NONSMOOTH OPTIMIZATION REFORMULATIONS OF PLAYER CONVEX GENERALIZED NASH EQUILIBRIUM PROBLEMS[1]

Axel Dreves[2], Christian Kanzow[2], and Oliver Stein[3]

Preprint 300 October 2010

[2]University of Würzburg
Institute of Mathematics
Am Hubland
97074 Würzburg
Germany
e-mail: dreves@mathematik.uni-wuerzburg.de
kanzow@mathematik.uni-wuerzburg.de

[3]Karlsruhe Institute of Technology
Institute of Operations Research
76131 Karlsruhe
Germany
stein@kit.edu

October 27, 2010

[1]This research was partially supported by the DFG (Deutsche Forschungsgemeinschaft) under grant KA 1296/17-1.

Abstract. Using a regularized Nikaido-Isoda function, we present a (nonsmooth) constrained optimization reformulation of a class of generalized Nash equilibrium problems (GNEPs). Further we give an unconstrained reformulation of a large subclass of all GNEPs which, in particular, includes the jointly convex GNEPs. Both approaches characterize all solutions of a GNEP as minima of optimization problems. The smoothness properties of these optimization problems are discussed in detail, and it is shown that the corresponding objective functions are continuous and piecewise continuously differentiable under mild assumptions. Some numerical results based on the unconstrained optimization reformulation being applied to player convex GNEPs are also included.

Key Words: Generalized Nash equilibrium problem; jointly convex; player convex; optimization reformulation; continuity; PC^1 mapping; 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 $x^\nu \in \mathbb{R}^{n_\nu}$, and the vector $x = (x^1, \dots, x^N)^T \in \mathbb{R}^n$ with $n = n_1 + \dots + n_N$ 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 his own strategy space $X_\nu(x^{-\nu}) \subseteq \mathbb{R}^{n_\nu}$ that depends on the other players' variables $x^{-\nu}$. Typically, these sets are defined explicitly via some constraint functions, say

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

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

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

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

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

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 is an important subclass of GNEPs which are often called *jointly convex* GNEPs. In that case there are convex feasible sets for all players which depend on the rivals' strategies, but they are defined via the same convex set for all players. More precisely, there exists a common convex strategy space $Y \subseteq \mathbb{R}^n$ such that the feasible set of player ν is given by

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

or in the setting of (1), we have $g^1 = g^2 = \dots = g^N =: g$ and

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

where g is convex ("jointly") in all variables x. This special case was discussed in a number of recent papers, see [6, 10] and references therein. Here, however, we consider the more general case where the functions g^ν may be different for all players and are only assumed to be convex with respect to x^ν.

Throughout this paper, we therefore 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) For all $\nu = 1, \dots, N$ and all $i = 1, \dots, m_\nu$, the component functions $g_i^\nu : \mathbb{R}^n \to \mathbb{R}$ are continuous and, as a function of x^ν alone, convex.

A GNEP satisfying Assumption 1.1 is called *player convex*. Assumption 1.1 is rather weak in the context of GNEPs, nevertheless, for the presentation of some counterexamples, we will sometimes be slightly more general and simply assume that the strategy spaces $X_\nu(x^{-\nu})$ are nonempty, closed, and convex, but not necessarily defined by a mapping g^ν satisfying the requirements from Assumption 1.1.

So far there is a very limited number of papers that deal with the player convex case of a GNEP, see [5, 6, 7, 9, 12, 20, 21]. These papers take different approaches using penalty methods, variational inequalities, and KKT systems. The one we follow here uses a regularized Nikaido-Isoda function and is similar to one mentioned in the survey article [6]. In contrast to the one given there, however, we obtain a "real" optimization problem and not a constrained "quasi-optimization" problem.

More specifically, the current work may be viewed as an extension of the two previous papers [14, 4]. In [14], some optimization reformulations of the jointly convex GNEP were considered, with a particular emphasis on differentiable formulations. However, these differentiable formulations can only be used in order to find so-called normalized (or variational) Nash equilibria of the jointly convex GNEP, whereas many other solutions cannot be found in this way. The subsequent work [4] therefore gives a much more detailed analysis of some nonsmooth optimization reformulations which themselves were already introduced in [14]. These nonsmooth optimization reformulations have the advantage that the minima of the corresponding minimization problems coincide with the entire solution set of the underlying GNEP, hence one can also find the non-normalized solutions in this way. However, the approach from [4] is restricted to jointly convex GNEPs. The aim of this paper is therefore to extend the techniques from [4] in order to obtain (nonsmooth) optimization reformulations of the much larger class of player convex GNEPs. In particular, we present a new unconstrained optimization reformulation of GNEPs.

We note that there exist some other recent approaches which try to characterize the complete solution set of a given GNEP by using different techniques based on parameterized variational inequalities, see [11, 19]. In principle, these parameterized variational inequalities might be easier to solve than our nonsmooth optimization approach, however, both papers [11, 19] are restricted to the class of jointly convex GNEPs, and both papers do not get a full characterization of the GNEP solution set. Furthermore, a higher degree of differentiability is required by the techniques from [11, 19].

