OPTIMIZATION REFORMULATIONS OF THE GENERALIZED NASH EQUILIBRIUM PROBLEM USING NIKAIDO-ISODA-TYPE FUNCTIONS

Anna von Heusinger and Christian Kanzow

Preprint 269 July 2006

University of Würzburg
Institute of Mathematics
Am Hubland
97074 Würzburg
Germany

e-mail: heusinger@mathematik.uni-wuerzburg.de
kanzow@mathematik.uni-wuerzburg.de

July 17, 2006

Abstract. We consider the generalized Nash equilibrium problem which, in contrast to the standard Nash equilibrium problem, allows joint constraints of all players involved in the game. Using a regularized Nikaido-Isoda-function, we then present three optimization problems related to the generalized Nash equilibrium problem. The first optimization problem is a complete reformulation of the generalized Nash game in the sense that the global minima are precisely the solutions of the game. However, this reformulation is non-smooth. We then modify this approach and obtain a smooth constrained optimization problem whose global minima correspond to so-called normalized Nash equilibria. The third approach uses the difference of two regularized Nikaido-Isoda-functions in order to get a smooth unconstrained optimization problem whose global minima are, once again, precisely the normalized Nash equilibria. Conditions for stationary points to be global minima of the two smooth optimization problems are also given. Some numerical results illustrate the behaviour of our approaches.

Key Words: Generalized Nash equilibria, normalized Nash equilibria, joint constraints, regularized Nikaido-Isoda-function, constrained optimization reformulation, unconstrained optimization reformulation.

1 Introduction

We consider the generalized Nash equilibrium problem, GNEP for short. To this end, we first recall the definition of the (standard) Nash equilibrium problem, NEP for short.

Let N be the number of players. Each player $\nu \in \{1, \dots, N\}$ controls the variables $x^\nu \in \mathbb{R}^{n_\nu}$. Let $x = (x^1, \dots, x^N)^T \in \mathbb{R}^n$ be the vector formed by all these decision variables, where $n := n_1 + \dots + n_N$. To emphasize the νth player's variables within the vector x, we sometimes write $x = (x^\nu, x^{-\nu})^T$, where $x^{-\nu}$ subsumes all the other players' variables.

Let $\theta_\nu : \mathbb{R}^n \to \mathbb{R}$ be the νth player's payoff (or loss) function. We assume that these payoff functions are at least continuous, and we further assume that the functions $\theta_\nu(x) = \theta_\nu(x^\nu, x^{-\nu})$ are convex in the variable x^ν. In the standard NEP, the variable x^ν belongs to a nonempty, closed and convex set $X_\nu \subseteq \mathbb{R}^{n_\nu}$, $\nu = 1, \dots, N$. Let

$$X := X_1 \times \dots \times X_N \tag{1}$$

be the Cartesian product of the strategy sets of each player. Then a vector $x^* \in X$ is called a Nash equilibrium, or a solution of the NEP, if the block component $x^{*,\nu}$ satisfies

$$\theta(x^{*,\nu}, x^{*,-\nu}) \leq \theta(x^\nu, x^{*,-\nu}) \quad \forall x^\nu \in X_\nu \tag{2}$$

for all $\nu = 1, \dots, N$.

The GNEP generalizes the situation to some extend since now the strategy sets of player ν are allowed to depend on the rival players' strategies, too. More precisely, we assume that $X \subseteq \mathbb{R}^n$ is a nonempty, closed (not necessarily compact) and convex set which represents the joint constraints of all players $\nu = 1, \dots, N$, so that

$$X_\nu(x^{-\nu}) := \left\{ x^\nu \,\middle|\, (x^\nu, x^{-\nu}) \in X \right\} \tag{3}$$

becomes the strategy set of player $\nu, \nu = 1, \dots, N$. Note that our assumptions on X imply that each set $X_\nu(x^{-\nu})$ is also closed and convex. Moreover, if X has the Cartesian product structure as in (1), then GNEP reduces to a standard NEP. Often, the set X is given by a set of inequalities like

$$X = \left\{ x \in \mathbb{R}^n \,\middle|\, g(x) \leq 0, h_\nu(x^\nu) \leq 0 \;\; \forall \nu = 1, \dots, N \right\}$$

for some functions $g : \mathbb{R}^n \to \mathbb{R}^m$ and $h_\nu : \mathbb{R}^{n_\nu} \to \mathbb{R}^{m_\nu}$ with some numbers $m, m_\nu \geq 0$. Here g represents the joint constraints of all players, whereas h_ν depends only on the decision variables of player ν. In this situation, we therefore have

$$X_\nu(x^{-\nu}) := \left\{ x^\nu \,\middle|\, g(x^\nu, x^{-\nu}) \leq 0, h_\nu(x^\nu) \leq 0 \right\}$$

for all $\nu = 1, \dots, N$.

In the context of GNEPs, we also need the set

$$\Omega(x) := X_1(x^{-1}) \times \dots \times X_N(x^{-N}). \tag{4}$$

Then a vector $x^* \in \Omega(x^*)$ is called a generalized Nash equilibrium, or simply a solution of the GNEP, if $x^{*,\nu}$ satisfies

$$\theta(x^{*,\nu}, x^{*,-\nu}) \leq \theta(x^{\nu}, x^{*,-\nu}) \quad \forall x^{\nu} \in X_{\nu}(x^{*,-\nu}) \tag{5}$$

for each $\nu = 1, \ldots, N$.

It is well-known that the standard NEP can be reformulated as a variational inequality problem, VIP for short, see, for example, [8]. In a similar way, it is possible to characterize the GNEP as a quasi-variational inequality (abbreviated as QVI in the following), see [5, 18]. However, since there are essentially no efficient methods for solving QVIs, such a characterization is not that interesting from a practical point of view. On the other hand, it was noted in [13, 10], for example, that certain solutions of the GNEP (the normalized Nash equilibria, to be defined later) can be found by solving a suitable standard VIP associated to the GNEP. A discussion of some local issues related to this formulation is given in [11]. A globally convergent augmented Lagrangian-type VIP method is presented in [28].

However, these VIP-based methods require a higher degree of smoothness of the payoff functions θ_ν than some other approaches that are based on the Nikaido-Isoda-function (see [27]) that also plays a central role in our paper, see Section 2 for a formal definition. Relaxation methods using this Nikaido-Isoda-function are investigated in [36, 24] (see also [3, 25]) for some similar ideas), and a proximal-like method on the basis of the Nikaido-Isoda-function is presented in [13].

Here we use a regularized version of the Nikaido-Isoda-function in order to get different optimization problems whose global minima are precisely the (normalized) solutions of the GNEP. Both, the Nikaido-Isoda-function and its regularized version, are defined formally in Section 2, where we also obtain a constrained optimization problem that is completely equivalent to the GNEP. However, the objective function of this optimization problem is nonsmooth. We then modify this approach in Section 3 and obtain a smooth optimization problem whose solutions characterize the class of normalized Nash equilibria of the GNEP. Section 4 then shows how the techniques from Section 3 can be used in order to get a smooth unconstrained optimization reformulation of the normalized GNEP solutions. Preliminary numerical results are presented in Section 5, and we close with some final remarks in Section 6.

The regularized Nikaido-Isoda-function was investigated earlier for standard NEPs in [17]. Somewhat related to our work are also the two recent papers [26, 38] on equilibrium programming. The standard NEP is known to be a special case of this equilibrium programming problem, see [12]. Formally, this special case is not even mentioned in any of the two papers [26, 38]. Nevertheless, the results from [26] are closely related to the material from Section 3, but [26] does not present, for example, conditions for a stationary point of the constrained optimization problem to be a global minimum. Furthermore, the material from [38] may be viewed as the counterpart of our results from Section 4 for equilibrium programming problems. Interestingly, one way to look at the results in our paper is the fact that also the GNEP may be interpreted as a special instance of an equilibrium programming problem. To the best of our knowledge, this observation has not been made elsewhere, at least not under our general assumptions (θ_ν convex in x^ν and X not necessarily compact).

The notation used in this paper is rather standard. Here we only mention that, given a differentiable function $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, the symbols $\nabla_x \Psi(x, y)$ and $\nabla_y \Psi(x, y)$ denote the partial derivatives with respect to the x- and y-variables. Finally, we stress that, in our setting, player ν tries to minimize (not maximize) his payoff function θ_ν. Hence the name loss function would be better in this context. However, since payoff (or utility) function is the standard name in game theory, we adopt this terminology throughout this paper.

2 A Nonsmooth Constrained Optimization Reformulation

The aim of this section is to present a (nonsmooth) constrained optimization reformulation of the GNEP from (5). To this end, we use the notation from the previous section, in particular, the sets $X_\nu(x^{-\nu})$ are given by (3), and $\Omega(x)$ denotes the Cartesian product of these sets, cf. (4).

We begin with a very simple, but important observation regarding the set $\Omega(x)$.

Lemma 2.1 *We have $x \in \Omega(x)$ if and only if $x \in X$. In particular, $\Omega(x) \neq \emptyset$ for all $x \in X$.*

Proof. Using the definitions of the sets $\Omega(x)$ and $X_\nu(x^{-\nu})$, we immediately obtain

$$\begin{aligned} x \in \Omega(x) \quad &\Longleftrightarrow \quad x^\nu \in X_\nu(x^{-\nu}) \quad \forall \nu = 1, \ldots, N \\ &\Longleftrightarrow \quad (x^\nu, x^{-\nu}) \in X \quad \forall \nu = 1, \ldots, N \\ &\Longleftrightarrow \quad x = (x^\nu, x^{-\nu}) \in X. \end{aligned}$$

The second part is now obvious. □

Note that, for $x \notin X$, we have either $\Omega(x) = \emptyset$ or $\Omega(x) \neq \emptyset$, but then necessarily $x \notin \Omega(x)$. Furthermore, given any $x \in X$, simple examples show that, in general, neither $\Omega(x)$ is a subset of X nor X is included in $\Omega(x)$.

