Comput Optim Appl (2011) 50:23–48
DOI 10.1007/s10589-009-9314-x

Nonsmooth optimization reformulations characterizing all solutions of jointly convex generalized Nash equilibrium problems

Axel Dreves · Christian Kanzow

Received: 22 April 2009 / Published online: 5 January 2010

Abstract Generalized Nash equilibrium problems (GNEPs) allow, in contrast to standard Nash equilibrium problems, a dependence of the strategy space of one player from the decisions of the other players. In this paper, we consider jointly convex GNEPs which form an important subclass of the general GNEPs. Based on a regularized Nikaido-Isoda function, we present two (nonsmooth) reformulations of this class of GNEPs, one reformulation being a constrained optimization problem and the other one being an unconstrained optimization problem. While most approaches in the literature compute only a so-called normalized Nash equilibrium, which is a subset of all solutions, our two approaches have the property that their minima characterize the set of all solutions of a GNEP. We also investigate the smoothness properties of our two optimization problems and show that both problems are continuous under a Slater-type condition and, in fact, piecewise continuously differentiable under the constant rank constraint qualification. Finally, we present some numerical results based on our unconstrained optimization reformulation.

Keywords Generalized Nash equilibrium problem · Jointly convex · Optimization reformulation · Continuity · PC^1 mapping · Semismoothness · Constant rank constraint qualification

1 Introduction

This paper considers the generalized Nash equilibrium problem, GNEP for short, with N players $\nu = 1, \dots, N$. Each player $\nu \in \{1, \dots, N\}$ controls the variables

A. Dreves · C. Kanzow (✉)
Institute of Mathematics, University of Würzburg, Am Hubland, 97074 Würzburg, Germany
e-mail: kanzow@mathematik.uni-wuerzburg.de

A. Dreves
e-mail: dreves@mathematik.uni-wuerzburg.de

$x^\nu \in \mathbb{R}^{n_\nu}$, and the vector $x = (x^1, \dots, x^N)^T \in \mathbb{R}^n$ with $n = n_1 + \cdots + n_N$ describes the decision vector of all players. To emphasize the role of player ν's variables within the vector x, we often write $(x^\nu, x^{-\nu})$ for this vector. Each player has a cost function $\theta_\nu : \mathbb{R}^n \to \mathbb{R}$ and, in the most general setting of a GNEP, its own strategy space $X_\nu(x^{-\nu}) \subseteq \mathbb{R}^{n_\nu}$ that depends on the other players. Typically, these sets are defined explicitly via some constraint functions, say

$$X_\nu(x^{-\nu}) := \{x^\nu \in \mathbb{R}^{n_\nu} \mid g^\nu(x^\nu, x^{-\nu}) \le 0\} \tag{1}$$

for suitable functions $g^\nu : \mathbb{R}^n \to \mathbb{R}^{m_\nu}$, $\nu = 1, \dots, N$. Let

$$\Omega(x) := X_1(x^{-1}) \times \cdots \times X_N(x^{-N}) \tag{2}$$

be the Cartesian product of these strategy spaces. Then a vector $x^* \in \Omega(x^*)$ is called a *generalized Nash equilibrium*, or simply a *solution* of the GNEP, if

$$\theta_\nu(x^{*,\nu}, x^{*,-\nu}) \le \theta_\nu(x^\nu, x^{*,-\nu}) \quad \text{for all } x^\nu \in X_\nu(x^{*,-\nu})$$

holds for all players $\nu = 1, \dots, N$, i.e. if $x^{*,\nu}$ solves the optimization problem

$$\min_{x^\nu} \theta_\nu(x^\nu, x^{*,-\nu}) \quad \text{s.t.} \quad x^\nu \in X_\nu(x^{*,-\nu})$$

for all $\nu = 1, \dots, N$. There are just a very few papers that deal with a GNEP in this general setting (see, in particular, [5, 6, 8, 9, 24]) where the feasible sets (besides their dependence on the rivals' strategies) are allowed to be different for each player. Here we consider the special case that is often called the *jointly convex case*. Therein we assume that there is a common strategy space $X \subseteq \mathbb{R}^n$ such that the feasible set of player ν is given by

$$X_\nu(x^{-\nu}) = \{x^\nu \in \mathbb{R}^{n_\nu} \mid (x^\nu, x^{-\nu}) \in X\}.$$

Throughout this paper, we assume that the following standard requirements are satisfied.

Assumption 1.1

(a) The cost functions $\theta_\nu : \mathbb{R}^n \to \mathbb{R}$ are continuous and, as a function of x^ν alone, convex.
(b) The set $X \subseteq \mathbb{R}^n$ is nonempty, closed and convex, and can be represented as $X = \{x \in \mathbb{R}^n \mid g(x) \le 0\}$ with a mapping $g : \mathbb{R}^n \to \mathbb{R}^m$ whose component functions g_i are convex for all $i = 1, \dots, m$.

Note that we do not require compactness of the set X. In view of Assumption 1.1, the strategy space of player ν is given by

$$X_\nu(x^{-\nu}) = \{x^\nu \in \mathbb{R}^{n_\nu} \mid g(x^\nu, x^{-\nu}) \le 0\}.$$

In the setting (1), this corresponds to the case where $g^1 = g^2 = \cdots = g^N = g$.

Note that this jointly convex case is still a very challenging problem. Although a number of methods have been developed for this problem during the last few years (see, in particular, [5] and references therein), most of these methods find a so-called normalized Nash equilibrium of the GNEP. Each normalized Nash equilibrium is, in particular, a solution of the GNEP, so these methods can be used to find a generalized Nash equilibrium, but the converse is not true in general. In fact, typically a GNEP has many solutions, but just one normalized Nash equilibrium. Unfortunately, this normalized Nash equilibrium is often not the solution economists, etc. are interested in. This observation is not new, and there exists a small number of approaches which try to deal with this problem.

One is described in the book [23], but only for the standard Nash equilibrium problem where the strategy spaces $X_\nu(x^{-\nu})$ do not depend on the rivals' strategies (making the entire problem considerably easier). The idea from [23] is to find specific solutions of a standard Nash problem by using a bilevel formulation which results into a mathematical program with equilibrium constraint (MPEC). Such an MPEC, however, is rather difficult to solve.

Another approach for GNEPs is very recent, see [22], and tries to use characterizations of all the solutions of a GNEP via certain parameterized variational inequality problems. The idea is quite nice and has the advantage of giving a smooth formulation, but a complete characterization of the whole solution set is not given, at least not without additional assumptions.

A classical idea to find a characterization of all solutions of a GNEP (under certain constraint qualifications) is to write down the KKT conditions for each player. This results in a large mixed complementarity system that can, in principle, be attacked by a semismooth solver, for example. However, this approach has some singularity problems since the joint constraints appear several times (one for each player). Deleting all the repeated constraints, on the other hand, gives an underdetermined system which again causes some troubles, see, in particular, the discussion in [4]. Moreover, this KKT approach involves both the original variables of the players and additional Lagrange multipliers. As a consequence, it is difficult to show that suitable methods generate bounded iterates (whereas this is trivial in our case at least for X being bounded, cf. the bounded level set result in [13]).

Another way to reformulate the GNEP in such a way that the solution sets coincide is to use the quasi-variational inequality (QVI) approach, see [11], for example. It has the advantage that it uses the original variables only, but a detailed discussion of this QVI approach in the context of GNEPs is still missing. In general, however, one can say that there are not many numerical methods known for the general QVI problem, and it is not clear how they work in practice when applied to GNEPs. In any case, this approach has differentiability problems due to projections onto sets that depend on the variable x. It might be possible, however, to overcome these problems under the assumptions that are also used in this paper for our optimization technique.

The approach we follow here was already settled in the paper [13], but not further discussed there simply because the focus on that paper was on some other (differentiable) formulations of a GNEP. The idea is to use a constrained optimization reformulation of the GNEP whose minima characterize the entire set of generalized Nash equilibria, and not only the normalized Nash equilibria. The price we have to pay is that this constrained optimization reformulation is nonsmooth. The precise reformulation and its elementary properties will be discussed in detail in Sect. 2. There, we also modify the constrained optimization reformulation in a suitable way to obtain a new unconstrained optimization reformulation of the GNEP whose solutions are, again, precisely the generalized Nash equilibria of the GNEP. The exact smoothness properties of these two reformulations, the constrained and the unconstrained optimization one, will be discussed in detail in Sects. 3 and 4, respectively. It turns out that both formulations are continuous in those points x where a Slater-condition for the sets $\Omega(x)$ holds. Moreover, it will be shown that the objective functions are PC^1 mappings under the additional assumption that the constant rank constraint qualification holds. This, in particular, implies that the functions are directionally differentiable, locally Lipschitz and semismooth. This paves the way for the application of suitable nonsmooth optimization solvers in order to find generalized Nash equilibria. Based on the unconstrained reformulation, we therefore present some numerical results in Sect. 5 using a sampling method from [1] for nonsmooth optimization. We then close with some final remarks in Sect. 6.

Notation: With $\|\cdot\|$ we denote the Euclidean norm. $P_X[x]$ is the (Euclidean) projection of a vector $x \in \mathbb{R}^n$ onto the nonempty, closed and convex set $X \subseteq \mathbb{R}^n$, i.e. it is the unique solution of

$$\min \frac{1}{2}\|z - x\|^2 \quad \text{s.t.} \quad z \in X.$$

A function $G : \mathbb{R}^n \to \mathbb{R}^m$ is called a PC^1 function in a neighbourhood of a given point x^* if G is continuous and there exists a neighborhood U of x^* and a finite number of continuously differentiable functions $G_1, G_2, \ldots, G_k$ defined on U such that, for all $x \in U$, we have $G(x) \in \{G_1(x), G_2(x), \ldots, G_k(x)\}$. For a locally Lipschitz function $H : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$, $(x, y) \mapsto H(x, y)$, we denote by $\partial H(x, y)$ the generalized Jacobian of H in the sense of Clarke [3], and by $\pi_y \partial H(x, y)$ the set of all matrices $M \in \mathbb{R}^{n\times n}$ such that, for a matrix $N \in \mathbb{R}^{n\times m}$, the matrix $[N, M] \in \mathbb{R}^{n\times(m+n)}$ is an element of $\partial H(x, y)$.

2 Constrained and unconstrained optimization reformulation

Here we first recall a constrained optimization reformulation of the GNEP as introduced in [13] and then present a new reformulation of the GNEP as an unconstrained optimization problem.

