u}_{c,j}$, the state constraint $y(x,t) \le \bar{y}_c$ (on a subset of the boundary $\{1\} \times [0,T]$), and the state constraint $y(1,t) \ge \sum_{i=1}^N u_i^1(t)$, i.e., the constraints are not redundant. The final state y can be seen in Figure 4. We observed that the wholesaler attempts to behave strategically by producing more of the product than is needed in periods of low-demand. We note that further experiments with smaller ε yielded similar results. Counting (u^0, u^1, y, p) as free variables, we have solved a GNEP with over 1e5 degrees of freedom.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

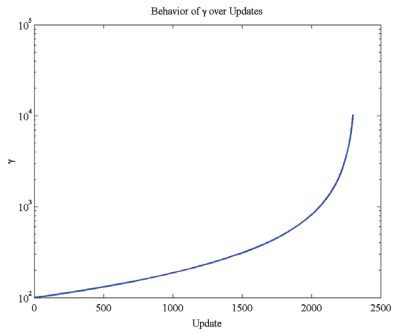
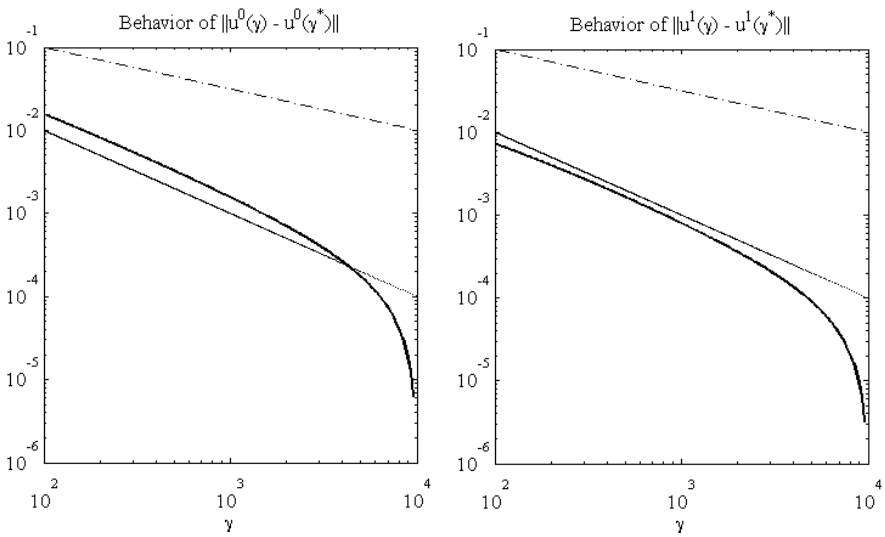


FIG. 5. γ *versus updates in Example* 5.2.



(a) (b)

FIG. 6. *Convergence rates of* u^0 *and* u^1 *(bold) versus* $\gamma^{-1/2}$ *(dashed-dotted) and* γ^{-1} *(dotted) in Example* 5.2.

The outer-loop/path-following scheme was observed to be initially very conservative, generally updating by less than 1% of γ. With the increasing feasibility of (u^0, u^1, y), however, the strategy became (predictably) more aggressive and selected successively larger steps η; see Figure 5. The number of iterations needed for the inner-loop, e.g., the projected gradient method, was generally stable with increases in γ. However, as (u^0, u^1, y) became more feasible, more iterations were required. This coincides with the typical behavior of steepest/gradient descent methods, in which the number of iterations increases as the iterates approach the solution. In particular, we observed "phases" in the number of inner iterations required for convergence. For example, for $\gamma \in (100, 150)$ an average of 4 iterations was needed, for $\gamma \in (150, 625)$ an average of 15, for $\gamma \in (625, 2300)$ between 29 and 45 inner iterations, for $\gamma \in (2300, 5000)$ roughly 54, etc. This behavior may be attributed to the fact that the control and state constraints are active on the same subsets of $\{0\} \times (0, T)$ and $\{1\} \times (0, T)$.

Finally, in Figure 6, we have plotted the convergence rates of $u^0(\gamma) \to u^0(\gamma^*)$ ($L^2(0,T)$-norm) and $u^1(\gamma) \to u^1(\gamma^*)$ ($L^2(0,T)^3$-norm) for $\gamma \to \gamma^*$. Here, $\gamma^* =$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

9870.5533 and $\gamma_0 = 100$. In both cases, the convergence rate appears to be close to $o(\gamma^{-1})$.