The main tool in order to obtain our optimization reformulations of the GNEP is the *Nikaido-Isoda-function*

$$\Psi(x, y) := \sum_{\nu=1}^{N} \left[\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu}) \right], \tag{6}$$

cf. [27]. Sometimes also the name *Ky-Fan-function* can be found in the literature, see [12, 13]. Using this Nikaido-Isoda-function, we define

$$V(x) := \sup_{y \in \Omega(x)} \Psi(x, y), \quad x \in X, \tag{7}$$

where, for the moment, we assume implicitly that the supremum is always attained for some $y \in \Omega(x)$. Later, this assumption will not be needed, so we do not state it here explicitly. Then it is not difficult to see that $V(x)$ is nonnegative for all $x \in \Omega(x)$, and that x^* is a solution of the GNEP if and only if $x^* \in \Omega(x^*)$ and $V(x^*) = 0$, see also the proof of Theorem 2.2 below. Therefore, finding a solution of the GNEP is equivalent to computing a global minimum of the optimization problem

$$\min V(x) \quad \text{s.t.} \quad x \in \Omega(x). \tag{8}$$

Note that this optimization problem has a complicated feasible set since $\Omega(x)$ explicitly depends on x. However, in view of Lemma 2.1, the program (8) is equivalent to the optimization problem

$$\min V(x) \quad \text{s.t.} \quad x \in X.$$

Although the Nikaido-Isoda-function is quite popular (especially for standard Nash games) in the economic and engineering literature, see, for example, [1, 2, 6, 23, 24], it has some disadvantages from a mathematical and practical point of view (also for the standard Nash game): On the one hand, given a vector x, the supremum in (6) may not exist unless additional assumptions (like the compactness of X) hold, and on the other hand, this supremum, if it exists, is usually not attained at a single point which, in turn, implies that the mapping V and, therefore, also the corresponding optimization reformulation (8) is nondifferentiable in general.

In order to overcome these deficiencies, we use a simple regularization of the Nikaido-Isoda-function. This idea was used earlier in several contexts, see, for example, Fukushima [15] (for variational inequalities), Gürkan and Pang [17] (for standard Nash games), and Mastroeni [26] (for equilibrium programming problems). Here we apply the regularization idea to GNEPs. To this end, let $\alpha > 0$ be a fixed parameter and define the *regularized Nikaido-Isoda-function* by

$$\Psi_\alpha(x,y) := \sum_{\nu=1}^N \Big[\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu}) - \frac{\alpha}{2}\|x^\nu - y^\nu\|^2\Big]. \tag{9}$$

Furthermore, for $x \in X$, let

$$\begin{aligned} V_\alpha(x) &:= \max_{y \in \Omega(x)} \Psi_\alpha(x,y) \\ &= \max_{y \in \Omega(x)} \sum_{\nu=1}^N \Big[\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu}) - \frac{\alpha}{2}\|x^\nu - y^\nu\|^2\Big] \qquad (10) \\ &= \sum_{\nu=1}^N \Big\{\theta_\nu(x^\nu, x^{-\nu}) - \min_{y^\nu \in X_\nu(x^{-\nu})} \Big[\theta_\nu(y^\nu, x^{-\nu}) + \frac{\alpha}{2}\|x^\nu - y^\nu\|^2\Big]\Big\}. \end{aligned}$$

be the corresponding value function.

A number of elementary properties of the mapping V_α are summarized in the following result.

Theorem 2.2 *The regularized function V_α has the following properties:*

(a) $V_\alpha(x) \geq 0$ *for all* $x \in \Omega(x)$.

(b) x^* *is a generalized Nash equilibrium if and only if* $x^* \in \Omega(x^*)$ *and* $V_\alpha(x^*) = 0$.

(c) For every $x \in X$, *there exists a unique vector* $y_\alpha(x) = \big(y_\alpha^1(x), \ldots, y_\alpha^N(x)\big)$ *such that for every* $\nu = 1, \ldots, N$,

$$argmin_{y^\nu \in X_\nu(x^{-\nu})} \big[\theta_\nu(y^\nu, x^{-\nu}) + \frac{\alpha}{2}\|x^\nu - y^\nu\|^2\big] = y_\alpha^\nu(x).$$

Proof. (a) For all $x \in \Omega(x)$, we have $V_\alpha(x) = \max_{y \in \Omega(x)} \Psi_\alpha(x, y) \geq \Psi_\alpha(x, x) = 0$.

(b) Suppose that x^* is a solution of the GNEP. Then $x^* \in \Omega(x^*)$ and

$$\theta_\nu(x^{*,\nu}, x^{*,-\nu}) \leq \theta_\nu(x^\nu, x^{*,-\nu}) \quad \forall x^\nu \in X_\nu(x^{*,-\nu})$$

for all $\nu = 1, \ldots, N$. Hence

$$\Psi_\alpha(x^*, y) = \sum_{\nu=1}^N \big[\underbrace{\theta_\nu(x^{*,\nu}, x^{*,-\nu}) - \theta_\nu(y^\nu, x^{*,-\nu})}_{\leq 0 \ \ \forall y^\nu \in X_\nu(x^{*,-\nu})} - \frac{\alpha}{2}\|x^{*,\nu} - y^\nu\|^2\big] \leq 0$$

for all $y \in \Omega(x^*)$. This implies

$$V_\alpha(x^*) = \max_{y \in \Omega(x^*)} \Psi_\alpha(x^*, y) \leq 0.$$

Together with part (a), we therefore have $V_\alpha(x^*) = 0$.

Conversely, assume that $x^* \in \Omega(x^*)$ and $V_\alpha(x^*) = 0$. Then $\Psi_\alpha(x^*, y) \leq 0$ holds for all $y \in \Omega(x^*)$. Let us fix a particular player $\nu \in \{1, \ldots, N\}$, and let $x^\nu \in X_\nu(x^{*,\nu})$ and $\lambda \in (0, 1)$ be arbitrary. Then define a vector $y = (y^1, \ldots, y^N) \in \mathbb{R}^n$ blockwise as follows:

$$y^\mu := \begin{cases} x^{*,\mu}, & \text{if } \mu \neq \nu, \\ \lambda x^{*,\nu} + (1-\lambda)x^\nu, & \text{if } \mu = \nu. \end{cases}$$

The convexity of the sets $X_\nu(x^{*,-\nu})$ imply that $y^\mu \in X_\mu(x^{*,-\mu})$ for all $\mu = 1, \ldots, N$, i.e., $y \in \Omega(x^*)$. For this particular y, we therefore obtain

$$\begin{aligned} 0 &\geq \Psi_\alpha(x^*, y) \\ &= \theta_\nu(x^{*,\nu}, x^{*,-\nu}) - \theta_\nu(\lambda x^{*,\nu} + (1-\lambda)x^\nu, x^{*,-\nu}) - \frac{\alpha}{2}(1-\lambda)^2\|x^{*,\nu} - x^\nu\|^2 \\ &\geq (1-\lambda)\theta_\nu(x^{*,\nu}, x^{*,-\nu}) - (1-\lambda)\theta_\nu(x^\nu, x^{*,-\nu}) - \frac{\alpha}{2}(1-\lambda)^2\|x^{*,\nu} - x^\nu\|^2 \end{aligned}$$

from the convexity of θ_ν with respect to x^ν. Dividing both sides by $1 - \lambda$ and then letting $\lambda \to 1^-$ shows that $\theta_\nu(x^{*,\nu}, x^{*,-\nu}) \leq \theta_\nu(x^\nu, x^{*,-\nu})$. Since this holds for all $x^\nu \in X_\nu(x^{*,-\nu})$

and all $\nu = 1, \dots, N$, it follows that x^* is a solution of the GNEP.

(c) This statement follows immediately from the fact that the mapping $y^\nu \mapsto \theta_\nu(y^\nu, x^{-\nu}) + \frac{\alpha}{2}\|x^\nu - y^\nu\|^2$ is strongly convex (for any given x), also taking into account that $X_\nu(x^{-\nu})$ is a nonempty, closed and convex set, cf. Lemma 2.1. □

Note that the previous result reduces to Proposition 3 in [17] for the standard Nash equilibrium problem. Using the first two statements of Theorem 2.2, we see that finding a solution of the GNEP is equivalent to computing a global minimum of the constrained optimization problem

$$\min V_\alpha(x) \quad \text{s.t.} \quad x \in \Omega(x), \tag{11}$$

which, in turn, can be reformulated as

$$\min V_\alpha(x) \quad \text{s.t.} \quad x \in X$$

in view of Lemma 2.1. The last statement of Theorem 2.2 shows that the new objective function overcomes one of the deficiencies of the mapping $V(x)$.

The following result shows that the definition of the mapping V_α can also be used in order to get a fixed point characterization of the GNEP.

Proposition 2.3 *Let $y_\alpha(x)$ be the vector defined in Theorem 2.2 (c) as the unique maximizer in the definition of the regularized function $V_\alpha(x)$, cf.* (10)*. Then x^* is a solution of GNEP if and only if x^* is a fixed point of the mapping $x \mapsto y_\alpha(x)$, i.e., if and only if $x^* = y_\alpha(x^*)$.*

Proof. First assume that x^* is a solution of GNEP. Then we obtain $x^* \in \Omega(x^*)$ (and, therefore, $x^* \in X$ in view of Lemma 2.1) and $V_\alpha(x^*) = 0$ from Theorem 2.2. In view of the definition of $y_\alpha(x^*)$, this implies

$$0 = V_\alpha(x^*) = \max_{y \in \Omega(x^*)} \Psi_\alpha(x^*, y) = \Psi_\alpha\big(x^*, y_\alpha(x^*)\big).$$

On the other hand, we also have $\Psi_\alpha(x^*, x^*) = 0$. Since $x^* \in \Omega(x^*)$ and the maximum $y_\alpha(x^*)$ is uniquely defined by Theorem 2.2, it follows that $x^* = y_\alpha(x^*)$.