This paper is organized in the following way: Section 2 contains the constrained and unconstrained optimization reformulation of a jointly convex GNEP together with some additional elementary observations. The precise smoothness properties of these two reformulations, the constrained and the unconstrained optimization one, will be discussed in Sections 3 and 4, respectively. There, we show that the two reformulations have continuous objective functions under fairly mild conditions which may also be viewed as generalizations of the corresponding result for jointly convex GNEPs shown in [4]. We also show that the objective functions are piecewise continuously differentiable under the constant rank constraint qualification. 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. We therefore present some numerical results for the unconstrained reformulation in Section 5 using a sampling method from [1] for nonsmooth optimization. As test examples, we consider player convex problems only, since numerical results for jointly convex GNEPs may already be found in [4] for a closely related approach. We then close with some final remarks in Section 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 (piecewise continuously differentiable) function in a neighbourhood of a given point x^*, if g is continuous and there exist a neighborhood U of x^* and a finite family of continuous differentiable functions $\{G_1, G_2, \dots, G_k\}$ defined on U, such that for all $x \in U$ we have $g(x) \in \{G_1(x), G_2(x), \dots, G_k(x)\}$. 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 with ∂H 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 there exists a matrix $N \in \mathbb{R}^{n\times m}$ with $[N, M] \in \partial H(x, y)$.

2 Constrained and Unconstrained Optimization Reformulation

Here we present two new reformulations of the GNEP, one as a constrained optimization problem and the other one as an unconstrained optimization problem. The constrained reformulation is similar to the one introduced in [14] and further discussed in [4] for the case of a jointly convex GNEP, whereas the unconstrained reformulation is a generalization of an approach from [4].

The basis of both reformulations is the *Nikaido-Isoda function* (also called *Ky Fan-function*) which is defined 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 by Assumption 1.1, it is easy to see that $\Psi(x, .)$ is concave for any fixed x. Hence the *regularized Nikaido-Isoda-function*, cf. [13],

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

is uniformly concave as a function of the second argument, where $\alpha > 0$ denotes a fixed parameter. Using this function, we define

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

Note that both $X_\nu(x^{-\nu})$ and $\Omega(x)$ are closed and convex sets in view of Assumption 1.1, and that all appearing objective functions are uniformly concave or convex, respectively. Hence the mapping $V_\alpha(x)$ is well-defined for all $x \in \mathbb{R}^n$ such that $\Omega(x) \neq \emptyset$. Note also that, formally, the definition of the mapping V_α is the same as the one from [4], where its smoothness properties are discussed in detail. However, here the sets $X_\nu(x^{-\nu})$ are more general since we consider a not necessarily jointly convex GNEP, hence the properties of V_α have to be investigated in this setting.

Now let us define the set

$$\begin{aligned} W &:= \{x \in \mathbb{R}^n \mid x^\nu \in X_\nu(x^{-\nu}) \text{ for all } \nu = 1, \ldots, N\} \\ &= \{x \in \mathbb{R}^n \mid g^\nu(x) \leq 0 \text{ for all } \nu = 1, \ldots, N\}, \end{aligned} \tag{3}$$

where the second equality follows from the representation of the sets $X_\nu(x^{-\nu})$, cf. Assumption 1.1. The set W is obtained by concatenating the constraints of all players. This set will play an important role in our subsequent discussion, and we therefore begin with some simple observations.

Remark 2.1 (a) Consider a Nash game with two players having arbitrary cost functions. Player 1 controls the single variable x_1, and player 2 controls the single variable x_2 (note that we use subscripts here since these variables are real numbers in this example). Let the strategy spaces $X_1(x_2)$ and $X_2(x_1)$ be defined by the mappings

$$g^1(x) := x_1^2 - x_2^2 \quad \text{and} \quad g^2(x) := x_1^2 + x_2^2 - 1,$$

respectively. Note that these functions satisfy the properties from Assumption 1.1, but that g^1 is not convex as a function of the whole vector x. The corresponding set W is shown in Figure 1. Obviously, this set is not convex. Note also that there is no (clear) connection to the sets $\Omega(x)$. For example, taking $x := (\frac{1}{2}, \frac{1}{2})^T \in W$, an easy calculation shows that $\Omega(x) = [-\frac{1}{2}, \frac{1}{2}] \times [-\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}]$, and this set is neither a subset of W nor vice versa.

(b) Consider once again a Nash game with two players. Once more, the cost functions are arbitrary, and player 1 has the single decision variable x_1, whereas player 2 has the single decision variable x_2. Let the strategy spaces be defined by

$$X_1(x_2) = \begin{cases} \mathbb{R}, & \text{if } x_2 \neq 0, \\ \emptyset, & \text{if } x_2 = 0, \end{cases} \quad \text{and} \quad X_2(x_1) = \begin{cases} \mathbb{R}, & \text{if } x_1 \neq 0, \\ \emptyset, & \text{if } x_1 = 0. \end{cases}$$

Note that these sets are always closed and convex. In this case, however, the set W is equal to $\mathbb{R}^2 \backslash \{(x_1, x_2) \mid x_1 x_2 = 0\}$, hence W is neither closed nor convex.