To this end, we first define the *Nikaido-Isoda function* (also called *Ky Fan-function*) by

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

Since θ_ν is convex in x^ν, it follows that $\Psi(x,.)$ is concave for any fixed x. Consequently, the *regularized Nikaido-Isoda-function*

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

originally introduced in [10] as a technical tool for the standard Nash equilibrium problem and afterwards used in [13–16] for the numerical solution of GNEPs, is uniformly concave as a function of the second argument, where $\alpha > 0$ denotes a fixed parameter. Using this function, we define

$$\begin{aligned} V_\alpha(x) &:= \max_{y\in\Omega(x)} \Psi_\alpha(x,y) \\ &= \max_{y\in\Omega(x)} \sum_{\nu=1}^{N} \Big[\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu}) - \frac{\alpha}{2}\|x^\nu - y^\nu\|^2\Big] \\ &= \sum_{\nu=1}^{N} \Big[\theta_\nu(x^\nu, x^{-\nu}) - \min_{y^\nu\in X_\nu(x^{-\nu})} \Big(\theta_\nu(y^\nu, x^{-\nu}) + \frac{\alpha}{2}\|x^\nu - y^\nu\|^2\Big)\Big], \end{aligned} \tag{3}$$

where the maximization is taken over the set $\Omega(x)$ defined in (2). Note that Assumption 1.1 implies that all sets $X_\nu(x^{-\nu})$ are closed and convex, hence $\Omega(x)$ is also closed and convex. Therefore, $V_\alpha(x)$ is well-defined for all $x \in \mathbb{R}^n$ such that $\Omega(x) \neq \emptyset$. According to the following result, the latter condition holds at least for all $x \in X$.

As shown in [13], there is a reformulation of the jointly convex GNEP as a constrained optimization problem based on the mapping V_α. The following is a summary of the corresponding results from [13].

Theorem 2.1 *Suppose that Assumption 1.1 holds. Then:*

(a) *$x \in X$ if and only if $x \in \Omega(x)$.*
(b) *$V_\alpha(x) \geq 0$ for all $x \in X$.*
(c) *x^* is a generalized Nash equilibrium if and only if $x^* \in X$ and $V_\alpha(x^*) = 0$.*
(d) *For all $x \in \mathbb{R}^n$ with $\Omega(x) \neq \emptyset$, there exists a unique vector $y_\alpha(x) := (y_\alpha^1(x), \ldots, y_\alpha^N(x))$ such that, for every $\nu = 1, \ldots, N$,*

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

(e) *x^* is a generalized Nash equilibrium if and only if x^* is a fixed point of the mapping $x \mapsto y_\alpha(x)$, i.e. if and only if $x^* = y_\alpha(x^*)$.*

Basically, this result says that finding a solution (i.e., an arbitrary generalized Nash equilibrium) of the GNEP is equivalent to solving the constrained optimization problem

$$\min V_\alpha(x) \quad \text{s.t.} \quad x \in X \tag{4}$$

with $V_\alpha(x) = 0$. Unfortunately, it turns out that this optimization problem has a non-smooth objective function even under very strong conditions. This observation was already made in [13], so that this reformulation was not further investigated there. The following example shows that V_α might even be discontinuous.

Example 2.2 Let the common strategy space of a two-player game be given by

$$X = \{x \in \mathbb{R}^3 \mid x_2^2 + (x_3 - x_1)^2 - x_1^2 \le 0, 0 \le x_1 \le 10, -10 \le x_2 \le 10, 0 \le x_3 \le 20\}.$$

The variable x_1 is controlled by the first player, and the two variables x_2, x_3 are the decision variables of the second player. The cost functions are defined by

$$\theta_1(x) := (x_1 + 10)^2 \quad \text{and} \quad \theta_2(x) := x_2^2 + x_3^2,$$

respectively. The corresponding Nikaido-Isoda function is given by

$$\Psi_\alpha(x, y) := (x_1 + 10)^2 + x_2^2 + x_3^2 - (y_1 + 10)^2 - y_2^2 - y_3^2 - \frac{\alpha}{2}\|x - y\|^2.$$

Its unconstrained maximum is $(\frac{-20+\alpha x_1}{2+\alpha}, \frac{\alpha x_2}{2+\alpha}, \frac{\alpha x_3}{2+\alpha})^T$. Now consider the sequence

$$x(\delta) := (10, \sqrt{20\delta - \delta^2}, \delta)^T \to (10, 0, 0)^T := x^*$$

with $\delta \downarrow 0$ (note that x^* belongs to X). Then an elementary calculation shows that, for all $\alpha > 0$ and all $\delta > 0$ sufficiently small, we have

$$y_\alpha\big(x(\delta)\big) = \left(10, \frac{\alpha\sqrt{20\delta - \delta^2}}{2+\alpha}, \frac{\alpha\delta}{2+\alpha}\right)^T \to (10, 0, 0) \quad \text{for } \delta \downarrow 0.$$

On the other hand, for the parameter $\alpha = 2$ or, more generally, for an arbitrary parameter $\alpha \in (0, 2]$, it can be shown that $y_\alpha(x^*) = (0, 0, 0)^T$, hence the function y_α is not continuous in $(10, 0, 0)^T$. Furthermore, we have

$$V_\alpha\big(x(\delta)\big) = \Psi_\alpha\big(x(\delta), y_\alpha(x(\delta))\big) = 20\delta\left(1 - \frac{\alpha^2}{(2+\alpha)^2} - \frac{\alpha}{2}\left(1 - \frac{\alpha}{2+\alpha}\right)^2\right) \to 0,$$

whereas $V_\alpha(x^*) = 20^2 - 10^2 - \frac{\alpha}{2}10^2 \neq 0$, which shows that V_α is not continuous in $(10, 0, 0)^T$. This example also shows that the Slater condition for the set X, i.e. the existence of an interior point of X, is not sufficient for continuity of V_α, since for example $\hat{x} := (2, 1, 2)^T$ is a Slater point.

Besides this negative observation, it turns out that the function V_α is continuous and even a PC^1 mapping under fairly mild conditions. This will be discussed in more detail in Sect. 3. Here, we now modify the previous approach and present a new unconstrained optimization reformulation of the GNEP which also characterizes all solutions of the GNEP.

In order to present an unconstrained reformulation of the GNEP which is close to the previous constrained one, we have to find a way to define the function

$V_\alpha(x) := \max_{y\in\Omega(x)} \Psi_\alpha(x,y)$ for those points $x \in \mathbb{R}^n$ where $\Omega(x)$ is empty. So far, we only know that $\Omega(x) \neq \emptyset$ for all $x \in X$. This fact will now be exploited in the following definition where, for an arbitrary $x \in \mathbb{R}^n$ (not necessarily belonging to X), we maximize over the set $\Omega(P_X[x])$ instead of $\Omega(x)$.

Definition 2.3 For all $x \in \mathbb{R}^n$ and $\alpha > 0$, we define

$$\bar{y}_\alpha(x) := \arg\max_{y\in\Omega(P_X[x])} \Psi_\alpha(x,y) \quad \text{and}$$

$$\bar{V}_\alpha(x) := \max_{y\in\Omega(P_X[x])} \Psi_\alpha(x,y) = \Psi_\alpha(x,\bar{y}_\alpha(x)).$$

Given two parameters $0 < \alpha < \beta$, we then define

$$\bar{V}_{\alpha\beta}(x) := \bar{V}_\alpha(x) - \bar{V}_\beta(x)$$

for all $x \in \mathbb{R}^n$ (where $\bar{y}_\beta(x)$ and $\bar{V}_\beta(x)$ are defined in an obvious way).

For all $x \in X$, we obviously have $\bar{y}_\alpha(x) = y_\alpha(x)$ and $\bar{V}_\alpha(x) = V_\alpha(x)$, so we leave the functions unchanged on X. On the other hand, for $x \notin X$, all functions are still well-defined since our previous discussion shows that, in particular, $P_X[x] \in \Omega(P_X[x])$, hence $\Omega(P_X[x]) \neq \emptyset$ and, therefore, $\bar{y}_\alpha(x)$ is well-defined and unique.

The next lemma will be crucial to prove that we get an unconstrained reformulation of the GNEP by the function $\bar{V}_{\alpha\beta}$.

Lemma 2.4 *For all $x \in \mathbb{R}^n$, the following inequalities hold*:

$$\frac{\beta-\alpha}{2}\|x-\bar{y}_\beta(x)\|^2 \leq \bar{V}_{\alpha\beta}(x) \leq \frac{\beta-\alpha}{2}\|x-\bar{y}_\alpha(x)\|^2.$$

Proof We have $\bar{y}_\alpha(x) \in \Omega(P_X[x])$ and $\bar{y}_\beta(x) \in \Omega(P_X[x])$. Therefore

$$\bar{V}_\beta(x) = \Psi_\beta(x,\bar{y}_\beta(x)) = \max_{y\in\Omega(P_X[x])} \Psi_\beta(x,y) \geq \Psi_\beta(x,\bar{y}_\alpha(x)), \tag{5}$$

$$\bar{V}_\alpha(x) = \Psi_\alpha(x,\bar{y}_\alpha(x)) = \max_{y\in\Omega(P_X[x])} \Psi_\alpha(x,y) \geq \Psi_\alpha(x,\bar{y}_\beta(x)). \tag{6}$$

On the one hand, this implies

$$\bar{V}_{\alpha\beta}(x) = \bar{V}_\alpha(x) - \bar{V}_\beta(x) \overset{(5)}{\leq} \Psi_\alpha(x,\bar{y}_\alpha(x)) - \Psi_\beta(x,\bar{y}_\alpha(x)) = \frac{\beta-\alpha}{2}\|x-\bar{y}_\alpha(x)\|^2,$$

and, on the other hand, we obtain

$$\bar{V}_{\alpha\beta}(x) = \bar{V}_\alpha(x) - \bar{V}_\beta(x) \overset{(6)}{\geq} \Psi_\alpha(x,\bar{y}_\beta(x)) - \Psi_\beta(x,\bar{y}_\beta(x)) = \frac{\beta-\alpha}{2}\|x-\bar{y}_\beta(x)\|^2$$

for all $x \in \mathbb{R}^n$. □

We are now in a position to show that the function $\bar{V}_{\alpha\beta}$ provides an unconstrained optimization reformulation of the GNEP.

Theorem 2.5 *The following statements hold*:

(a) $\bar{V}_{\alpha\beta}(x) \geq 0$ *for all* $x \in \mathbb{R}^n$.
(b) x^* *is a generalized Nash equilibrium if and only if* x^* *is a minimum of* $\bar{V}_{\alpha\beta}$ *with* $\bar{V}_{\alpha\beta}(x^*) = 0$.

Proof Lemma 2.4 (left inequality) shows that $\bar{V}_{\alpha\beta}(x) \geq \frac{\beta-\alpha}{2}\|x - \bar{y}_\beta(x)\|^2 \geq 0$ for all $x \in \mathbb{R}^n$, hence statement (a) holds.

In order to verify the second statement, first assume that x^* is a generalized Nash equilibrium. Then $x^* \in \Omega(x^*)$, and Theorem 2.1 (a) implies $x^* \in X$. This, in turn, gives $P_X[x^*] = x^*$, and together with the fixed point characterization of Theorem 2.1 (e), we get $x^* = y_\alpha(x^*) = \bar{y}_\alpha(x^*)$. Lemma 2.4 (right inequality) then implies $\bar{V}_{\alpha\beta}(x^*) \leq 0$. In view of part (a), we therefore have $\bar{V}_{\alpha\beta}(x^*) = 0$.