Conversely, let x^* be a fixed point of the mapping y_α. Then $x^* = y_\alpha(x^*) \in \Omega(x^*)$ and

$$0 = \Psi_\alpha(x^*, x^*) = \Psi_\alpha\big(x^*, y_\alpha(x^*)\big) = V_\alpha(x^*).$$

Consequently, the statement follows from Theorem 2.2. □

We next consider a simple example which shows that, in general, the objective function from (11) is nondifferentiable.

Example 2.4 Consider the GNEP with $N = 2$ players and the following optimization problems:

$$\begin{array}{lll} \min_{x_1} & \theta_1(x_1,x_2) := -x_1 \\ \text{s.t.} & x_1 + x_2 \leq 1, \\ & 2x_1 + 4x_2 \leq 3, \\ & x_1, x_2 \geq 0, \end{array} \quad \Bigg| \quad \begin{array}{lll} \min_{x_2} & \theta_2(x_1,x_2) := 0 \\ \text{s.t.} & x_1 + x_2 \leq 1, \\ & 2x_1 + 4x_2 \leq 3, \\ & x_1, x_2 \geq 0. \end{array}$$

Hence we have $X = \{(x_1,x_2)^T \,|\, x_1 + x_2 \leq 1, 2x_1 + 4x_2 \leq 3, x_1 \geq 0, x_2 \geq 0\}$. An elementary calculation shows that the solution set is given by

$$\mathcal{S} = \Big\{ x^* = (x_1^*, x_2^*) \,\Big|\, x_2^* \in \big[0, \frac{3}{4}\big], x_1^* = \left\{ \begin{array}{ll} 1 - x_2^*, & \text{if } x_2^* \in [0, \frac{1}{2}], \\ \frac{3}{2} - 2x_2^*, & \text{if } x_2^* \in [\frac{1}{2}, \frac{3}{4}] \end{array} \right\} \Big\}.$$

We want to compute $V_\alpha(x)$. To this end, we first note that the regularized Nikaido-Isoda-function for this game is

$$\Psi_\alpha(x,y) = -x_1 + y_1 - \frac{\alpha}{2}(x_1 - y_1)^2 - \frac{\alpha}{2}(x_2 - y_2)^2.$$

Moreover, for this example, we have

$$\begin{array}{rcl} X_1(x^{-1}) & = & \big\{x_1 \,\big|\, x_1 \leq 1 - x_2, x_1 \leq \frac{3}{2} - 2x_2, x_1 \geq 0\big\} = \big[0, \min\{1 - x_2, \frac{3}{2} - 2x_2\}\big] \quad \text{and} \\ X_2(x^{-2}) & = & \big\{x_2 \,\big|\, x_2 \leq 1 - x_1, x_2 \leq \frac{3}{4} - \frac{1}{2}x_1, x_2 \geq 0\big\} = \big[0, \min\{1 - x_1, \frac{3}{4} - \frac{1}{2}x_1\}\big] \end{array}$$

and, therefore

$$V_\alpha(x) = -x_1 - \min_{y_1 \in X_1(x^{-1})} \big[-y_1 + \frac{\alpha}{2}(x_1 - y_1)^2\big] - \min_{y_2 \in X_2(x^{-2})} \big[\frac{\alpha}{2}(x_2 - y_2)^2\big].$$

Given $x = (x_1, x_2) \in \mathbb{R}^2$, the solution of the first minimization problem is given by

$$y_\alpha^1(x) = \left\{ \begin{array}{ll} 0, & \text{if } \frac{1}{\alpha} + x_1 \leq 0, \\ \frac{1}{\alpha} + x_1, & \text{if } \frac{1}{\alpha} + x_1 \in \big[0, \min\{1 - x_2, \frac{3}{2} - 2x_2\}\big], \\ \min\{1 - x_2, \frac{3}{2} - 2x_2\}, & \text{if } \frac{1}{\alpha} + x_1 \geq \min\{1 - x_2, \frac{3}{2} - 2x_2\}, \end{array} \right.$$

and the solution of the second problem is

$$y_\alpha^2(x) = \left\{ \begin{array}{ll} 0, & \text{if } x_2 \leq 0, \\ x_2, & \text{if } x_2 \in \big[0, \min\{1 - x_1, \frac{3}{4} - \frac{1}{2}x_1\}\big], \\ \min\{1 - x_1, \frac{3}{4} - \frac{1}{2}x_1\}, & \text{if } x_2 \geq \min\{1 - x_1, \frac{3}{4} - \frac{1}{2}x_1\}. \end{array} \right.$$

However, since we are only interested in $x \in X$, the above formula simplify to

$$\begin{array}{rcl} y_\alpha^1(x) & = & \left\{ \begin{array}{ll} \frac{1}{\alpha} + x_1, & \text{if } \frac{1}{\alpha} + x_1 \in \big[0, \min\{1 - x_2, \frac{3}{2} - 2x_2\}\big], \\ \min\{1 - x_2, \frac{3}{2} - 2x_2\}, & \text{if } \frac{1}{\alpha} + x_1 \geq \min\{1 - x_2, \frac{3}{2} - 2x_2\}, \end{array} \right. \\ & = & \min\big\{\frac{1}{\alpha} + x_1, 1 - x_2, \frac{3}{2} - 2x_2\big\} \end{array}$$

and

$$y_\alpha^2(x) = x_2,$$

respectively. Now it is easy to see that the corresponding mapping

$$V_\alpha(x) = -x_1 - \big[- y_\alpha^1(x) + \frac{\alpha}{2}(x_1 - y_\alpha^1(x))^2 \big]$$

is not everywhere differentiable on the feasible set X.

The nondifferentiability of the mapping V_α is a major disadvantage if one wants to apply suitable optimization methods to the corresponding reformulation (11). In the following section, we therefore describe a modification of our current approach which results into a smooth optimization reformulation of the GNEP.

We stress, however, that the situation is much more favourable if we specialize our results to the standard NEP. Then it can be shown that the mapping V_α is continuously differentiable provided all payoff functions θ_ν are smooth. This follows from the observation given in Remark 3.10 below.

3 A Smooth Constrained Optimization Reformulation

In this section, we modify the idea of the previous one and obtain another constrained optimization reformulation of the GNEP which has significantly different properties than the reformulation discussed in Section 2. In particular, the reformulation to be given here is smooth. However, it does not give a complete reformulation of all solutions of the GNEP, but it provides a characterization of the so-called normalized Nash equilibria.

Definition 3.1 *A vector $x^* \in X$ is called a* normalized Nash equilibrium *of the GNEP, if* $\sup_{y \in X} \Psi(x^*, y) \leq 0$ *holds, where Ψ denotes the Nikaido-Isoda-function from* (6).

The above definition of a normalized Nash equilibrium corresponds to one given in, e.g., [13, 36]. Note that it is slightly different from the original definition of a normalized equilibrium given in [35], see, however, the corresponding results in [13, 10]. It is not difficult to see that a normalized Nash equilibrium is always a solution of the GNEP, whereas the converse is not true in general.

We next state a simple property of the Nikaido-Isoda-function which follows immediately from the fact that the payoff functions $\theta_\nu(x) = \theta_\nu(x^\nu, x^{-\nu})$ are convex with respect to x^ν.

Lemma 3.2 *For any given $x \in X$, the Nikaido-Isoda-function $\Psi(x, y)$ is concave in $y \in X$.*

In order to derive a smooth reformulation of the GNEP, our basic tool is, once again, the regularized Nikaido-Isoda-function $\Psi_\alpha(x, y)$ from (9). Based on this mapping, we define

$$\hat{V}_\alpha(x) \;:=\; \max_{y \in X} \Psi_\alpha(x, y)$$

$$
\begin{aligned}
&= \max_{y\in X} \sum_{\nu=1}^{N} \left[\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu}) - \frac{\alpha}{2}\|x^\nu - y^\nu\|^2\right] && (12)\\
&= \max_{y\in X} \left[\Psi(x,y) - \frac{\alpha}{2}\|x-y\|^2\right].
\end{aligned}
$$

Note that, due to Lemma 3.2, given an arbitrary $x \in X$, we take the maximum of a uniformly concave function in y, hence $\hat{V}_\alpha(x)$ is well-defined. Comparing the definition of $\hat{V}_\alpha$ with the one of V_α in (10), we see that the only difference is that the maximum is taken over all $y \in X$ instead of all $y \in \Omega(x)$.

This minor change has a number of important consequences. We first state the counterpart of Theorem 2.2 for the mapping $\hat{V}_\alpha$.

Theorem 3.3 *The regularized function $\hat{V}_\alpha$ has the following properties:*

(a) $\hat{V}_\alpha(x) \geq 0$ for all $x \in X$.

(b) x^* is a normalized Nash equilibrium if and only if $x^* \in X$ and $\hat{V}_\alpha(x^*) = 0$.

(c) For every $x \in X$, there exists a unique maximizer $\hat{y}_\alpha(x)$ such that

$$
\textit{argmax}_{y\in X} \left[\Psi(x,y) - \frac{\alpha}{2}\|x-y\|^2\right] = \hat{y}_\alpha(x),
$$

and $\hat{y}_\alpha(x)$ is continuous in x.

Proof. (a) For any $x \in X$, we have $\hat{V}_\alpha(x) = \max_{y\in X} \Psi_\alpha(x,y) \geq \Psi_\alpha(x,x) = 0$.

(b) First let x^* be a normalized Nash equilibrium. Then $x^* \in X$ and $\sup_{y\in X} \Psi(x^*,y) \leq 0$. Hence $\Psi(x^*,y) \leq 0$ for all $y \in X$. Since

$$
\Psi_\alpha(x^*,y) = \Psi(x^*,y) - \frac{\alpha}{2}\|x^* - y\|^2 \leq \Psi(x^*,y) \leq 0 \quad \forall y \in X,
$$

it follows that $\hat{V}_\alpha(x^*) = \max_{y\in X} \Psi_\alpha(x^*,y) \leq 0$. Together with statement (a), this implies $\hat{V}_\alpha(x^*) = 0$.