REFERENCES

[1] R. A. Adams and J. J.-F. Fournier, *Sobolev Spaces*, 2nd ed., Elsevier, Amsterdam, 2008.

[2] J.-P. Aubin, *Optima and Equilibria*, Grad. Texts in Math. 140, Springer, Berlin, 2003.

[3] P. D. Barba, *Multiobjective Shape Design in Electricity and Magnetism*, Lect. Notes Electr. Eng. 47, Springer, Berlin, 2010.

[4] 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.

[5] J. B. Conway, *A Course in Functional Analysis*, Grad. Texts in Math. 96, Springer, Berlin, 1985.

[6] F. Facchinei and C. Kanzow, *Generalized Nash equilibrium problems*, 4OR, 5 (2007), pp. 173–210.

[7] F. Facchinei, V. Piccialli, and M. Sciandrone, *Decomposition algorithms for generalized potential games*, Comput. Optim. Appl., 50 (2011), pp. 237–262.

[8] W. Hackbusch, *Multigrid Methods and Applications*, Springer Ser. Comput. Math. 4, Springer, Berlin, 1985.

[9] E. Hille and R. S. Phillips, *Functional Analysis and Semi-Groups*, Amer. Math. Soc. Colloq. Publi. 31. AMS, Providence, RI, 1957.

[10] M. Hintermüller and M. Hinze, *Moreau–Yosida regularization in state constrained elliptic control problems: Error estimates and parameter adjustment*, SIAM J. Numer. Anal., 47 (2009), pp. 1666–1683.

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

[12] M. Hintermüller and K. Kunisch, *Feasible and noninterior path-following in constrained minimization with low multiplier regularity*, SIAM J. Control Optim., 45 (2006), pp. 1198–1221.

[13] M. Hintermüller and K. Kunisch, *Path-following methods for a class of constrained minimization problems in function space*, SIAM J. Optim., 17 (2006), pp. 159–187.

[14] M. Hintermüller, A. Schiela, and W. Wollner, *The length of the primal-dual path in Moreau–Yosida-based path-following methods for state constrained optimal control*, SIAM J. Optim., 24 (2014), pp. 108–126.

[15] 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.

[16] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints*, Math. Model. Theory Appl. 23, Springer, Berlin, 2009.

[17] A. D. Ioffe and V. M. Tihomirov, *Theory of Extremal Problems*, Stud. Math. Appl. 6, North-Holland, Amsterdam, 1979.

[18] C. Meyer, A. Rösch, and F. Tröltzsch, *Optimal control of PDEs with regularized pointwise state constraints*, Comput. Optim. Appl., 33 (2006), pp. 209–228.

[19] D. Monderer and L. S. Shapley, *Potential games*, Games Econom. Behav., 14 (1996), pp. 124–143.

[20] H. Nikaidô and K. Isoda, *Note on non-cooperative convex games*, Pacific J. Math., 5 (1955), pp. 807–815.

[21] Á. 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.

[22] Á. 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.

[23] Á. M. Ramos and T. Roubíček, *Nash equilibria in noncooperative predator-prey games*, Appl. Math. Optim., 56 (2007), pp. 211–241.

[24] J. B. Rosen, *Existence and uniqueness of equilibrium points for concave N-person games*, Econometrica, 33 (1965), pp. 520–534.

[25] 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.

[26] T. Roubíček, *On noncooperative nonlinear differential games*, Kybernetika (Prague), 35 (1999), pp. 487–498.

[27] A. Schiela and W. Wollner, *Barrier methods for optimal control problems with convex nonlinear gradient state constraints*, SIAM J. Optim., 21 (2011), pp. 269–286.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[28] 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.

[29] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications*, Grad. Stud. Math. 112, AMS, Providence, RI, 2010.

[30] M. Ulbrich and S. Ulbrich, *Primal-dual interior-point methods for PDE-constrained optimization*, Math. Program. Ser. B, 117 (2009), pp. 435–485.

Downloaded 03/20/17 to 141.20.52.151. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.