Conversely, assume that $\bar{V}_{\alpha\beta}(x^*) = 0$ for some $x^* \in \mathbb{R}^n$. Then we obtain

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

from Lemma 2.4. Consequently, we have $x^* = \bar{y}_\beta(x^*) \in \Omega(P_X[x^*])$, i.e.

$$x^{*,\nu} \in X_\nu((P_X[x^*])^{-\nu}) = \{x^\nu \mid (x^\nu, (P_X[x^*])^{-\nu}) \in X\}$$

for all $\nu = 1, \ldots, N$. Let $\bar{\nu} \in \{1, \ldots, N\}$ be arbitrarily given.

Then we have $(x^{*,\bar{\nu}}, (P_X[x^*])^{-\bar{\nu}}) \in X$ and

$$\begin{aligned}\|x^* - (x^{*,\bar{\nu}}, (P_X[x^*])^{-\bar{\nu}})\|^2 &= \sum_{\nu=1,\nu\neq\bar{\nu}}^{N} \|x^{*,\nu} - (P_X[x^*])^\nu\|^2 \\ &\leq \sum_{\nu=1}^{N} \|x^{*,\nu} - (P_X[x^*])^\nu\|^2 = \|x^* - P_X[x^*]\|^2.\end{aligned}$$

Since the projection $P_X[x^*]$ onto the nonempty, closed and convex set X is the unique solution of the problem

$$\min \frac{1}{2}\|x^* - z\|^2 \quad \text{s.t.} \quad z \in X,$$

we must have $x^{*,\bar{\nu}} = (P_X[x^*])^{\bar{\nu}}$. Since $\bar{\nu} \in \{1, \ldots, N\}$ was arbitrarily chosen, this is true for all components and hence $x^* = P_X[x^*]$, i.e. $x^* \in X$. Thus we get $y_\beta(x^*) = \bar{y}_\beta(x^*) = x^*$. Therefore, x^* is a generalized Nash equilibrium by the fixed point characterization from Theorem 2.1 (e). □

This theorem shows that the generalized Nash equilibria x^* are exactly the minima of the function $\bar{V}_{\alpha\beta}$ satisfying $\bar{V}_{\alpha\beta}(x^*) = 0$. We therefore have the unconstrained optimization reformulation

$$\min \bar{V}_{\alpha\beta}(x), \quad x \in \mathbb{R}^n, \tag{7}$$

in order to find solutions of a GNEP. Again, the minima of this problem (with zero objective function value) characterize all solutions of the GNEP (not only the normalized Nash equilibria). However, similar to the constrained reformulation, also this unconstrained one is nondifferentiable in general. The smoothness properties of this unconstrained problem will be discussed in more detail in Sect. 4.

3 Smoothness properties of the constrained reformulation

Here we come back to the constrained reformulation (4) of the GNEP with the objective function V_α from (3). Knowing that this objective function is nondifferentiable, we take a closer look at the smoothness properties of this mapping. Our aim is to show the following statements:

- V_α is continuous at $x \in X$ provided that $\Omega(x)$ satisfies a Slater condition;
- V_α is a PC^1 function provided that g and θ_ν are twice continuously differentiable and, in addition to the Slater condition, also a constant rank constraint qualification holds.

In order to verify the continuity of V_α, we need some terminology and results from set-valued analysis. Let us begin with the following well known definitions.

Definition 3.1 Suppose $X \subseteq \mathbb{R}^n$, $Y \subseteq \mathbb{R}^m$, and $\Phi : X \rightrightarrows Y$ is a point-to-set mapping. Then Φ is called

(a) *lower semicontinuous* in $x^* \in X$, if for all sequences $\{x^k\} \subseteq X$ with $x^k \to x^*$ and all $y^* \in \Phi(x^*)$, there exists a number $m \in \mathbb{N}$ and a sequence $\{y^k\} \subseteq Y$ such that $y^k \in \Phi(x^k)$ for all $k \geq m$ and $y^k \to y^*$;
(b) *closed* in $x^* \in X$, if for all sequences $\{x^k\} \subseteq X$ with $x^k \to x^*$ and all sequences $y^k \to y^*$ with $y^k \in \Phi(x^k)$ for all $k \in \mathbb{N}$ sufficiently large, we have $y^* \in \Phi(x^*)$;
(c) *lower semicontinuous* or *closed on* X if it is lower semicontinuous or closed in every $x \in X$.

The definition of a lower semicontinuous set-valued mapping is in the sense of Berge. Alternative names used in the literature are "open mapping" (see [17]) and "inner semicontinuous mapping" (see [27]). A useful result for our subsequent analysis is the following one which is an immediate consequence of [17, Corollaries 8.1 and 9.1].

Lemma 3.2 *Let $X \subseteq \mathbb{R}^n$ arbitrary, $Y \subseteq \mathbb{R}^m$ convex, and $f : X \times Y \to \mathbb{R}$ be concave in y for fixed x and continuous on $X \times Y$. Let $\Phi : X \rightrightarrows Y$ be a point-to-set map which is closed in a neighborhood of $\bar{x}$ and lower semicontinuous in $\bar{x}$, and $\Phi(x)$ convex in a neighbourhood of $\bar{x}$. Define*

$$Y(x) := \left\{ z \in \Phi(x) \mid \sup_{y \in \Phi(x)} f(x, y) = f(x, z) \right\}$$

and assume that $Y(\bar{x})$ has exactly one element. Then the point-to-set mapping $x \mapsto Y(x)$ is lower semicontinuous and closed in $\bar{x}$.

We can use Lemma 3.2 to prove continuity of V_α.

Theorem 3.3 *Suppose that Assumption 1.1 holds and that the point-to-set mapping $x \to \Omega(x)$ from (2) is closed on X and lower semicontinuous in $x \in X$. Then the functions y_α and V_α are continuous at $x \in X$.*

Proof Assumption 1.1 implies that the function $\Psi_\alpha(x, .)$ is concave for fixed x and continuous on $\mathbb{R}^n \times \mathbb{R}^n$. By Theorem 2.1 (a), $\Omega(x)$ is nonempty for all $x \in X$, and Theorem 2.1 (d) shows that the sets $Y_\alpha(x) := \{z \in \Omega(x) \mid \sup_{y\in\Omega(x)} \Psi_\alpha(x, y) = \Psi_\alpha(x, z)\}$ consist of exactly one element for all $x \in X$, namely $y_\alpha(x)$. Taking into account the convexity of $\Omega(x)$, Lemma 3.2 therefore implies that $x \to \{y_\alpha(x)\}$, viewed as a point-to-set mapping, is lower semicontinuous and closed at $x \in X$. This implies that the single-valued function $x \mapsto y_\alpha(x)$ is continuous at x. Hence, the composition $V_\alpha(x) = \Psi_\alpha(x, y_\alpha(x))$ is also continuous at x. □

Theorem 3.3 shows that the continuity of the functions y_α and V_α follows immediately if we can show that the set-valued mapping $x \mapsto \Omega(x)$ is lower semicontinuous and closed. The following result states that this mapping is always closed on X.

Lemma 3.4 *Suppose that Assumption 1.1 holds. Then the point-to-set mapping $x \mapsto \Omega(x)$ is closed on X.*

Basically, Lemma 3.4 follows from the fact that X is closed. We therefore skip the proof of Lemma 3.4.

Next we want to show that the point-to-set mapping $x \mapsto \Omega(x)$ is also lower semicontinuous. To this end, it will be useful to define the function

$$h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{mN} \quad \text{by} \quad h(x, y) := \begin{pmatrix} g(y^1, x^{-1}) \\ \vdots \\ g(y^N, x^{-N}) \end{pmatrix},$$

where g is the mapping from Assumption 1.1. The function h has the following obvious properties:

- h is locally Lipschitz continuous (since all g_i are convex);
- The component functions $h_i(x, \cdot)$ are convex in y for any given x;
- For any given x, we have $y \in \Omega(x) \iff h(x, y) \leq 0$.

In view of Theorem 3.3 and (the Counter-) Example 2.2, it is clear that we cannot expect lower semicontinuity of $x \mapsto \Omega(x)$ without any further condition. The missing assumption is the *Slater condition* for the set $\Omega(x) = \{y \in \mathbb{R}^n \mid h(x, y) \leq 0\}$ saying that, for the given vector x, there exists a vector $\hat{y} \in \mathbb{R}^n$ with $h(x, \hat{y}) < 0$.

Lemma 3.5 *Suppose that Assumption 1.1 holds. Then the point-to-set mapping $x \mapsto \Omega(x)$ is lower semicontinuous in every $x \in X$ where $\Omega(x)$ satisfies the Slater condition.*

Proof Let $x^* \in X$ be given, such that the Slater condition holds with $\hat{y} \in \Omega(x^*)$, i.e. $h(x^*, \hat{y}) < 0$. Consider an arbitrary sequence $\{x^k\} \subseteq X$ converging to x^*, and let $y^* \in \Omega(x^*)$ and hence $h(x^*, y^*) \leq 0$ be given. To prove lower semicontinuity of $x \mapsto \Omega(x)$ in $x = x^*$, we have to show the existence of a sequence $\{y^k\}$ converging to y^* with $y^k \in \Omega(x^k)$ for k sufficiently large. To this end, let us define $y^k := t_k\hat{y} + (1 - t_k)y^*$ with a suitable sequence $\{t_k\} \downarrow 0$. Then we obviously obtain $y^k \to y^*$. By convexity and the local Lipschitz property of the function h, we obtain for all $i = 1, \dots, mN$

$$\begin{aligned}
h_i(x^k, y^k) &= h_i\left(x^k, t_k\hat{y} + (1 - t_k)y^*\right) \\
&\leq t_k h_i(x^k, \hat{y}) + (1 - t_k)h_i(x^k, y^*) \\
&= t_k\left(h_i(x^k, \hat{y}) - h_i(x^*, \hat{y})\right) + t_k h_i(x^*, \hat{y}) \\
&\quad + (1 - t_k)\left(h_i(x^k, y^*) - h_i(x^*, y^*)\right) + (1 - t_k)h_i(x^*, y^*) \\
&\leq t_k L_i^1\|x^k - x^*\| + t_k h_i(x^*, \hat{y}) + (1 - t_k)L_i^2\|x^k - x^*\| \\
&\quad + (1 - t_k)\underbrace{h_i(x^*, y^*)}_{\leq 0} \\
&\leq L\|x^k - x^*\| + t_k h_i(x^*, \hat{y}),
\end{aligned}$$

where L_i^1 and L_i^2 are the two local Lipschitz constants of h_i around $(x^*, \hat{y})$ and (x^*, y^*), respectively, and $L := \max\{\max\{L_i^1, L_i^2\} \mid i = 1, \dots, mN\}$. Since $x^k \to x^*$ and $h(x^*, \hat{y}) < 0$, we have

$$t_k := -2L\frac{\|x^k - x^*\|}{\max_i h_i(x^*, \hat{y})} \downarrow 0.$$

Using this particular sequence $\{t_k\}$ in the previous calculations, we get

$$h_i(x^k, y^k) \leq -L\|x^k - x^*\| \leq 0,$$

for all $i = 1, \dots, mN$ and, therefore, $y^k \in \Omega(x^k)$. □

Taking these two lemmas and Theorem 3.3 together, we immediately get the following continuity result.