Conversely, let $x^* \in X$ be such that $\hat{V}_\alpha(x^*) = 0$. Then

$$
\Psi_\alpha(x^*,y) \leq 0 \quad \forall y \in X. \tag{13}
$$

Assume there is a vector $\hat{y} \in X$ such that $\Psi(x^*,\hat{y}) > 0$. Then $\lambda x^* + (1-\lambda)\hat{y} \in X$ for all $\lambda \in (0,1)$, and Lemma 3.2 implies

$$
\Psi(x^*, \lambda x^* + (1-\lambda)\hat{y}) \geq \lambda \Psi(x^*,x^*) + (1-\lambda)\Psi(x^*,\hat{y}) = (1-\lambda)\Psi(x^*,\hat{y}) > 0 \quad \forall \lambda \in (0,1).
$$

Therefore, we obtain

$$
\Psi_\alpha(x^*, \lambda x^* + (1-\lambda)\hat{y}) \;=\; \Psi(x^*, \lambda x^* + (1-\lambda)\hat{y}) - \frac{\alpha}{2}\|x^* - \lambda x^* - (1-\lambda)\hat{y}\|^2
$$

$$
\begin{aligned}
&= \Psi(x^*, \lambda x^* + (1-\lambda)\hat{y}) - \frac{\alpha}{2}(1-\lambda)^2\|x^* - \hat{y}\|^2 \\
&\geq (1-\lambda)\Psi(x^*, \hat{y}) - \frac{\alpha}{2}(1-\lambda)^2\|x^* - \hat{y}\|^2 \\
&> 0
\end{aligned}
$$

for all $\lambda \in (0,1)$ sufficiently close to 1. This, however, is a contradiction to (13).

(c) In view of Lemma 3.2, the mapping $y \mapsto \Psi(x,y) - \frac{\alpha}{2}\|x-y\|^2$ is strongly concave (uniformly in x). Hence statement (c) is a consequence of standard sensitivity results, see, for example, [20, Corollaries 8.1 and 9.1]. □

Theorem 3.3 shows that we can characterize the normalized Nash equilibria of a GNEP as the global minima of the constrained optimization problem

$$\min \hat{V}_\alpha(x) \quad \text{s.t.} \quad x \in X. \tag{14}$$

In contrast to the corresponding reformulation in (11), we do not get a reformulation of all generalized Nash equilibria.

We next state the counterpart of Proposition 2.3. Its proof is omitted here since it is essentially the same as the one for Proposition 2.3 (using Theorem 3.3 instead of Theorem 2.2).

Proposition 3.4 *Let $\hat{y}_\alpha(x)$ be the vector defined in Theorem 3.3 (c) as the unique maximizer in the definition of the regularized function $\hat{V}_\alpha(x)$, cf.* (12)*. Then x^* is a normalized Nash equilibrium of GNEP if and only if x^* is a fixed point of the mapping $x \mapsto \hat{y}_\alpha(x)$.*

Our next aim is to show that the regularized function $\hat{V}_\alpha$ is continuously differentiable, provided that the payoff functions θ_ν are continuously differentiable for each player $\nu = 1,\dots,N$. The continuous differentiability of the functions θ_ν will therefore be assumed implicitly throughout the rest of this section.

Theorem 3.5 *The regularized function $\hat{V}_\alpha$ is continuously differentiable for every $x \in X$, and its gradient is given by*

$$\nabla \hat{V}_\alpha(x) = \sum_{\nu=1}^N \left[\nabla\theta_\nu(x^\nu, x^{-\nu}) - \nabla\theta_\nu(\hat{y}^\nu_\alpha(x), x^{-\nu})\right] + \begin{pmatrix} \nabla_{x^1}\theta_1(\hat{y}^1_\alpha(x), x^{-1}) \\ \vdots \\ \nabla_{x^N}\theta_N(\hat{y}^N_\alpha(x), x^{-N}) \end{pmatrix} - \alpha\big(x - \hat{y}_\alpha(x)\big),$$

where $\hat{y}_\alpha(x)$ denotes the unique maximizer from Theorem 3.3 (c) associated to the given vector x.

Proof. We first recall that the regularized function $\hat{V}_\alpha$ can be represented as in the last line of (12), and that the mapping

$$y \mapsto \Psi_\alpha(x,y) = \Psi(x,y) - \frac{\alpha}{2}\|x-y\|^2$$

is strongly concave for any fixed x in view of Lemma 3.2. Hence it follows from Danskin's Theorem (see, for example, [9]) that $\hat{V}_\alpha$ is differentiable with gradient $\nabla \hat{V}_\alpha(x) = \nabla_x \Psi_\alpha(x,y)\big|_{y=\hat{y}_\alpha(x)}$. Using the definition of the mapping Ψ_α, an elementary calculation shows that

$$\nabla_x \Psi_\alpha(x,y) = \sum_{\nu=1}^N \big[\nabla\theta_\nu(x^\nu, x^{-\nu}) - \nabla\theta_\nu(y^\nu, x^{-\nu})\big] + \begin{pmatrix} \nabla_{x^1}\theta_1(y^1, x^{-1}) \\ \vdots \\ \nabla_{x^N}\theta_N(y^N, x^{-N}) \end{pmatrix} - \alpha(x-y),$$

Inserting $y = \hat{y}_\alpha(x)$ then gives the desired formula for the gradient of $\hat{V}_\alpha$. Since all payoff functions θ_ν are continuously differentiable, and since $\hat{y}_\alpha(x)$ is also a continuous mapping of x in view of Theorem 3.3, we finally get that the gradient $\nabla \hat{V}_\alpha(x) = \nabla_x \Psi_\alpha(x,y)\big|_{y=\hat{y}_\alpha(x)}$ is continuous, i.e., the regularized function $\hat{V}_\alpha$ is continuously differentiable. □

So far, we know that (14) gives a reformulation of the GNEP as a smooth constrained optimization problem. In order to get a solution of the GNEP, however, we need to compute a global minimum of (14). Since most algorithms only find stationary points, the question arises under which conditions such a stationary point is already a global minimum. Such a condition is introduced in the following assumption.

Assumption 3.6 *For given $x \in X$ with $x \neq \hat{y}_\alpha(x)$, the inequality*

$$\sum_{\nu=1}^N \big[\nabla\theta_\nu(x^\nu, x^{-\nu}) - \nabla\theta_\nu(\hat{y}^\nu_\alpha(x), x^{-\nu})\big]^T \big(x - \hat{y}_\alpha(x)\big) > 0$$

holds.

We postpone a discussion of this assumption until the end of this section. The following result first shows that Assumption 3.6 provides a sufficient condition for a stationary point to be a global minimum and, therefore, a normalized Nash equilibrium.

Theorem 3.7 *Let $x^* \in X$ be a stationary point of* (14) *in the sense that*

$$\nabla \hat{V}_\alpha(x^*)^T (x - x^*) \geq 0 \quad \forall x \in X. \tag{15}$$

If Assumption 3.6 holds at $x = x^*$, then x^* is a normalized Nash equilibrium of the GNEP.

Proof. Using (15) and the representation of the gradient $\nabla V_\alpha(x^*)$ from Theorem 3.5, we obtain

$$\begin{aligned} 0 &\leq \nabla \hat{V}_\alpha(x^*)^T (x - x^*) \\ &= \sum_{\nu=1}^N \big[\nabla\theta_\nu(x^{*,\nu}, x^{*,-\nu}) - \nabla\theta_\nu(\hat{y}^\nu_\alpha(x^*), x^{*,-\nu})\big]^T (x - x^*) \end{aligned}$$

$$
\begin{aligned}
&+\sum_{\nu=1}^{N} \nabla_{x^\nu}\theta_\nu(\hat{y}_\alpha^\nu(x^*), x^{*,-\nu})^T(x^\nu - x^{*,\nu}) - \alpha\big(x^* - \hat{y}_\alpha(x^*)\big)^T(x - x^*) \\
= &\sum_{\nu=1}^{N} \big[\nabla\theta_\nu(x^{*,\nu}, x^{*,-\nu}) - \nabla\theta_\nu(\hat{y}_\alpha^\nu(x^*), x^{*,-\nu})\big]^T(x - x^*) \\
&+\sum_{\nu=1}^{N} \big[\nabla_{x^\nu}\theta_\nu(\hat{y}_\alpha^\nu(x^*), x^{*,-\nu}) - \alpha\big(x^{*,\nu} - \hat{y}_\alpha^\nu(x^*)\big)\big]^T\big(x^\nu - x^{*,\nu}\big)
\end{aligned}
$$

for all $x \in X$. Choosing $x = \hat{y}_\alpha(x^*)$, we therefore get

$$
\begin{aligned}
0 \;\le\; &\textstyle\sum_{\nu=1}^{N} \big[\nabla\theta_\nu(x^{*,\nu}, x^{*,-\nu}) - \nabla\theta_\nu(\hat{y}_\alpha^\nu(x^*), x^{*,-\nu})\big]^T\big(\hat{y}_\alpha(x^*) - x^*\big) \\
&+\textstyle\sum_{\nu=1}^{N} \big[\nabla_{x^\nu}\theta_\nu(\hat{y}_\alpha^\nu(x^*), x^{*,-\nu}) - \alpha(x^{*,\nu} - \hat{y}_\alpha^\nu(x^*))\big]^T\big(\hat{y}_\alpha^\nu(x^*) - x^{*,\nu}\big).
\end{aligned} \tag{16}
$$

Now recall that $\hat{y}_\alpha(x^*)$ is the unique solution of the optimization problem

$$
\max \sum_{\nu=1}^{N} \big[\theta_\nu(x^{*,\nu}, x^{*,-\nu}) - \theta_\nu(y^\nu, x^{*,-\nu}) - \frac{\alpha}{2}\|x^{*,\nu} - y^\nu\|^2\big] \quad \text{s.t.} \quad y \in X.
$$