Corollary 3.6 *Suppose that Assumption 1.1 holds. Then the functions y_α and V_α are continuous in $x^* \in X$ provided the Slater condition holds for $\Omega(x^*)$.*

Hence the optimization reformulation (4) of the GNEP is at least a continuous problem. Continuity alone, however, is not sufficient for the application of suitable nonsmooth optimization solvers to this problem. What is typically needed is at least the local Lipschitz continuity of the objective function and, if possible, the semismoothness of this mapping. Our next aim is therefore to show that these additional properties hold under fairly mild conditions. In fact, we will prove the stronger property that V_α is a PC^1 mapping.

To this end, we need a stronger smoothness property in addition to Assumption 1.1.

Assumption 3.7 The functions $\theta_\nu : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are twice continuously differentiable.

Note that Assumption 3.7 implies that the function h is twice continuously differentiable. Hence $y_\alpha(x)$ is the unique solution of the twice continuously differentiable optimization problem

$$\max_y \ \Psi_\alpha(x, y) \quad \text{s.t.} \quad h(x, y) \le 0. \tag{8}$$

Let

$$I(x) := \{i \in \{1, \dots, mN\} \mid h_i(x, y_\alpha(x)) = 0\}$$

be the set of active constraints. Consider, for a fixed subset $I \subseteq I(x)$, the problem (which has equality constraints only)

$$\max_y \ \Psi_\alpha(x, y) \quad \text{s.t.} \quad h_i(x, y) = 0 \ \ (i \in I). \tag{9}$$

Let

$$L_\alpha^I(x, y, \lambda) := -\Psi_\alpha(x, y) + \sum_{i \in I} \lambda_i h_i(x, y)$$

be the Lagrangian of the optimization problem (9). Then the KKT-system of this problem reads

$$\begin{aligned} &\nabla_y L_\alpha^I(x, y, \lambda) = -\nabla_y \Psi_\alpha(x, y) + \sum_{i \in I} \lambda_i \nabla_y h_i(x, y) = 0, \\ &h_i(x, y) = 0 \ \forall i \in I. \end{aligned} \tag{10}$$

This can be written as a nonlinear system of equations

$$\Phi_\alpha^I(x, y, \lambda) = 0 \quad \text{with} \quad \Phi_\alpha^I(x, y, \lambda) := \begin{pmatrix} \nabla_y L_\alpha^I(x, y, \lambda) \\ h_I(x, y) \end{pmatrix}, \tag{11}$$

where h_I consists of all components h_i of h with $i \in I$. The function Φ_α^I is continuously differentiable since Ψ_α and g are twice continuously differentiable, and we have

$$\nabla \Phi_\alpha^I(x, y, \lambda) = \begin{pmatrix} \nabla_{yx}^2 L_\alpha^I(x, y, \lambda)^T & \nabla_{yy}^2 L_\alpha^I(x, y, \lambda) & \nabla_y h_I(x, y)^T \\ \nabla_x h_I(x, y) & \nabla_y h_I(x, y) & 0 \end{pmatrix}.$$

Therefore, we obtain

$$\nabla_{(y,\lambda)} \Phi_\alpha^I(x, y, \lambda) = \begin{pmatrix} \nabla_{yy}^2 L_\alpha^I(x, y, \lambda) & \nabla_y h_I(x, y)^T \\ \nabla_y h_I(x, y) & 0 \end{pmatrix}.$$

Then we have the following result whose proof is standard so we skip it.

Lemma 3.8 *Suppose that Assumption 3.7 holds, that $\nabla^2_{yy} L^I_\alpha(x, y, \lambda)$ is positive definite and that the gradients $\nabla_y h_i(x, y)$ $(i \in I)$ are linearly independent. Then $\nabla_{(y,\lambda)} \Phi^I_\alpha(x, y, \lambda)$ is nonsingular.*

Note that the assumed positive definiteness of the Hessian $\nabla^2_{yy} L^I(x, y, \lambda)$ is an assumption that can easily be relaxed in Lemma 3.8, but that this condition automatically holds in our situation, so we do not really need a weaker assumption here. Furthermore, we stress that the assumed linear independence of the gradients $\nabla_y h_i(x, y)$ $(i \in I)$ is a very strong condition for certain index sets I, however, in our subsequent application of Lemma 3.8, we will only consider index sets I where this assumption holds automatically, so this condition is not really crucial in our context.

We next introduce another assumption that will be used in order to show that our objective function V_α is a PC^1 mapping.

Assumption 3.9 The *(feasible) constant rank constraint qualification* (CRCQ) holds at $x^* \in X$ if there exists a neighbourhood N of x^* such that for every subset $I \subseteq I(x^*) := \{i \mid h_i(x^*, y_\alpha(x^*)) = 0\}$, the set of gradient vectors $\{\nabla_y h_i(x, y_\alpha(x)) \mid i \in I\}$ has the same rank (depending on I) for all $x \in N \cap X$.

Note that the previous definition requires the same rank only for those $x \in N$ which also belong to the common strategy space X; this is important in our case since for $x \notin X$, the vector $y_\alpha(x)$ is not necessarily defined. Moreover, this is the only difference compared to the standard CRCQ as introduced in [18] and the reason why we call this assumption the feasible CRCQ, although, in our subsequent discussion, we will often speak of the CRCQ condition when we refer to Assumption 3.9. This feasible CRCQ has also been used before in [7], for example, where the authors simply call this condition the CRCQ.

The following result is motivated by [26] (see also [16]) and states that both y_α and V_α are piecewise continuously differentiable functions.

Theorem 3.10 *Suppose that Assumptions 1.1 and 3.7 hold, let $x^* \in X$ be given, and suppose that the solution mapping $y_\alpha : X \to \mathbb{R}^n$ of (8) is continuous in a neighbourhood of x^* (see Corollary 3.6 for a sufficient condition). Then there exists a neighbourhood $\hat{N}$ of $x^* \in X$ such that y_α is a PC^1 function on $\hat{N} \cap X$ provided that the (feasible) CRCQ condition from Assumption 3.9 holds at x^*.*

Proof We divide the proof into several steps.

Step 1: Here we introduce some notation and summarize some preliminary statements that will be useful later on.

First let $x^* \in X$ be fixed such that Assumption 3.9 holds in a neighbourhood N of x^*. Recall that $I(x) := \{i \mid h_i(x, y_\alpha(x)) = 0\}$ for all $x \in N \cap X$. Furthermore, for any such $x \in N \cap X$, let us denote by

$$\mathcal{M}(x) := \{\lambda \in \mathbb{R}^{mN} \mid (y_\alpha(x), \lambda) \text{ is a KKT point of } (8)\}$$

the set of all Lagrange multipliers of the optimization problem (8). Since CRCQ holds at x^*, it is easy to see that CRCQ also holds for all $x \in X$ sufficiently close to x^*. Without loss of generality, let us say that CRCQ holds for all $x \in N \cap X$ with the same neighbourhood N as before. Then it follows from a result in [18] that the set $\mathcal{M}(x)$ is nonempty for all $x \in N \cap X$. This, in turn, implies that the set

$$\mathcal{B}(x) := \big\{ I \subseteq I(x) \mid \nabla_y h_i(x, y_\alpha(x))\ (i \in I) \text{ are linearly independent and} \\ \text{supp}(\lambda) \subseteq I \text{ for some } \lambda \in \mathcal{M}(x) \big\}$$

is also nonempty for all x in a sufficiently small neighbourhood of x^*, say, again, for all $x \in N \cap X$ (see [16] for a formal proof), where supp(λ) denotes the support of the nonnegative vector λ, i.e. $\text{supp}(\lambda) := \{i \mid \lambda_i > 0\}$. Furthermore, it can be shown that, in a suitable neighbourhood of x^* (which we assume to be N once again), we have $\mathcal{B}(x) \subseteq \mathcal{B}(x^*)$, see, e.g., [16, 26].

Step 2: Here we show that, for every $x \in N \cap X$ and every $I \in \mathcal{B}(x)$, there is a unique multiplier $\lambda_\alpha^I(x) \in \mathcal{M}(x)$ such that $\Phi_\alpha^I(x, y_\alpha(x), \lambda_\alpha^I(x)) = 0$, where $N, \mathcal{M}(x)$, and $\mathcal{B}(x)$ are defined as in Step 1.

To this end, let $x \in N \cap X$ and $I \in \mathcal{B}(x)$ be arbitrarily given. The definition of $\mathcal{B}(x)$ implies that there is a Lagrange multiplier $\lambda_\alpha^I(x) \in \mathcal{M}(x)$ with $\text{supp}(\lambda_\alpha^I(x)) \subseteq I$. Since $(x, y_\alpha(x), \lambda_\alpha^I(x))$ satisfies the KKT conditions of the optimization problem (8), $[\lambda_\alpha^I(x)]_i = 0$ for all $i \notin I$, and $h_i(x, y_\alpha(x)) = 0$ for all $i \in I$ (since $I \subseteq I(x)$), it follows that $\Phi_\alpha^I(x, y_\alpha(x), \lambda_\alpha^I(x)) = 0$. Moreover, the linear independence of the gradients $\nabla_y h_i(x, y_\alpha(x))$ for $i \in I$ shows that the multiplier $\lambda_\alpha^I(x)$ is unique.

Step 3: Here we claim that, for any given $x^* \in X$ satisfying Assumption 3.9 and an arbitrary $I \in \mathcal{B}(x^*)$ with corresponding multiplier λ^*, there exist open neighbourhoods $N^I(x^*)$ and $N^I(y_\alpha(x^*), \lambda^*)$ as well as a C^1-diffeomorphism $(y^I(\cdot), \lambda^I(\cdot)) : N^I(x^*) \to N^I(y_\alpha(x^*), \lambda^*)$ such that $y^I(x^*) = y_\alpha(x^*), \lambda^I(x^*) = \lambda^*$ and $\Phi_\alpha^I(x, y^I(x), \lambda^I(x)) = 0$ for all $x \in N^I(x^*)$.

To verify this statement, let $x^* \in X$ be given such that the CRCQ holds, choose $I \in \mathcal{B}(x^*)$ arbitrarily, and let $\lambda^* \in \mathcal{M}(x^*)$ with $\text{supp}(\lambda^*) \subseteq I$ be a corresponding multiplier coming from the definition of the set $\mathcal{B}(x^*)$. Now, consider once again the nonlinear system of equations $\Phi_\alpha^I(x, y, \lambda) = 0$ with Φ_α^I being defined in (11). The function Φ_α^I is continuously differentiable, and the triple $(x^*, y_\alpha(x^*), \lambda^*)$ satisfies this system. The convexity of θ_ν with respect to x^ν implies that $-\Psi_\alpha^I(x^*, .)$ is strongly convex with respect to the second argument and, therefore, $\nabla_{yy}^2(-\Psi_\alpha^I(x^*, y_\alpha(x^*)))$ is positive definite. Moreover, the convexity of $h_i(x^*, .)$ in the second argument implies the positive semidefiniteness of $\nabla_{yy}^2 h_i(x^*, y_\alpha(x^*))$. Since $\lambda^* \geq 0$, it follows that the Hessian of the Lagrangian L_α^I evaluated in $(x^*, y_\alpha(x^*), \lambda^*)$, i.e. the matrix

$$\nabla_{yy}^2 L_\alpha^I(x^*, y_\alpha(x^*), \lambda^*) = -\nabla_{yy}^2 \Psi_\alpha(x^*, y_\alpha(x^*)) + \sum_{i \in I} \lambda_i^* \nabla_{yy}^2 h_i(x^*, y_\alpha(x^*))$$

is positive definite. Since, in addition, $\nabla_y h_i(x^*, y_\alpha(x^*))$ $(i \in I)$ are linearly independent in view of our choice of $I \in \mathcal{B}(x^*)$, the matrix $\nabla_{(y,\lambda)} \Phi_\alpha^I(x^*, y_\alpha(x^*), \lambda^*)$ is nonsingular by Lemma 3.8. The statement therefore follows from the standard implicit function theorem, where, without loss of generality, we can assume that $N^I(x^*) \subseteq N$.

Step 4: Here we verify the statement of our theorem.

Let $x^* \in X$ be given such that CRCQ holds in x^*. Define $\hat{N} := \bigcap_{I \in \mathcal{B}(x^*)} N^I(x^*)$ with the neighbourhoods $N^I(x^*)$ from Step 3. Since $\mathcal{B}(x^*)$ is a finite set, $\hat{N}$ is a neighborhood of x^*.

Choose $x \in \hat{N} \cap X$ arbitrarily. Step 2 shows that, for each $I \in \mathcal{B}(x)$, there exists a unique multiplier $\lambda_\alpha^I(x) \in \mathcal{M}(x)$ satisfying $\Phi_\alpha^I(x, y_\alpha(x), \lambda_\alpha^I(x)) = 0$. On the other hand, Step 3 guarantees that there exists neighbourhoods $N^I(x^*)$ and $N^I(y_\alpha(x^*), \lambda^*)$ and a C^1-diffeomorphism $y^I(\cdot), \lambda^I(\cdot) : N^I(x^*) \to N^I(y_\alpha(x^*), \lambda^*)$ such that $\Phi_\alpha^I(x, y^I(x), \lambda^I(x)) = 0$ for all $x \in N^I(x^*)$. In particular, $y^I(x), \lambda^I(x)$ is the locally unique solution of the system of equations $\Phi_\alpha^I(x, y, \lambda) = 0$ for all $x \in N^I(x^*)$. Hence, as soon as we can show that $(y_\alpha(x), \lambda_\alpha^I(x))$ belongs to the neighbourhood $N^I(y_\alpha(x^*), \lambda^*)$ for all $x \in X$ sufficiently close to x^*, the local uniqueness then implies $y_\alpha(x) = y^I(x)$ (for all $I \in \mathcal{B}(x) \subseteq \mathcal{B}(x^*)$).

Suppose this is not true in a sufficiently small neighbourhood. Then there is a sequence $\{x^k\} \subseteq X$ with $\{x^k\} \to x^*$ and a corresponding sequence of index sets $I_k \in \mathcal{B}(x^k)$ such that $(y_\alpha(x^k), \lambda_\alpha^{I_k}(x^k)) \notin N^{I_k}(y_\alpha(x^*), \lambda^*)$ for all $k \in \mathbb{N}$. Since $\mathcal{B}(x^k) \subseteq \mathcal{B}(x^*)$ contains only finitely many index sets, we may assume that I_k is the same index set for all k which we denote by I.

By the continuity of y_α, we have $y_\alpha(x^k) \to y_\alpha(x^*)$. On the other hand, for every x^k with associated $y_\alpha(x^k)$ and $\lambda_\alpha^I(x^k)$ from Step 2, we have

$$-\nabla_y \Psi_\alpha(x^k, y_\alpha(x^k)) + \sum_{i \in I} [\lambda_\alpha^I(x^k)]_i \nabla_y h_i(x^k, y_\alpha(x^k)) = 0 \tag{12}$$

for all k. The continuity of all functions involved, together with the linear independence of the vectors $\nabla_y h_i(x^*, y_\alpha(x^*))$ (which is a consequence of $I \in \mathcal{B}(x^k) \subseteq \mathcal{B}(x^*)$ and the assumed CRCQ condition) implies that the sequence $\{\lambda_\alpha^I(x^k)\}$ is convergent, say $\{\lambda_\alpha^I(x^k)\} \to \bar{\lambda}^I$ for some limiting vector $\bar{\lambda}^I$. Taking the limit in (12) and using once again the continuity of the solution mapping $y_\alpha(\cdot)$ then gives $-\nabla_y \Psi_\alpha(x^*, y_\alpha(x^*)) + \sum_{i \in I} \bar{\lambda}_i^I \nabla_y h_i(x^*, y_\alpha(x^*)) = 0$. Note that the CRCQ condition implies that $\bar{\lambda}^I$ is uniquely defined by this equation and the fact that $\bar{\lambda}_i^I = 0$ for all $i \notin I$. However, by definition, the vector λ^* also satisfies this equation, hence we have $\lambda_\alpha^I(x^k) \to \lambda^*$. But then it follows that $(y_\alpha(x^k), \lambda_\alpha^I(x^k)) \in N^I(y_\alpha(x^*), \lambda^*)$, and this contradiction implies the desired statement. □

Thus we get the following corollary.

Corollary 3.11 *Suppose that Assumptions 1.1 and 3.7 hold. Moreover, suppose that Assumption 3.9 holds in $x^* \in X$ and that the sets $\Omega(x)$ satisfy a Slater condition for all $x \in X$ sufficiently close to x^*. Then y_α and V_α are PC^1 functions in a neighbourhood of x^*.*

Proof From Corollary 3.6, we obtain the continuity of y_α, whereas Theorem 3.10 implies the PC^1 property of y_α near x^*. Hence the composite mapping $V_\alpha(x) = \Psi_\alpha(x, y_\alpha(x))$ is also continuous and a PC^1 mapping in a neighbourhood of x^*. □

4 Smoothness properties of the unconstrained reformulation

Here we consider the unconstrained reformulation (7) with the objective function $\bar{V}_{\alpha\beta}$ from Definition 2.3. We will show that the smoothness properties of the constrained reformulation can be transfered to the unconstrained one. This means we can prove continuity under a Slater-type condition and, moreover, that $\bar{V}_{\alpha\beta}$ is a PC^1 function provided g and θ_ν are twice continuously differentiable and a constant rank constraint qualification holds. Although the proofs for these results are similar to the analysis from the previous section, there are also some significant differences. In order to keep this section as short as possible, we will, more or less, only stress those points where these differences occur.

For the unconstrained reformulation, we first define the function

$$\bar{h} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{mN} \quad \text{by} \quad \bar{h}(x, y) := \begin{pmatrix} g(y^1, (P_X[x])^{-1}) \\ \vdots \\ g(y^N, (P_X[x])^{-N}) \end{pmatrix}$$

which is the analogue of the mapping h used in the previous section. Then we have

$$y \in \Omega(P_X[x]) \quad \Longleftrightarrow \quad \bar{h}(x, y) \leq 0$$

for any given $x \in \mathbb{R}^n$. Note, however, that in contrast to the mapping h, the function $\bar{h}$ is nondifferentiable in general, even if g itself is differentiable, simply because the projection mapping is nonsmooth. However, $\bar{h}$ is continuously differentiable with respect to y, at least under the smoothness condition from Assumption 3.7.

Our first aim is to show continuity of $\bar{V}_{\alpha\beta}$. Similar to the constrained reformulation, the continuity of $\bar{y}_\alpha$ (hence of $\bar{V}_{\alpha\beta}$) follows directly from the point-to-set mapping $x \mapsto \Omega(P_X[x])$ being lower semicontinuous and closed. The proofs for this mapping being lower semicontinuous and closed are along the lines of the proofs of Lemmas 3.4 and 3.5 by using the continuity and Lipschitz property of the projection mapping. Hence Corollary 3.6 transfers to the unconstrained reformulation and shows continuity of $\bar{y}_\alpha$ and $\bar{V}_{\alpha\beta}$, i.e. we have the following result.

Corollary 4.1 *Suppose that Assumption 1.1 holds. Then $\bar{y}_\alpha$ and $\bar{V}_{\alpha\beta}$ are continuous in every $x^* \in \mathbb{R}^n$ where $\Omega(P_X[x^*])$ satisfies the Slater condition.*

Hence the unconstrained reformulation (7) of the GNEP is also a continuous problem. Now we want to show that the function $\bar{V}_{\alpha\beta}$ is a PC^1 mapping. To this end, recall that $\bar{y}_\alpha(x)$ is the unique solution of

$$\max_y \ \Psi_\alpha(x, y) \quad \text{s.t.} \quad \bar{h}(x, y) \leq 0. \tag{13}$$

The function $\bar{h}$ is not continuously differentiable, but it is a PC^1 function if the projection mapping itself is a PC^1 mapping. This PC^1 property of the projection mapping is shown in [25] under the smoothness conditions of Assumption 3.7 and a constant rank constraint qualification. Hence, we first define the constant rank constraint qualification in a way it will be used within this section.

Assumption 4.2 The *constant rank constraint qualification* (CRCQ) holds at $x^* \in \mathbb{R}^n$ if there exists a neighbourhood N of x^* such that for every subset $I \subseteq \bar{I}(x^*) := \{i \mid \bar{h}_i(x^*, \bar{y}_\alpha(x^*)) = 0\}$, the set of gradient vectors $\{\nabla_y \bar{h}_i(x, \bar{y}_\alpha(x)) \mid i \in I\}$ has the same rank (depending on I) for all $x \in N$.

Note that there are some minor differences between Assumptions 3.9 and 4.2: Here we use $\bar{h}$ and $\bar{y}_\alpha$ instead of h and y_α, respectively. Furthermore, we assume the same rank for all $x \in N$, whereas in Assumption 3.9 is was enough to consider a feasible neighbourhood $N \cap X$. The latter is not possible in our context now since we use an unconstrained reformulation here, so x could be any vector from $\mathbb{R}^n$.

To get an analogous result to Theorem 3.10, we need an implicit function theorem for PC^1 functions.