Consequently, $\hat{y}_\alpha(x^*)$ satisfies the corresponding optimality conditions

$$
\begin{pmatrix}
\nabla_{x^1}\theta_1(\hat{y}_\alpha^1(x^*), x^{*,-1}) - \alpha(x^{*,1} - \hat{y}_\alpha^1(x^*)) \\
\vdots \\
\nabla_{x^N}\theta_N(\hat{y}_\alpha^N(x^*), x^{*,-N}) - \alpha(x^{*,N} - \hat{y}_\alpha^N(x^*))
\end{pmatrix}^T \big(z - \hat{y}_\alpha(x^*)\big) \ge 0 \quad \forall z \in X.
$$

Using $z = x^*$, we therefore obtain

$$
\sum_{\nu=1}^{N} \big[\nabla_{x^\nu}\theta_\nu(\hat{y}_\alpha^\nu(x^*), x^{*,-\nu}) - \alpha(x^{*,\nu} - \hat{y}_\alpha^\nu(x^*))\big]^T\big(x^{*,\nu} - \hat{y}_\alpha^\nu(x^*)\big) \ge 0.
$$

Taking this into account, we get

$$
0 \le \sum_{\nu=1}^{N} \big[\nabla\theta_\nu(x^{*,\nu}, x^{*,-\nu}) - \nabla\theta_\nu(\hat{y}_\alpha^\nu(x^*), x^{*,-\nu})\big]^T\big(\hat{y}_\alpha(x^*) - x^*\big) \tag{17}
$$

from (16). Now assume that $x^* \neq \hat{y}_\alpha(x^*)$. Then (17) and Assumption 3.6 together imply $0 < 0$. This contradiction shows that $x^* = \hat{y}_\alpha(x^*)$. Hence x^* is a normalized Nash equilibrium of the GNEP because of Proposition 3.4. □

Assumption 3.6 may be viewed as a kind of strict monotonicity or positive definiteness assumption. To illustrate this point, consider the case where all payoff functions θ_ν are quadratic, say

$$
\theta_\nu(x) = (x^\nu)^T A_{\nu\nu} x^\nu + \sum_{\substack{\mu=1 \\ \mu\neq\nu}}^{N} (x^\nu)^T A_{\nu\mu} x^\mu \tag{18}
$$

for certain matrices $A_{\nu\mu} \in \mathbb{R}^{n_\nu \times n_\mu}$ with $A_{\nu\nu}$ symmetric (without loss of generality). Additional linear terms are also allowed, but they do not change the subsequent discussion. Now let $A \in \mathbb{R}^{n \times n}$ be the matrix with (ν, μ)-block component $A_{\nu\mu}$, so that $A = (A_{\nu\mu})_{\nu,\mu=1}^N$. Then the following result holds.

Proposition 3.8 *Assume that the payoff functions θ_ν are given by (18) for $\nu = 1, \ldots, N$, and suppose that the matrix $A = (A_{\nu\mu})_{\nu,\mu=1}^N$ is positive definite. Then Assumption 3.6 is satisfied at an arbitrary point $x \in \mathbb{R}^n$.*

Proof. Let $x \in \mathbb{R}^n$ be arbitrarily given. Then

$$\nabla_{x^\mu} \theta_\nu(x^\nu, x^{-\nu}) = A_{\nu\mu}^T x^\nu \quad \forall \nu \neq \mu$$

and

$$\nabla_{x^\nu} \theta_\nu(x^\nu, x^{-\nu}) = 2A_{\nu\nu}x^\nu + \sum_{\substack{\mu=1 \\ \mu\neq\nu}}^N A_{\nu\mu}x^\mu = A_{\nu\nu}x^\nu + \sum_{\mu=1}^N A_{\nu\mu}x^\mu.$$

Consequently, by an elementary calculation, we obtain

$$\begin{aligned}
&\sum_{\nu=1}^N \left[\nabla\theta_\nu(x^\nu, x^{-\nu}) - \nabla\theta_\nu(\hat{y}_\alpha^\nu(x), x^{-\nu}\right]^T \big(x - \hat{y}_\alpha(x)\big) \\
&= \sum_{\nu=1}^N \sum_{\mu=1}^N \left[\nabla_{x^\mu}\theta_\nu(x^\nu, x^{-\nu}) - \nabla_{x^\mu}\theta_\nu(\hat{y}_\alpha^\nu(x), x^{-\nu})\right]^T \big(x^\mu - \hat{y}_\alpha^\mu(x)\big) \\
&= 2\sum_{\nu,\mu=1}^N \big(x^\nu - \hat{y}_\alpha^\nu(x)\big)^T A_{\nu\mu} \big(x^\mu - \hat{y}_\alpha^\mu(x)\big) \\
&= 2\big(x - \hat{y}_\alpha(x)\big)^T A \big(x - \hat{y}_\alpha(x)\big) \\
&> 0
\end{aligned}$$

whenever $x \neq \hat{y}_\alpha(x)$. Hence Assumption 3.6 holds. □

The following note shows that no regularization of the Nikaido-Isoda-function is necessary if the payoff functions θ_ν have some stronger properties than those mentioned so far.

Remark 3.9 Suppose that the functions $\theta_\nu(x) = \theta_\nu(x^\nu, x^{-\nu})$ are strongly convex in x^ν (for any given $x^{-\nu}$). Then the mapping

$$\hat{V}(x) := \max_{y \in X} \Psi(x, y)$$

is well-defined and gives a reformulation of the GNEP as a smooth optimization problem

$$\min \hat{V}(x) \quad \text{s.t.} \quad x \in X.$$

This means that there is no need to regularize the function Ψ for strongly convex payoff functions. The proof of the above statement follows by simple inspection of the proofs given in this section. Also a stationary point result can be derived similar to Theorem 3.7. Note, however, that the unconstrained optimization reformulation to be presented in Section 4 needs a regularized Nikaido-Isoda-function even in the case of strongly convex functions θ_ν.

We close this section with a simple note on the application of our results to the standard NEP.

Remark 3.10 Suppose that the nonempty, closed, and convex set $X \subseteq \mathbb{R}^n$ has the Cartesian product structure from (1). Then $\Omega(x) = X$ for all x, and the GNEP reduces to the standard NEP. Moreover, it follows that

$$V_\alpha(x) = \max_{y \in \Omega(x)} \Psi_\alpha(x, y) = \max_{y \in X} \Psi_\alpha(x, y) = \hat{V}_\alpha(x)$$

for all $x \in X$, i.e., the two functions V_α from the previous section and $\hat{V}_\alpha$ from the current section coincide. In particular, the mapping V_α is therefore also continuously differentiable when applied to a standard NEP.

4 An Unconstrained Optimization Reformulation

We use the regularized Nikaido-Isoda-function in order to obtain an unconstrained optimization reformulation of the GNEP in this section. To this end, let $0 < \alpha < \beta$ be two given parameters, let

$$\begin{aligned}
\Psi_\alpha(x, y) &:= \sum_{\nu=1}^{N} \Big[\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu}) - \frac{\alpha}{2}\|x^\nu - y^\nu\|^2\Big], \\
\Psi_\beta(x, y) &:= \sum_{\nu=1}^{N} \Big[\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu}) - \frac{\beta}{2}\|x^\nu - y^\nu\|^2\Big]
\end{aligned}$$

be the associated regularized Nikaido-Isoda functions, and let

$$\begin{aligned}
\hat{V}_\alpha(x) &:= \max_{y \in X} \Psi_\alpha(x, y) = \Psi_\alpha\big(x, \hat{y}_\alpha(x)\big), \\
\hat{V}_\beta(x) &:= \max_{y \in X} \Psi_\beta(x, y) = \Psi_\beta\big(x, \hat{y}_\beta(x)\big)
\end{aligned}$$

be the corresponding regularized value functions. Formally, these functions are defined only for $x \in X$ in the previous section. However, it is easy to see that they can be defined for any $x \in \mathbb{R}^n$.

Similar to the way the D-gap function was derived from the regularized gap function in the context of variational inequalities, see [32, 37], we then define

$$\hat{V}_{\alpha\beta}(x) := \hat{V}_\alpha(x) - \hat{V}_\beta(x), \quad x \in \mathbb{R}^n.$$

In order to show that this difference of two regularized Nikaido-Isoda-functions gives an unconstrained optimization reformulation of the GNEP, we first state the following result.

Lemma 4.1 *The inequality*

$$\frac{\beta-\alpha}{2}\|x-\hat{y}_\beta(x)\|^2 \leq \hat{V}_{\alpha\beta}(x) \leq \frac{\beta-\alpha}{2}\|x-\hat{y}_\alpha(x)\|^2 \tag{19}$$

holds for all $x \in \mathbb{R}^n$.

Proof. By definition, we have for any $x \in \mathbb{R}^n$

$$\hat{V}_\beta(x) = \Psi_\beta\big(x, \hat{y}_\beta(x)\big) = \max_{y\in X} \Psi_\beta(x,y)$$

and, therefore

$$\hat{V}_\beta(x) \geq \Psi_\beta\big(x, \hat{y}_\alpha(x)\big).$$

This implies

$$\begin{aligned}
\hat{V}_{\alpha\beta}(x) &= \hat{V}_\alpha(x) - \hat{V}_\beta(x) \\
&\leq \Psi_\alpha\big(x, \hat{y}_\alpha(x)\big) - \Psi_\beta\big(x, \hat{y}_\alpha(x)\big) \\
&= \frac{\beta-\alpha}{2}\sum_{\nu=1}^{N}\|x^\nu - \hat{y}_\alpha^\nu(x)\|^2 \\
&= \frac{\beta-\alpha}{2}\|x-\hat{y}_\alpha(x)\|^2
\end{aligned}$$

for all $x \in \mathbb{R}^n$. This proves the right-hand inequality in (19). The other inequality can be verified in a similar way. □

Note that, similar to an observation in [21], Lemma 4.1 immediately implies that the level sets of the function $\hat{V}_{\alpha\beta}$ are compact for compact sets X. This observation guarantees that any sequence $\{x^k\}$ generated by a descent method for $\hat{V}_{\alpha\beta}$ will remain bounded and, therefore, has at least one accumulation point.