Theorem 4.3 *Assume $H : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ is a PC^1 function in a neighborhood of $(\bar{x}, \bar{y})$ with $H(\bar{x}, \bar{y}) = 0$ and all matrices in $\pi_y \partial H(\bar{x}, \bar{y})$ have the same nonzero orientation. Then there exists an open neighborhood U of $\bar{x}$ and a function $g : U \to \mathbb{R}^n$ which is a PC^1 function on U such that $g(\bar{x}) = \bar{y}$ and $H(x, g(x)) = 0$ for all $x \in U$.*

Proof We will derive this implicit function theorem from an inverse function theorem in [7]. To do so, define

$$F : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n \quad \text{by} \quad F(x, y) := \begin{pmatrix} x - \bar{x} \\ H(x, y) \end{pmatrix}.$$

Then we have

$$\partial F(\bar{x}, \bar{y}) \subseteq \begin{pmatrix} I_m & 0 \\ \pi_x \partial H(\bar{x}, \bar{y}) & \pi_y \partial H(\bar{x}, \bar{y}) \end{pmatrix},$$

and all elements in $\partial F(\bar{x}, \bar{y})$ have the same nonzero orientation, because the matrices in $\pi_y \partial H(\bar{x}, \bar{y})$ have. With H also the function F is a PC^1 function in a neighborhood of $(\bar{x}, \bar{y})$. By Lemma 2.2 in [19], we get for the index $ind(F, (\bar{x}, \bar{y})) \in \{+1, -1\}$. Now we can use the inverse function theorem from [7, Theorem 4.6.5] which implies the existence of open neighborhoods V of $(\bar{x}, \bar{y})$ and W of $(0, 0) = F(\bar{x}, \bar{y})$ such that $F : V \to W$ is a homeomorphism and the local inverse $G : W \to V$ is a PC^1 function. Define the set

$$U := \{x \in \mathbb{R}^n \mid (x - \bar{x}, 0) \in W\}.$$

U is nonempty and open (in $\mathbb{R}^n$) since $(0, 0) \in W$ and W is open. Let $x \in U$ arbitrarily be given. Then we have $(x - \bar{x}, 0) \in W$ and hence, by the definition of a homeomorphism, we obtain the existence of a unique y with $(x, y) \in V$ and $F(x, y) = (x - \bar{x}, 0)$. Thus we have $H(x, y) = 0$. Since y depends on x, we write $y =: g(x)$ which defines a function $g : U \to \mathbb{R}^n$ such that $H(x, g(x)) = 0$ for each $x \in U$. Therefore we have

$$F(x, g(x)) = \begin{pmatrix} x - \bar{x} \\ H(x, g(x)) \end{pmatrix} = \begin{pmatrix} x - \bar{x} \\ 0 \end{pmatrix}$$

for all $x \in U$. Applying the inverse function G on both sides, we obtain $(x, g(x)) = G(x - \bar{x}, 0)$ for all $x \in U$. Since g coincides with some component functions of the PC^1 function G, it is a PC^1 function itself which completes the proof. □

Now we are able to show an analogous result to Theorem 3.10.

Theorem 4.4 *Suppose that Assumptions 1.1 and 3.7 hold. Let $x^* \in \mathbb{R}^n$ be given and suppose that the solution mapping $\bar{y}_\alpha : \mathbb{R}^n \to \mathbb{R}^n$ of (13) is continuous in a neighbourhood of x^* (see Corollary 4.1 for a sufficient condition). Then $\bar{y}_\alpha$ is a PC^1 function in a neighbourhood of x^* provided that the CRCQ condition from Assumption 4.2 holds at x^*.*

Proof We follow the proof of Theorem 3.10 by dividing the proof into four steps. Rather than giving all the details, however, we more or less only mention the differences.

Step 1: Similar to the discussion in Sect. 3, let us introduce the sets

$$\bar{I}(x) := \{i \mid \bar{h}_i(x, \bar{y}_\alpha(x)) = 0\},$$
$$\bar{\mathcal{M}}(x) := \{\lambda \in \mathbb{R}^{mN} \mid (\bar{y}_\alpha(x), \lambda) \text{ is a KKT point of (13)}\}$$

and

$$\bar{\mathcal{B}}(x) := \Big\{I \subseteq \bar{I}(x) \mid \nabla_y \bar{h}_i(x, \bar{y}_\alpha(x))\ (i \in I) \text{ are linearly independent and } \operatorname{supp}(\lambda) \subseteq I \text{ for some } \lambda \in \bar{\mathcal{M}}(x)\Big\}.$$

Then Assumption 4.2 implies that there is a neighbourhood N of x^* such that $\bar{\mathcal{M}}(x) \neq \emptyset$, $\bar{\mathcal{B}}(x) \neq \emptyset$ and $\bar{\mathcal{B}}(x) \subseteq \bar{\mathcal{B}}(x^*)$ for all $x \in N$.

Step 2: For an arbitrary vector $x \in \mathbb{R}^n$ and an index set $I \subseteq \bar{I}(x)$, consider the optimization problem

$$\max_y \ \Psi_\alpha(x, y) \quad \text{s.t.} \quad \bar{h}_i(x, y) = 0 \ \ (i \in I).$$

The corresponding Lagrangian is given by

$$\bar{L}_\alpha^I(x, y, \lambda) := -\Psi_\alpha(x, y) + \sum_{i \in I} \lambda_i \bar{h}_i(x, y),$$

so that the KKT conditions can be rewritten as

$$\bar{\Phi}^I_\alpha(x, y, \lambda) = 0 \quad \text{with} \quad \bar{\Phi}^I_\alpha(x, y, \lambda) := \begin{pmatrix} \nabla_y \bar{L}^I_\alpha(x, y, \lambda) \\ \bar{h}_I(x, y) \end{pmatrix}.$$

Using this notation, it follows as in the proof of Theorem 3.10 that, for every $x \in N$ and every $I \in \bar{\mathcal{B}}(x)$, there is a unique multiplier $\lambda^I_\alpha(x) \in \bar{\mathcal{M}}(x)$ such that $\bar{\Phi}^I_\alpha(x, \bar{y}_\alpha(x), \lambda^I_\alpha(x)) = 0$, where N, $\bar{\mathcal{M}}(x)$, and $\bar{\mathcal{B}}(x)$ are the sets defined in Step 1.

Step 3: Here we have the main difference to the proof of Theorem 3.10 since the mapping $\bar{\Phi}^I_\alpha$ defined in Step 2 is only a PC^1 function, but not continuously differentiable (in contrast to the mapping Φ^I_α from the previous section which was continuously differentiable). Therefore, we have to use an implicit function theorem for PC^1 functions instead of the standard implicit function theorem. Let any $x^* \in \mathbb{R}^n$ satisfying Assumption 4.2 and an arbitrary $I \in \bar{\mathcal{B}}(x^*)$ with corresponding multiplier λ^* be given. Since $\bar{\Phi}^I_\alpha(x, y, \lambda)$ is continuously differentiable with respect to y and λ, it follows that $\pi_{(y,\lambda)} \partial \Phi^I_\alpha(x^*, \bar{y}_\alpha(x^*), \lambda^*)$ has only one element, whose nonsingularity can be shown as in the proof of Theorem 3.10. In particular, the same nonzero orientation of all the elements is guaranteed. Using the PC^1 implicit function Theorem 4.3, we get the existence of open neighbourhoods $N^I(x^*)$ and $N^I(\bar{y}_\alpha(x^*), \lambda^*)$ as well as a PC^1 function $(y^I(\cdot), \lambda^I(\cdot)) : N^I(x^*) \to N^I(\bar{y}_\alpha(x^*), \lambda^*)$ such that $y^I(x^*) = \bar{y}_\alpha(x^*)$, $\lambda^I(x^*) = \lambda^*$ and $\Phi^I_\alpha(x, y^I(x), \lambda^I(x)) = 0$ for all $x \in N^I(x^*)$.

Step 4: Repeating the arguments from Step 4 of the proof of Theorem 3.10, we obtain $\bar{y}_\alpha(x) \in \{y^I(x) \mid I \in \bar{\mathcal{B}}(x^*)\}$ for all x in a sufficiently small neighborhood of x^*. Since all y^I are PC^1 functions, it follows that also $\bar{y}_\alpha$ is a PC^1 mapping in a neighborhood of any x^* satisfying the CRCQ condition from Assumption 4.2. □

Thus we get the following corollary.

Corollary 4.5 *Suppose that Assumptions 1.1 and 3.7 hold. Moreover, suppose that Assumption 4.2 holds in $x^* \in \mathbb{R}^n$ and that the sets $\Omega(P_X[x])$ satisfy the Slater condition for all x sufficiently close to x^*. Then $\bar{y}_\alpha$ and $\bar{V}_\alpha$ are PC^1 functions in a neighbourhood of x^*.*

Proof From Corollary 4.1 we obtain the continuity of $\bar{y}_\alpha$. Theorem 4.4 therefore implies the PC^1 property of $\bar{y}_\alpha$ near x^* satisfying the CRCQ condition from Assumption 4.2. Hence the composite mapping $\bar{V}_\alpha(x) = \Psi_\alpha(x, \bar{y}_\alpha(x))$ and therefore also $\bar{V}_{\alpha\beta}$ are PC^1 mappings in a neighborhood of x^*. □

Thus we have shown that also the PC^1 property transfers from the constrained to the unconstrained reformulation. In particular, it follows that the objective function $\bar{V}_{\alpha\beta}$ is directionally differentiable, locally Lipschitz continuous, and semismooth under the assumptions of Corollary 4.5, cf. [2].

A good property that we would like to have is a condition which guarantees that a Clarke stationary point is already a global minimum and, therefore, a solution of our

GNEP, i.e. we would like to have the implication

$$0 \in \partial \bar{V}_{\alpha\beta}(x^*) \quad \Longrightarrow \quad \bar{V}_{\alpha\beta}(x^*) = 0.$$

However, a proof of this result is difficult due to the fact that suitable estimates (like the one from [21]) for the Clarke subdifferential of $\bar{V}_{\alpha\beta}$ are by far too large. In fact, even in very simple examples, it turns out that these estimates are equal to the entire set $\mathbb{R}^n$. It is therefore an interesting question whether (much) better estimates can be obtained for our particular function, but this is not the scope of this paper. We stress, however, that the numerical results in the following section indicate that our method is usually able to find solutions of the GNEP.

5 Numerical results

Here we present some numerical results that are obtained by applying the robust gradient sampling algorithm from [1] to our unconstrained optimization reformulation using the objective function $\bar{V}_{\alpha\beta}$. The MATLAB® implementation used for our numerical tests is the one written by the authors of [1] which is available online at the following address: http://www.cs.nyu.edu/overton/papers/gradsamp. The method involves a random sampling strategy which implies that it (usually) generates different iterates (hence possibly different solutions) even if we use the same starting point. The limit point of any sequence generated by this method is a Clarke stationary point with probability 1. The algorithm stops if the norm of the vector with the smallest Euclidean norm in the convex hull of the sampled gradients is less than 10^{-6}. Apart from using standard parameter settings, we use the two values $\alpha = 0.02$ and $\beta = 0.05$ which define our objective function. In order to evaluate this objective function, we have to compute the vectors $\bar{y}_\alpha(x)$ and $\bar{y}_\beta(x)$. This is done by using the `fmincon` solver from the MATLAB® Optimization Toolbox. In a similar way, projections onto the convex set X are computed by using suitable methods from the same toolbox.

Regarding the examples that are used for our numerical tests, we only took problems from the literature which are known to have multiple solutions since otherwise the examples would be uninteresting for our method. In the first four 2-dimensional examples all generalized Nash equilibria are known analytically. By taking 100 randomly generated starting vectors in a neighbourhood of the feasible set X, we show that the computed generalized Nash equilibria spread over the whole solution set. Then we consider four examples with more than two players or more than one variable for each player which also show that the algorithm finds different solutions.

Example 5.1 This problem is a two player game from [4]. Each player has a one-dimensional variable $x^\nu \in \mathbb{R}$. The problem uses the cost functions

$$\theta_1(x) := (x^1 - 1)^2 \quad \text{and} \quad \theta_2(x) := \left(x^2 - \frac{1}{2}\right)^2$$

and the feasible set $X := \{x \in \mathbb{R}^2 \mid x^1 + x^2 \leq 1\}$. There are infinitely many solutions given by $\{(\lambda, 1 - \lambda) \mid \lambda \in [\frac{1}{2}, 1]\}$. We tested 100 random starting points in $[-2, 2]^2$.

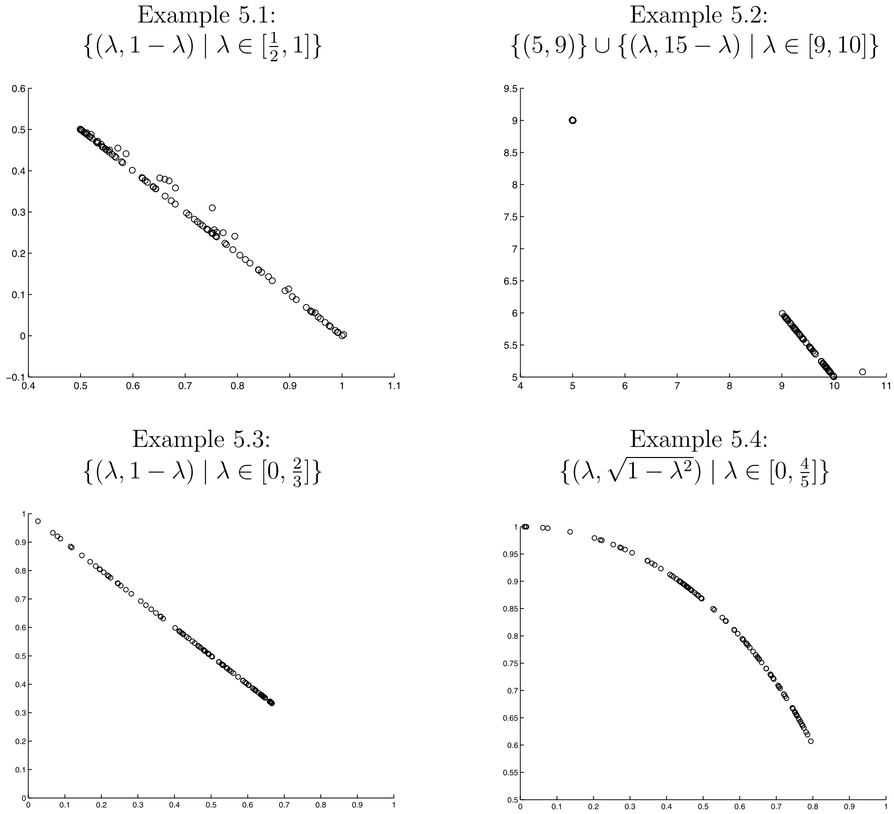

Fig. 1 Numerical results

We have convergence for all starting points in a range from 8 to 23 iterations. 77 computed solution have a marginal function value less than $1.5 * 10^{-10}$, further 11 solutions a value less than $5.0 * 10^{-7}$. The rest has a value less than $3.0 * 10^{-5}$. The computed solutions are displayed in Fig. 1.

Example 5.2 The following example is taken from [11]. In the game we have two players, each controlling a single variable $x_i \in \mathbb{R}$. The objective functions are given by

$$\theta_1(x) = x_1^2 + \frac{8}{3}x_1x_2 - 34x_1, \quad \text{and} \quad \theta_2(x) = x_2^2 + \frac{5}{4}x_1x_2 - 24.25x_2.$$

The common strategy set is $X = \{x \in \mathbb{R}^2 \mid 0 \le x_1 \le 10, 0 \le x_2 \le 10, x_1 + x_2 \le 15\}$ and the solution set is $\{(5, 9)\} \cup \{(\lambda, 15 - \lambda) \mid \lambda \in [9, 10]\}$. We tested 100 starting vectors randomly distributed over $[0, 10]^2$. 99 times the algorithm converged to a point x^* with function value $\bar{V}_{\alpha\beta}(x^*) < 1.6 * 10^{-11}$ and the iteration number was between 12 and 31. Once the algorithm terminates at a point which is not a solution; 48 times the algorithm converged to the isolated generalized Nash equilibrium (5, 9) and in the other cases the computed equilibria spread over the set of solutions. The results are displayed in Fig. 1.

Example 5.3 Here we consider a two-player game, where each player controls one variable $x_i \in \mathbb{R}$. The example is taken from [22]. The objective functions are

$$\theta_1(x) = x_1^2 - x_1 x_2 - x_1, \quad \text{and} \quad \theta_2(x) = x_2^2 - \frac{1}{2} x_1 x_2 - 2x_2.$$

The feasible set is $X = \{x \in \mathbb{R}^2 \mid x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 1\}$. The solution set is given by $\{(\lambda, 1 - \lambda) \mid \lambda \in [0, \frac{2}{3}]\}$. The algorithm was tested with 100 random starting vectors in $[0, 1]^2$ and converged in all cases. The iteration number was between 9 and 49 and the marginal function value was less than $1.7 * 10^{-11}$. Figure 1 shows the computed solutions.

Example 5.4 This is a modification of the Example 5.3 with the same objective functions but a different feasible set given by $X = \{x \in \mathbb{R}^2 \mid x_1 \geq 0, x_2 \geq 0, x_1^2 + x_2^2 \leq 1\}$. The solution set modifies to $\{(\lambda, \sqrt{1 - \lambda^2}) \mid \lambda \in [0, \frac{4}{5}]\}$. This modification was already considered in [22]. The algorithm was again tested with 100 random starting vectors in $[0, 1] \times [0, 2]$. We have convergence for all starting vectors, marginal function values less than $1.7 * 10^{-11}$ and iteration numbers between 9 and 22. For the distribution of the computed solutions, see again Fig. 1.

Example 5.5 This example is the river basin pollution game, taken from [20]. There are three players, each controlling a single variable $x^\nu \in \mathbb{R}$. The objective functions are

$$\theta_\nu(x) := x^\nu \left(c_{1\nu} + c_{2\nu} x^\nu - d_1 + d_2(x^1 + x^2 + x^3) \right)$$

for $\nu = 1, 2, 3$ with certain parameters specified in [6, 20]. The strategy space X is defined by some linear constraints, see again [6, 20] for more details. We used 100 different starting vectors, randomly distributed on $[0, 5]^3$ and found 100 different generalized Nash equilibria. The largest marginal function value $\bar{V}_{\alpha\beta}(x^*)$ was $1.6 * 10^{-11}$ and the number of iterations was between 21 and 38. We made a second run with three nonrandom starting vectors, each one used three times and we found 9 different equilibria, see Table 1 for some numerical results.

Example 5.6 This problem is an oligopoly model for $N = 5$ players, each player controlling a single variable $x^\nu \in \mathbb{R}$. The objective functions are highly nonlinear and given by

$$\theta_\nu(x) := f_\nu(x^\nu) - 5000^{1/\gamma} x^\nu (x^1 + \cdots + x^N)^{-1/\gamma}$$

for all $\nu = 1, \ldots, N$ with

$$f_\nu(x^\nu) := c_\nu x^\nu + \frac{\delta_\nu}{1 + \delta_\nu} K_\nu^{-1/\delta_\nu} (x^\nu)^{(1+\delta_\nu)/\delta_\nu}$$

Table 1 Results for Example 5.5

x^0	It.	x^*	$\bar{V}_{\alpha\beta}(x^*)$
(0, 0, 0)	34	(9.6424, 9.5651, 13.7469)	$4.4 * 10^{-13}$
(0, 0, 0)	34	(9.1568, 7.7046, 14.6932)	$2.2 * 10^{-12}$
(0, 0, 0)	38	(11.6080, 9.1545, 12.3226)	$5.4 * 10^{-13}$
(1, 1, 1)	33	(10.5010, 9.4778, 13.0969)	$8.2 * 10^{-13}$
(1, 1, 1)	38	(12.1166, 11.6582, 11.1633)	$3.6 * 10^{-13}$
(1, 1, 1)	31	(10.3501, 8.8811, 13.3966)	$1.8 * 10^{-12}$
(1, 2, 3)	29	(9.9927, 10.6657, 13.1374)	$6.4 * 10^{-14}$
(1, 2, 3)	35	(9.3339, 9.8344, 13.9084)	$5.3 * 10^{-14}$
(1, 2, 3)	33	(11.1988, 10.8160, 12.1415)	$4.6 * 10^{-13}$

Table 2 Results for Example 5.6

P	It.	x^*	$\bar{V}_{\alpha\beta}(x^*)$
75	59	(13.8905, 14.5065, 15.1038, 15.3253, 16.1811)	$1.5 * 10^{-7}$
75	25	(7.9908, 11.5942, 15.1505, 18.5750, 21.6896)	$1.4 * 10^{-10}$
100	116	(18.4518, 19.5977, 20.4044, 20.6551, 20.8941)	$2.8 * 10^{-8}$
100	36	(13.8988, 17.2366, 20.3282, 23.1469, 25.3894)	$6.9 * 10^{-13}$
150	108	(27.3865, 30.2110, 31.5907, 31.1618, 29.6528)	$2.3 * 10^{-8}$
150	60	(23.7846, 28.2614, 31.6395, 33.3193, 32.9956)	$5.5 * 10^{-10}$
200	78	(35.7400, 40.4412, 42.7681, 42.1063, 38.9446)	$2.7 * 10^{-10}$
200	82	(34.7850, 40.2821, 43.0930, 42.6119, 39.2279)	$7.1 * 10^{-12}$

for all $\nu = 1, \dots, N$. For the precise values of the parameters involved in these functions, the reader is referred to [6, 23]. The constraints are linear:

$$x^1 + \dots + x^N \leq P, \quad x^\nu \geq 0 \quad \text{for all } \nu = 1, \dots, N.$$

We tested this problem with different total production parameters P and for each P we tested two different starting vectors $x^0 = (10, \dots, 10)^T$ and $x^0 = (0, 5, 10, 15, 20)^T$. Table 2 contains the corresponding results. The first column gives the production parameters P, the second column the number of iterations until convergence, the third column gives the computed solution, and the final column shows the value of the marginal function $\bar{V}_{\alpha\beta}$ at the computed solution x^*.

Example 5.7 Here we consider an electricity market model which is originally proposed in [24] and further discussed in [22]. The details of the problem are from the latter reference. It is a two player game where each player has six variables, $(x_1, \dots, x_6)^T$ for player 1 and $(x_7, \dots, x_{12})^T$ for player 2. All constraints are linear and the objective functions are quadratic. Table 3 shows the results of four test runs with symmetric starting vectors $x^0 = (0, \dots, 0)^T$, $x^0 = (10, \dots, 10)^T$, $x^0 = (100, 0, 0, 50, 0, 0, 100, 0, 0, 50, 0, 0)^T$, and $x^0 = (50, 25, 25, 25, 12.5, 12.5, 50, 25, 25, 25, 12.5, 12.5)^T$, and three test runs with random starting vectors. The symmetric starting vectors lead to symmetric solutions, the random ones do not.