As another consequence of Lemma 4.1, we obtain the following result.

Theorem 4.2 *The following statements about the function $\hat{V}_{\alpha\beta}$ hold:*

(a) $\hat{V}_{\alpha\beta}(x) \geq 0$ for all $x \in \mathbb{R}^n$.

(b) x^* is a normalized Nash equilibrium of the GNEP if and only if x^* is a global minimum of $\hat{V}_{\alpha\beta}$ with $\hat{V}_{\alpha\beta}(x^*) = 0$.

Proof. (a) Using Proposition 3.4, we have

$$\hat{V}_{\alpha\beta}(x) \geq \frac{\beta-\alpha}{2}\|x-\hat{y}_\beta(x)\|^2 \geq 0$$

for all $x \in \mathbb{R}^n$.

(b) First assume that x^* is a normalized Nash equilibrium. Then Proposition 3.4 implies $x^* = \hat{y}_\alpha(x^*)$ and $x^* = \hat{y}_\beta(x^*)$. Hence (19) immediately gives $\hat{V}_{\alpha\beta}(x^*) = 0$.

Conversely, let x^* be such that $\hat{V}_{\alpha\beta}(x^*) = 0$. Then (19) implies $x^* = \hat{y}_\beta(x^*)$. Hence x^* solves the GNEP in view of Proposition 3.4. □

Theorem 4.2 shows that the normalized Nash equilibria of GNEP are precisely the global minima of the *unconstrained* optimization problem

$$\min \hat{V}_{\alpha\beta}(x), \quad x \in \mathbb{R}^n. \tag{20}$$

We next note that this is a smooth problem. To this end, however, we need to assume, for the remainder of this section, that all payoff functions θ_ν are continuously differentiable. Then we have the following result.

Theorem 4.3 *The function $\hat{V}_{\alpha\beta}$ is continuously differentiable for every $x \in \mathbb{R}^n$, and its gradient is given by*

$$\begin{aligned}\nabla \hat{V}_{\alpha\beta}(x) &= \sum_{\nu=1}^{N} \left[\nabla\theta_\nu(\hat{y}^\nu_\beta(x), x^{-\nu}) - \nabla\theta_\nu(\hat{y}^\nu_\alpha(x), x^{-\nu})\right] \\ &\quad + \begin{pmatrix} \nabla_{x^1}\theta_1(\hat{y}^1_\alpha(x), x^{-1}) - \nabla_{x^1}\theta_1(\hat{y}^1_\beta(x), x^{-1}) \\ \vdots \\ \nabla_{x^N}\theta_N(\hat{y}^N_\alpha(x), x^{-N}) - \nabla_{x^N}\theta_N(\hat{y}^N_\beta(x), x^{-N}) \end{pmatrix} \\ &\quad -\alpha\big(x - \hat{y}_\alpha(x)\big) + \beta\big(x - \hat{y}_\beta(x)\big).\end{aligned}$$

Proof. First recall that $\hat{V}_\alpha(x)$ and $\hat{V}_\beta(x)$ are defined for all $x \in \mathbb{R}^n$. Then observe that the formula for the gradients of these two functions, as given in Theorem 3.5 for $x \in X$, remain true for all $x \in \mathbb{R}^n$. Since we have $\nabla\hat{V}_{\alpha\beta}(x) = \nabla\hat{V}_\alpha(x) - \nabla\hat{V}_\beta(x)$, the statement follows from Theorem 3.5. □

We now know that (20) is a smooth unconstrained reformulation of the GNEP. However, we need to compute the global minimum of $\hat{V}_{\alpha\beta}$. Since standard optimization software is usually only able to find a stationary point, we next want to give a result saying that such a stationary point is already a normalized Nash equilibrium under certain conditions. To this end, we first state the following preliminary result.

Lemma 4.4 *The inequality*

$$\sum_{\nu=1}^{N} \left[\nabla_{x^\nu}\theta_\nu(\hat{y}^\nu_\alpha(x), x^{-\nu}) - \nabla_{x^\nu}\theta_\nu(\hat{y}^\nu_\beta(x), x^{-\nu}) - \alpha(x^\nu - \hat{y}^\nu_\alpha(x)) + \beta(x^\nu - \hat{y}^\nu_\beta(x))\right]^T \big(\hat{y}^\nu_\beta(x) - \hat{y}^\nu_\alpha(x)\big) \geq 0$$

holds for any $x \in \mathbb{R}^n$.

Proof. As noted in the proof of Theorem 3.7, $\hat{y}^\nu_\alpha(x)$ satisfies the optimality condition

$$\sum_{\nu=1}^N \big[\nabla_{x^\nu}\theta_\nu(\hat{y}^\nu_\alpha(x), x^{-\nu}) - \alpha(x^\nu - \hat{y}^\nu_\alpha(x))\big]^T \big(z^\nu - \hat{y}^\nu_\alpha(x)\big) \geq 0 \quad \forall z \in X.$$

In a similar way, it follows that $\hat{y}^\nu_\beta(x)$ satisfies

$$\sum_{\nu=1}^N \big[\nabla_{x^\nu}\theta_\nu(\hat{y}^\nu_\beta(x), x^{-\nu}) - \beta(x^\nu - \hat{y}^\nu_\beta(x))\big]^T \big(z^\nu - \hat{y}^\nu_\beta(x)\big) \geq 0 \quad \forall z \in X.$$

Using $z = \hat{y}_\beta(x)$ in the first inequality and $z = \hat{y}_\alpha(x)$ in the second inequality, we get

$$\sum_{\nu=1}^N \big[\nabla_{x^\nu}\theta_\nu(\hat{y}^\nu_\alpha(x), x^{-\nu}) - \alpha(x^\nu - \hat{y}^\nu_\alpha(x))\big]^T \big(\hat{y}^\nu_\beta(x) - \hat{y}^\nu_\alpha(x)\big) \geq 0$$

and

$$\sum_{\nu=1}^N \big[\nabla_{x^\nu}\theta_\nu(\hat{y}^\nu_\beta(x), x^{-\nu}) - \beta(x^\nu - \hat{y}^\nu_\beta(x))\big]^T \big(y^\nu_\alpha(x) - \hat{y}^\nu_\beta(x)\big) \geq 0,$$

respectively. Adding these two inequalities gives the desired result. □

In order to state a result that a stationary point is, automatically, a global minimum of $\hat{V}_{\alpha\beta}$, we need a certain condition which is quite similar to the one stated in Assumption 3.6.

Assumption 4.5 *For given $x \in \mathbb{R}^n$ with $\hat{y}_\alpha(x) \neq \hat{y}_\beta(x)$, the inequality*

$$\sum_{\nu=1}^N \big[\nabla\theta_\nu(\hat{y}^\nu_\beta(x), x^{-\nu}) - \nabla\theta_\nu(\hat{y}^\nu_\alpha(x), x^{-\nu})\big]^T \big(\hat{y}_\beta(x) - \hat{y}_\alpha(x)\big) > 0$$

holds.

It is easy to see from the proof of Proposition 3.8 that, under the assumption of that result, we also have a sufficient condition for Assumption 4.5 to hold. Using Assumption 4.5, we are now able to state the following result.

Theorem 4.6 *Let x^* be a stationary point of $\hat{V}_{\alpha\beta}$. If Assumption 4.5 holds at $x = x^*$, then x^* is a normalized Nash equilibrium of the GNEP.*

Proof. Since x^* is a stationary point of $\hat{V}_{\alpha\beta}$, we obtain from Theorem 4.3

$$\begin{aligned} 0 &= \nabla\hat{V}_{\alpha\beta}(x^*) \\ &= \sum_{\nu=1}^N \big[\nabla\theta_\nu(\hat{y}^\nu_\beta(x^*), x^{*,-\nu}) - \nabla\theta_\nu(\hat{y}^\nu_\alpha(x^*), x^{*,-\nu})\big] \end{aligned}$$

$$
+\begin{pmatrix} \nabla_{x^1}\theta_1(\hat{y}^1_\alpha(x^*),x^{*,-1}) - \nabla_{x^1}\theta_1(\hat{y}^1_\beta(x^*),x^{*,-1}) \\ \vdots \\ \nabla_{x^N}\theta_N(\hat{y}^N_\alpha(x^*),x^{*,-N}) - \nabla_{x^N}\theta_N(\hat{y}^N_\beta(x^*),x^{*,-N}) \end{pmatrix} \tag{21}
$$
$$
-\alpha\big(x^* - \hat{y}_\alpha(x^*)\big) + \beta\big(x^* - \hat{y}_\beta(x^*)\big).
$$

Multiplication with $\big(\hat{y}_\beta(x^*) - \hat{y}_\alpha(x^*)\big)^T$ and using Lemma 4.4, we therefore get

$$
\begin{aligned}
0 \;&=\; \sum_{\nu=1}^N \big[\nabla\theta_\nu(\hat{y}^\nu_\beta(x^*),x^{*,-\nu}) - \nabla\theta_\nu(\hat{y}^\nu_\alpha(x^*),x^{*,-\nu})\big]^T\big(\hat{y}_\beta(x^*) - \hat{y}_\alpha(x^*)\big) \\
&\quad + \sum_{\nu=1}^N \big[\nabla_{x^\nu}\theta_\nu(\hat{y}^\nu_\alpha(x^*),x^{*,-\nu}) - \nabla_{x^\nu}\theta_\nu(\hat{y}^\nu_\beta(x^*),x^{*,-\nu}) \\
&\qquad\quad -\alpha(x^{*,\nu} - \hat{y}^\nu_\alpha(x^*)) + \beta(x^{*,\nu} - \hat{y}^\nu_\beta(x^*))\big]^T\big(\hat{y}^\nu_\beta(x^*) - \hat{y}^\nu_\alpha(x^*)\big) \\
&\geq\; \sum_{\nu=1}^N \big[\nabla\theta_\nu(\hat{y}^\nu_\beta(x^*),x^{*,-\nu}) - \nabla\theta_\nu(\hat{y}^\nu_\alpha(x^*),x^{*,-\nu})\big](\hat{y}_\beta(x^*) - \hat{y}_\alpha(x^*)).
\end{aligned}
$$

Assume that $\hat{y}_\beta(x^*) - \hat{y}_\alpha(x^*) \neq 0$. Then the previous chain of inequalities together with Assumption 4.5 gives the contradiction $0 > 0$. Hence $\hat{y}_\alpha(x^*) = \hat{y}_\beta(x^*)$. But then (21) simplifies to $(\beta - \alpha)(x^* - \hat{y}_\alpha(x^*)) = 0$. Since $\alpha < \beta$, this implies $x^* = \hat{y}_\alpha(x^*)$. Consequently, x^* is a normalized Nash equilibrium in view of Proposition 3.4. □

5 Numerical Illustrations

Here we want to illustrate the performance of our unconstrained optimization reformulation on some GNEPs taken from the literature. To this end, we use the Barzilai-Borwein (BB) gradient method [4] (see also [33, 34, 14, 7, 16] for some subsequent modifications and investigations of this first-order method) for the unconstrained minimization of the objective function $\hat{V}_{\alpha\beta}$. This method uses the iterative procedure

$$
x^{k+1} := x^k - \alpha_k \nabla\hat{V}_{\alpha\beta}(x^k), \quad k = 0,1,2,\ldots
$$

with the stepsize

$$
\alpha_k := \frac{s^T y}{s^T s},
$$

where

$$
s := x^k - x^{k-1}, \quad y := \nabla\hat{V}_{\alpha\beta}(x^k) - \nabla\hat{V}_{\alpha\beta}(x^{k-1}).
$$

Hence the BB method has the advantage of using an explicit formula for the stepsize, so no extra line search is required which would be very expensive in our case since this would require further evaluations of the mapping $\hat{V}_{\alpha\beta}$. Each function evaluation of $\hat{V}_{\alpha\beta}$, however, needs the solution of two constrained optimization problems in order to compute $\hat{y}_\alpha(x)$ and $\hat{y}_\beta(x)$. We therefore believe that the BB method fits perfectly into our setting.

The implementation is done in MATLAB using the function `fmincon` from the Optimization Toolbox to compute the values $\hat{y}_\alpha(x)$ and $\hat{y}_\beta(x)$. We use the parameters $\alpha = 0.02$ and $\beta = 0.05$ for all test examples, and terminate the iteration if $\hat{V}_{\alpha\beta}(x^k) \leq \varepsilon$ with $\varepsilon := 10^{-8}$.

Example 5.1 This test problem is taken from [10, Example 3.8]. There are two players, each player has one decision variable. The example has infinitely many Nash equilibria, but only one normalized Nash equilibrium at $x^* := \left(\frac{3}{4}, \frac{1}{4}\right)^T$. The iteration history of our method applied to this example with the starting point $x^0 := (0,0)^T$ is given in Table 1.

k	x_1^k	x_2^k	$\hat{V}_{\alpha\beta}(x^k)$	$\|\nabla\hat{V}_{\alpha\beta}(x^k)\|$
0	0.00000000	0.00000000	0.0093111567	0.0235577850
1	0.00010481	0.00003654	0.0093085421	0.0235544767
2	0.74630253	0.26018955	0.0000017132	0.0003161316
3	0.74978601	0.25017191	0.0000000011	0.0000079447

Table 1: Numerical results for Example 5.1

Example 5.2 This is the duopoly example from [24], so there are two players. Each player controls one variable. The solution is $x^* := \left(\frac{16}{3}, \frac{16}{3}\right)^T$. The iteration history of our method, using the starting point $x^0 := (2,0)^T$ from [24], is summarized in Table 2.

k	x_1^k	x_2^k	$\hat{V}_{\alpha\beta}(x^k)$	$\|\nabla\hat{V}_{\alpha\beta}(x^k)\|$
0	2.00000000	0.00000000	1.2315865733	0.1366926294
1	2.00198061	0.00316894	1.2301234437	0.1366113010
2	5.32890551	5.32619359	0.0000021945	0.0001825423
3	5.33333413	5.33333003	0.0000000000	0.0000000907

Table 2: Numerical results for Example 5.2

Example 5.3 Here we use the river basin pollution game from [24]. This time there are three players, and once again each player controls only one variable. Table 3 contains the iteration history for our unconstrained minimization approach using the starting point $x^0 := (0,0,0)^T$ from [24].

k	x_1^k	x_2^k	x_3^k	$\hat{V}_{\alpha\beta}(x^k)$	$\|\nabla\hat{V}_{\alpha\beta}(x^k)\|$
0	0.00000000	0.00000000	0.0000000000	7.4511392248	0.6336400961
1	0.04928389	0.04525843	0.0443943119	7.4003515320	0.6313081588
2	13.39298590	12.30468520	12.0524627775	0.5374357435	0.0993373542
3	13.95131677	14.34063963	9.3604015563	0.2743881869	0.0596227368
4	16.45591073	16.66153325	6.3640055287	0.0915913193	0.0310976805
5	19.25548426	15.68602395	4.2473780863	0.0163221419	0.0143759817
6	21.14761847	16.89516211	3.1383210958	0.0094477244	0.0212229515
7	20.33577610	15.10976412	1.9706797290	0.0293617720	0.0415040130
8	21.22675171	15.91622108	2.9990735639	0.0009997184	0.0071599812
9	21.09176520	15.94309622	2.8023980917	0.0000769253	0.0015397590
10	21.09657997	15.99462888	2.7807787818	0.0000190059	0.0007236726
11	21.10779025	16.03326200	2.7539403754	0.0000050857	0.0002581068
12	21.12180189	16.02691746	2.7432959196	0.0000001780	0.0000750139
13	21.12831611	16.02837413	2.7408150983	0.0000000804	0.0000614958
14	21.12884798	16.02676549	2.7327078937	0.0000003541	0.0001442056
15	21.13481718	16.02889292	2.7376402083	0.0000001665	0.0000981303
16	21.13355539	16.02799321	2.7346145670	0.0000002679	0.0000688952
17	21.13582739	16.02805299	2.7330809781	0.0000002084	0.0000578498
18	21.14671036	16.02782075	2.7242447250	0.0000000010	0.0000069063

Table 3: Numerical results for Example 5.3

Example 5.4 This is the electricity market example from [19], also used in [6]. There are three players (companies) and altogether six variables, the first player controls one variable (corresponding to one electricity plant), the second player controls two variables (he owns two electricity plants), and the third player controls the remaining three variables (electricity plants). Starting at $x^0 := (0, \dots, 0)^T$, we get the iteration history given in Table 4 (where only the first three components of the iteration vector x^k are shown). The method terminates at the approximate solution

$$x^{13} \approx \big(46.599, 32.133, 14.935, 22.072, 12.439, 12.438\big)^T.$$

Note that the final value of $\hat{V}_{\alpha\beta}(x^k)$ is slightly negative. This is due to the fact that the computation of the values $\hat{y}_\alpha(x^k)$ and $\hat{y}_\beta(x^k)$ are (necessarily) done inexactly.

Example 5.5 Here we consider the internet switching model from [22] in a slightly modified version that was also analysed in [11]. More precisely, we use the constraints $x_\nu \geq l_\nu$ for all players with the lower bounds $l_\nu := 0.01$ for all $\nu = 1, \dots, N$ and the additional capacity constraint $\sum_{\nu=1}^N x_\nu \leq B$ for some positive constant B. Note that we use positive lower bounds here since otherwise the payoff functions of the players used in [22, 11] are

k	x_1^k	x_2^k	x_3^k	$\hat{V}_{\alpha\beta}(x^k)$	$\|\nabla\hat{V}_{\alpha\beta}(x^k)\|$
0	0.00000000	0.00000000	0.0000000000	212.9967936440	5.4266419193
1	3.97795020	2.37873182	1.6949517306	182.3524379329	4.9797766535
2	49.04627557	28.53584600	20.2046625231	6.0343418007	1.2218601926
3	42.99451338	23.68477781	14.9208122921	1.5059663257	0.5429641149
4	45.02683837	25.85239314	16.1809149899	0.3868362085	0.0821865582
5	45.12470353	26.26870608	15.9640738490	0.3374590618	0.0761857866
6	46.50488593	31.59216381	13.5591488174	0.0289298316	0.0265612735
7	46.56155685	31.53880993	15.8250292617	0.0084523325	0.0178959824
8	46.83875964	32.38449530	15.0597168570	0.0201040891	0.0727633378
9	45.99036086	31.60132317	14.4448223076	0.1060890498	0.1687710947
10	46.59695610	32.13906308	14.9259196777	0.0001640659	0.0020086656
11	46.59310677	32.12701437	14.9267478916	0.0002346250	0.0034718130
12	46.59407813	32.13732663	14.9337793376	0.0001464873	0.0019340344
13	46.59853223	32.13333385	14.9345457848	-0.0000262409	0.0033345202

Table 4: Numerical results for Example 5.4

not defined everywhere on the feasible set. We take $N = 10$ and $B = 1$ for our computations, together with the starting point $x^0 = (0.1, \ldots, 0.1)^T$ which is close to the solution at $x^* = (0.09, \ldots, 0.09)^T$. We therefore expect that our method converges after just a few steps, and this is indeed the case, see Table 5.

k	x_1^k	x_2^k	$\hat{V}_{\alpha\beta}(x^k)$	$\|\nabla\hat{V}_{\alpha\beta}(x^k)\|$
0	0.10000000	0.10000000	0.0003822155	0.0247426159
1	0.09996128	0.09996128	0.0003791924	0.0246404400
2	0.09062399	0.09062399	0.0000014200	0.0014458597
3	0.09004195	0.09004193	0.0000000081	0.0000220902