Table 3 Results for Example 5.7

x^0	It.	x^*	$\bar{V}_{\alpha\beta}(x^*)$
$(0,\dots,0)$	118	(43.536, 28.139, 28.325, 26.868, 11.471, 11.661, 43.538, 28.138, 28.324, 26.870, 11.471, 11.658)	$3.3 * 10^{-12}$
(100, 0, 0, 50, 0, 0, 100, 0, 0, 50, 0, 0)	120	(60.203, 19.805, 19.992, 10.203, 19.805, 19.992, 60.203, 19.805, 19.992, 10.203, 19.805, 19.992)	$6.3 * 10^{-13}$
(50, 25, 25, 25, 12.5, 12.5, 50, 25, 25, 25, 12.5, 12.5)	41	(47.703, 26.055, 26.242, 22.703, 13.555, 13.742, 47.703, 26.055, 26.242, 22.703, 13.555, 13.742)	$2.9 * 10^{-12}$
random in $[0, 100]^{12}$	148	(43.207, 30.734, 26.059, 30.435, 8.8760, 10.689, 35.655, 26.724, 37.620, 31.515, 12.885, 5.5999)	$3.6 * 10^{-12}$
random in $[0, 100]^{12}$	157	(38.889, 39.603, 21.508, 31.122, 0.0067, 18.871, 23.997, 38.693, 37.311, 46.804, 0.9167, 2.2788)	$7.2 * 10^{-12}$
random in $[0, 100]^{12}$	159	(34.116, 37.595, 28.289, 35.904, 2.0147, 12.081, 36.863, 39.606, 23.530, 33.929, 0.0032, 16.068)	$7.3 * 10^{-12}$

Example 5.8 The last example is an electricity market model taken from [12]. There are two players and the five variables $(x_1,\dots,x_5)$ are controlled by player 1 and $(x_6,\dots,x_{10})$ are controlled of player 2. The objective functions are given by

$$\theta_1(x) = c^A \sum_{i=1}^4 x_i + e_1 x_1 + e_2(x_4+x_5) + e_3(x_3-x_5) + e_4 x_5 - \sum_{i=1}^4 \frac{x_i C_i^\gamma}{(x_i+x_{i+5})^\gamma},$$

$$\theta_2(x) = c^B \sum_{i=1}^4 x_{i+5} + e_1 x_6 + e_2(x_6+x_7+x_8-x_{10}) + e_3(x_8-x_{10}) + e_4 x_{10} - \sum_{i=1}^4 \frac{x_{i+5} C_i^\gamma}{(x_i+x_{i+5})^\gamma}.$$

The common strategy set is

$$X = \{x \in \mathbb{R}^{10} \mid 0.1 \le x_i, i = 1,\dots,10, x_5 \le x_3, x_{10} \le x_8, h(x) \le 0\}$$

where

$$h(x) := \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & -1 & 1 & 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} x - \begin{pmatrix} k_1 \\ k_2 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix}$$

defines the joint constraints. The parameters used are $e = (5, 2, 5, 2)$, $k = (300, 300, 300, 300)$, $C = (20000, 50000, 30000, 50000)$, $c^A = 26$, $c^B = 28$, $\gamma = 10/11$. In all solutions we found, six components have within a tolerance of 10^{-4} the same value.

Table 4 Results for Example 5.8

$x^0 * 10^{-2}$	It.	$x^*_{3,5,8,10}$	$\bar{V}_{\alpha\beta}(x^*)$
$(1,\dots,1)$	502	(176.1087, 121.6257, 178.3743, 178.3743)	$-1.4 * 10^{-11}$
$(2,\dots,2)$	326	(176.4614, 150.0000, 152.5344, 150.0000)	$2.0 * 10^{-14}$
(1, 1, 1, 1, 1, 3, 3, 3, 3, 3)	477	(175.8595, 117.0743, 182.9243, 182.9242)	$3.8 * 10^{-8}$
(3, 3, 3, 3, 3, 1, 1, 1, 1, 1)	439	(182.7254, 182.7254, 151.8423, 117.2746)	$6.2 * 10^{-13}$
(1, 3, 1, 3, 1, 1, 3, 1, 3, 1)	185	(176.1012, 121.4731, 178.5269, 178.5269)	$1.8 * 10^{-12}$
(1.17, 3.56, 1, 3.28, 1, 1.01, 3, 1, 3.28, 1)	171	(176.1003, 121.4544, 178.5456, 178.5456)	$-8.5 * 10^{-12}$
random in $[200, 500]^2$	596	(176.4614, 159.4029, 152.5344, 140.5971)	$-1.4 * 10^{-13}$
random in $[200, 500]^2$	511	(175.8600, 117.0816, 182.9184, 182.9183)	$8.7 * 10^{-12}$
random in $[200, 500]^2$	582	(182.7260, 182.7259, 151.8422, 117.2740)	$4.1 * 10^{-11}$

These values are

$$x_1 = 117.6409, \quad x_2 = 356.6286, \quad x_4 = 328.4737, \quad x_6 = 101.6896,$$
$$x_7 = 300.3188, \quad x_9 = 328.4737.$$

Table 4 reports starting vectors x^0, the number of iterations, solutions we found for the remaining variables $x_3^*, x_5^*, x_8^*, x_{10}^*$ and the corresponding marginal function value.

The previous examples show that the method finds different solutions, in particular, it computes non-normalized solutions. Furthermore, the tests indicate that the computed solutions spread over the whole set of solutions. Moreover, using the standard termination criterion for the software from [1], the accuracy is surprisingly high for most test runs (or, to be more precise, the function value $\bar{V}_{\alpha\beta}$ at termination is always relatively close to zero) which is an interesting observation since the software itself is, in general, not a fast converging method.

6 Final remarks

This paper discusses the smoothness properties of a known (see [13]) constrained reformulation of a jointly convex GNEP as well as of a new unconstrained reformulation. Both reformulations have the properties that they characterize all solutions of the GNEP (and not just the normalized ones) and that their objective functions are continuous under a Slater-type condition. Under an additional constant rank constraint qualification, the objective functions are, in fact, piecewise continuously differentiable. This allows the application of suitable nonsmooth optimization software in order to get a solution of GNEPs. So far, the investigations were restricted to the jointly convex class of GNEPs. An interesting future research topic is to see whether these results can be extended to a general (not necessarily jointly convex) GNEP.

Acknowledgement The authors would like to thank Defeng Sun for several helpful comments regarding the smoothness properties of the projection operator. They also thank the referees for some very detailed comments.

References

1. Burke, J.V., Lewis, A.S., Overton, M.L.: A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. SIAM J. Opt. **15**, 751–779 (2005)
2. Chaney, R.W.: Piecewise C^k functions in nonsmooth analysis. Nonlinear Anal., Theory Methods Appl. **15**, 649–660 (1990)
3. Clarke, F.H.: Optimization and Nonsmooth Analysis. SIAM, Philadelphia (1990)
4. Facchinei, F., Fischer, A., Piccialli, V.: Generalized Nash equilibrium problems and Newton methods. Math. Program. **117**, 163–194 (2009)
5. Facchinei, F., Kanzow, C.: Generalized Nash equilibrium problems. 4OR **5**, 173–210 (2007)
6. Facchinei, F., Kanzow, C.: Penalty methods for the solution of generalized Nash equilibrium problems. Technical Report, Institute of Mathematics, University of Würzburg, Germany, January 2009
7. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems I. Springer, New York (2003)
8. Facchinei, F., Pang, J.-S.: Exact penalty functions for generalized Nash problems. In: Di Pillo, G., Roma, M. (eds.) Large Scale Nonlinear Optimization, pp. 115–126. Springer, Berlin (2006)
9. Fukushima, M.: Restricted generalized Nash equilibria and controlled penalty algorithm. Technical Report 2008-007, Department of Applied Mathematics and Physics, Kyoto University, July 2008
10. Gürkan, G., Pang, J.-S.: Approximations of Nash equilibria. Math. Program. **117**, 223–253 (2009)
11. Harker, P.T.: Generalized Nash games and quasi-variational inequalities. Eur. J. Oper. Res. **54**, 81–94 (1991)
12. von Heusinger, A.: Numerical methods for the solution of the generalized Nash equilibrium problem. Ph.D. Thesis, Institute of Mathematics, University of Würzburg, August 2009
13. von Heusinger, A., Kanzow, C.: Optimization reformulations of the generalized Nash equilibrium problem using Nikaido-Isoda-type functions. Comput. Optim. Appl. **43**, 353–377 (2009)
14. von Heusinger, A., Kanzow, C.: SC^1 optimization reformulations of the generalized Nash equilibrium problem. Optim. Methods Softw. **23**, 953–973 (2008)
15. von Heusinger, A., Kanzow, C.: Relaxation methods for generalized Nash equilibrium problems with inexact line search. J. Optim. Theory Appl. **143**, 159–183 (2009)
16. von Heusinger, A., Kanzow, C., Fukushima, M.: Newton's method for computing a normalized equilibrium in the generalized Nash game through fixed point formulation. Technical Report, Institute of Mathematics, University of Würzburg, Germany, January 2009
17. Hogan, W.W.: Point-to-set maps in mathematical programming. SIAM Rev. **15**, 591–603 (1973)
18. Janin, R.: Directional derivative of the marginal function in nonlinear programming. Math. Program. Study **21**, 110–126 (1984)
19. Jongen, H.Th., Klatte, D., Kummer, B.: Implicit functions and sensitivity of stationary points. Math. Program. **49**, 123–138 (1990)
20. Krawczyk, J.B., Uryas'ev, S.P.: Relaxation algorithms to find Nash equilibria with economic applications. Environ. Model. Assess. **5**, 63–73 (2000)
21. Luderer, B., Minchenko, L., Satsura, T.: Multivalued Analysis and Nonlinear Programming Problems with Perturbations. Kluwer, Dordrecht (2002)
22. Nabetani, K., Tseng, P., Fukushima, M.: Parametrized variational inequality approaches to generalized Nash equilibrium problems with shared constraints. Technical Report, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Japan, October 2008
23. Outrata, J.V., Kočvara, M., Zowe, J.: Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Kluwer, Dordrecht (1998)
24. Pang, J.-S., Fukushima, M.: Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games. Comput. Manag. Sci. **2**, 21–56 (2005)
25. Pang, J.-S., Ralph, D.: Piecewise smoothness, local invertibility, and parametric analysis of normal maps. Math. Oper. Res. **21**, 401–426 (1996)
26. Ralph, D., Dempe, S.: Directional derivatives of the solution of a parametric nonlinear program. Math. Program. **70**, 159–172 (1995)
27. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. A Series of Comprehensive Studies in Mathematics, vol. 317. Springer, Berlin (1998)