Table 5: Numerical results for Example 5.5

We next give a short comparison of our method with the relaxation method from [36]. This relaxation method computes a sequence $\{x^k\}$ according to the formula

$$x^{k+1} := (1 - \alpha_k)x^k + \alpha_k y(x^k), \quad k = 0, 1, 2, \ldots,$$

where $\alpha_k \in (0, 1]$ is a suitable parameter and $y(x^k)$ is the same as $y_\alpha(x^k)$ from Theorem 3.3 with $\alpha = 0$, i.e., we obtain $y(x^k)$ by maximizing the Nikaido-Isoda function without regularization. Theoretically, the relaxation method is well-defined only under some stronger assumptions that are not necessary for our approaches. However, if these conditions are met and the parameter α_k satisfies

$$\alpha_k \to 0 \quad \text{and} \quad \sum_{k=0}^{\infty} \alpha_k = \infty, \tag{22}$$

the method converges to a normalized Nash equilibrium, see [36] for details. The two requirements (22) suggest to take $\alpha_k := \frac{1}{k+1}$ or a related updating. Since numerical results indicate that the relaxation method converges much better if α_k is taken to be a fixed number during the first few iterations, we use an idea similar to [2] and choose

$$\alpha_k := \begin{cases} 0.5, & \text{if } k \leq 40, \\ 0.5\frac{1}{k-40}, & \text{if } k > 40. \end{cases}$$

This choice guarantees that, theoretically, (22) holds, whereas in practice this means that we take the constant value $\alpha_k = 0.5$ for all four examples.

We terminate the relaxation method as soon as $\hat{V}(x^k) \leq \varepsilon$ with $\varepsilon := 10^{-8}$, where $\hat{V}(x) := \hat{V}_\alpha(x)$ is the function from (12) with $\alpha = 0$. Hence the termination criterion for the relaxation method is similar to the one used before in our method, although they are not directly comparable.

Table 6 compares the number of iterations required by the relaxation method applied to Examples 5.1–5.4 with our unconstrained minimization approach.

Method	Ex. 5.1	Ex. 5.2	Ex. 5.3	Ex. 5.4	Ex. 5.5
Relaxation method	20	26	34	40	50
Our method	3	3	18	13	3

Table 6: Number of iterations of relaxation method and our method

To put the results from Table 6 in the right perspective, one should take into account that the main computational burden of the relaxation method is the solution of one constrained optimization problem, whereas our method has to solve two constrained optimization problems at each iteration. The remaining overhead of both methods is neglectable. However, even if we multiply the iteration numbers of our method with the factor two, we are usually still better (usually much better) than the relaxation method. Moreover, it is interesting to observe that the relaxation method takes relatively many iterations for Example 5.5 where the starting point was chosen close to the solution.

6 Final Remarks

We presented some optimization reformulations of the generalized Nash equilibrium problem such that the global minima of these reformulations correspond to (normalized) Nash equilibria of the underlying game. Two of these optimization reformulations were smooth with a continuously differentiable objective function, and their stationary points were shown to be global minima under suitable assumptions.

One of our future projects is to establish conditions under which the smooth objective functions have a (strongly) semismooth gradient in the sense of [30, 31, 29]. The advantage of such a result would be that it would allow the application of nonsmooth Newton methods to the corresponding optimization reformulation which then should be locally superlinearly or quadratically convergent under certain assumptions.

Acknowledgement. The authors would like to thank Masao Fukushima and Diethard Klatte for some very helpful comments on a preliminary version of this paper.

References

[1] E. Adida and G. Perakis: *Dynamic pricing and inventory control: Uncertainty and competition. Part A: Existence of Nash equilibrium.* Technical Report, Operations Research Center, Sloan School of Management, MIT, January 2006.

[2] E. Adida and G. Perakis: *Dynamic pricing and inventory control: Uncertainty and competition. Part B: An algorithm for the normalized Nash equilibrium.* Technical Report, Operations Research Center, Sloan School of Management, MIT, December 2005.

[3] T. Başar: *Relaxation techniques and asynchronous algorithms for on-line computation of non-cooperative equilibria.* Journal of Economic Dynamics and Control 11, 1987, pp. 531–549.

[4] J. Barzilai and J.M. Borwein: *Two point step size gradient method.* IMA Journal on Numerical Analysis 8, 1988, pp. 141–148.

[5] A. Bensoussan: *Points de Nash dans le cas de fonctionelles quadratiques et jeux differentiels lineaires a N personnes.* SIAM Journal on Control 12, 1974, pp. 460–499.

[6] J. Contreras, M. Klusch, and J.B. Krawczyk: *Numerical solutions to Nash-Cournot equilibria in coupled constraint electricity markets.* IEEE Transactions on Power Systems 19, 2004, pp. 195–206.

[7] Y.H. Dai and L.Z. Liao: *R-linear convergence of the Barzilai and Borwein gradient method.* IMA Journal on Numerical Analysis 22, 2002, pp. 1–10.

[8] F. Facchinei and J.-S. Pang: *Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume I.* Springer, New York, NY, 2003.

[9] F. Facchinei and J.-S. Pang: *Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume II.* Springer, New York, NY, 2003.

[10] F. Facchinei, A. Fischer, and V. Piccialli: *On generalized Nash games and VIs.* Technical Report, Department of Computer and System Sciences "A. Ruberti", Università di Roma "La Sapienza", Rome, Italy, January 2006.

[11] F. Facchinei, A. Fischer, and V. Piccialli: *Generalized Nash equilibrium problems and Newton methods.* Technical Report, Department of Computer and System Sciences "A. Ruberti", Università di Roma "La Sapienza", Rome, Italy, February 2006.

[12] S.D. Flåm and A.S. Antipin: *Equilibrium programming using proximal-like algorithms.* Mathematical Programming 78, 1997, pp. 29–41.

[13] S.D. Flåm and A. Ruszczyński: *Noncooperative convex games: Computing equilibria by partial regularization.* IIASA Working Paper 94-42, Laxenburg, Austria, May 1994.

[14] R. Fletcher: *On the Barzilai-Borwein method.* Numerical Analysis Report NA/207, Department of Mathematics, University of Dundee, Dundee, United Kingdom, 2001.

[15] M. Fukushima: *Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems.* Mathematical Programming 53, 1992, pp. 99–110.

[16] L. Grippo and M. Sciandrone: *Nonmonotone globalization techniques for the Barzilai-Borwein gradient method.* Computational Optimization and Applications 23, 2002, pp. 143–169.

[17] G. Gürkan and J.-S. Pang: *Approximations of Nash equilibria.* Technical Report, Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY, January 2006.

[18] P.T. Harker: *Generalized Nash games and quasivariational inequalities.* European Journal of Operations Research 54, 1991, pp. 81–94.

[19] B.F. Hobbs: *Linear complementarity models of Nash-Cournot competition in bilateral and POOLCO power markets.* IEEE Transactions on Power Systems 16, 2001, pp. 194–202.

[20] W.W. Hogan: *Point-to-set maps in mathematical programming.* SIAM Review 15, 1973, pp. 591–603.

[21] C. Kanzow and M. Fukushima: *Theoretical and numerical investigation of the D-gap function for box constrained variational inequalities.* Mathematical Programming 83, 1998, pp. 55–87.

[22] A. Kesselman, S. Leonardi, and V. Bonifaci: *Game-theoretic analysis of internet switching with selfish users.* Lecture Notes in Computer Science 3828, 2005, pp. 236–245.

[23] J.B. Krawczyk: *Coupled constraint Nash equilibria in environmental games.* Resource and Energy Economics 27, 2005, pp. 157–181.

[24] J.B. Krawczyk and S. Uryasev: *Relaxation algorithms to find Nash equilibria with economic applications.* Environmental Modeling and Assessment 5, 2000, pp. 63–73.

[25] S. Li and T. Başar: *Distributed algorithms for the computation of noncooperative equilibria.* Automatica 23, 1987, pp. 523–533.

[26] G. Mastroeni: *Gap functions for equilibrium problems.* Journal of Global Optimization 27, 2003, pp. 411–426.

[27] H. Nikaido and K. Isoda: *Note on noncooperative convex games.* Pacific Journal of Mathematics 5, 1955, pp. 807–815.

[28] J.-S. Pang and M. Fukushima: *Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games.* Computational Management Science 2, 2005, pp. 21–56.

[29] J.-S. Pang and L. Qi: *Nonsmooth equations: motivation and algorithms.* SIAM Journal on Optimization 3, 1993, pp. 443–465.

[30] L. Qi: *Convergence analysis of some algorithms for solving nonsmooth equations.* Mathematics of Operations Research 18, 1993, pp. 227–244.

[31] L. Qi and J. Sun: *A nonsmooth version of Newton's method.* Mathematical Programming 58, 1993, pp. 353–367.

[32] J.-M. Peng: *Equivalence of variational inequality problems to unconstrained optimization.* Mathematical Programming 78, 1997, pp. 347–356.

[33] M. Raydan: *On the Barzilai and Borwein choice of the steplength for the gradient method.* IMA Journal on Numerical Analysis 13, 1993, pp. 321–326.

[34] M. Raydan: *The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem.* SIAM Journal on Optimization 7, 1997, pp. 26–33.

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

[36] S. Uryasev and R.Y. Rubinstein: *On relaxation algorithms in computation of noncooperative equilibria.* IEEE Transactions on Automatic Control 39, 1994, pp. 1263–1267.

[37] N. Yamashita, K. Taji, and M. Fukushima: *Unconstrained optimization reformulations of variational inequality problems.* Journal of Optimization Theory and Applications 92, 1997, pp. 439–456.

[38] L. Zhang and J. Han: *Unconstrained optimization reformulation of equilibrium problems.* Technical Report, Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing, China, 2006.