(c) The previous counterexample is somewhat artificial since the sets $X_\nu(x^{-\nu})$ were not defined by some functions g^ν. Of course, if we have $X_\nu(x^{-\nu}) = \{x^\nu \mid g^\nu(x^\nu, x^{-\nu}) \leq 0\}$ for all $\nu = 1, \ldots, N$ for some continuous functions g^ν (as required in Assumption 1.1), then the set W is obviously closed. Recall, however, that Figure 1 shows that W is nonconvex in general.

(d) Let x^* be a solution of the GNEP. Then $x^* \in \Omega(x^*)$, and this implies $x^* \in W$ (see Theorem 2.2 (a) for a formal proof of this statement). In particular, W is nonempty whenever the GNEP has at least one solution. $\Diamond$

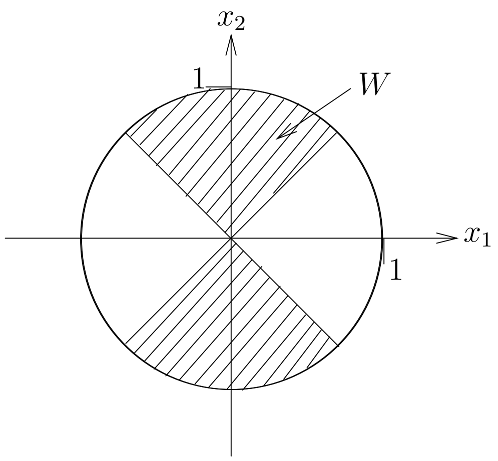

Figure 1: The set W for the example from Remark 2.1 (a)

For the case of a jointly convex GNEP, the mapping V_α was already used in [14, 4] in order to get a reformulation of this GNEP as a constrained optimization problem. In the following result, we show that this is also possible for the player convex GNEP.

Theorem 2.2 *Let Assumption 1.1 be satisfied. Then the following statements hold:*

(a) $x \in W$ *if and only if* $x \in \Omega(x)$.

(b) $V_\alpha(x) \geq 0$ *for all* $x \in W$.

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

(d) For all $x \in \mathbb{R}^n$ *with* $\Omega(x) \neq \emptyset$*, there exists a unique vector* $y_\alpha(x) := \big(y_\alpha^1(x), \ldots, y_\alpha^N(x)\big)$ *such that*
$$\arg\min_{y^\nu \in X_\nu(x^{-\nu})} \Big[\theta_\nu(y^\nu, x^{-\nu}) + \frac{\alpha}{2}\|x^\nu - y^\nu\|^2\Big] = y_\alpha^\nu(x)$$
for all $\nu = 1, \ldots, N$.

(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 $x^* = y_\alpha(x^*)$.

Proof. (a) By definition, $x \in \Omega(x)$ means $x^\nu \in X_\nu(x^{-\nu})$ for all $\nu = 1, \dots, N$, which is equivalent to $x \in W$.

(b) For all $x \in W$ we have $x \in \Omega(x)$ by part (a). Therefore

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

(c) Let x^* be a generalized Nash equilibrium. Then we have $x^* \in \Omega(x^*)$ (hence $x^* \in W$ by part (a)) and

$$\theta_\nu(x^{*,\nu}, x^{*,-\nu}) \leq \theta_\nu(x^\nu, x^{*,-\nu}) \qquad \text{for all} \quad x^\nu \in X_\nu(x^{*,-\nu}), \nu = 1, \dots, N.$$

This implies

$$\Psi_\alpha(x^*, y) = \sum_{\nu=1}^N \underbrace{\left(\theta_\nu(x^{*,\nu}, x^{*,-\nu}) - \theta_\nu(y^\nu, x^{*,-\nu})\right)}_{\leq 0} - \frac{\alpha}{2}\|x^* - y\|^2 \leq 0$$

for all $y \in \Omega(x^*)$. Therefore, we get $V_\alpha(x^*) = \max_{y \in \Omega(x^*)} \Psi_\alpha(x^*, y) \leq 0$. Together with part (b), we obtain $V_\alpha(x^*) = 0$.

Conversely assume that $x^* \in W$ (which is equivalent to $x^* \in \Omega(x^*)$ by part (a)) and $V_\alpha(x^*) = 0$. Then $\Psi_\alpha(x^*, y) \leq 0$ holds for all $y \in \Omega(x^*)$. Let $\nu \in \{1, \dots, N\}$ be a fixed player, $x^\nu \in X_\nu(x^{*,-\nu})$ and $\lambda \in (0, 1)$ arbitrary. Define $y = (y^1, \dots, y^N) \in \mathbb{R}^n$ by

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

The convexity of $X_\nu(x^{*,-\nu})$ implies $y^\mu \in X_\mu(x^{*,-\mu})$ for all $\mu = 1, \dots, N$ and therefore $y \in \Omega(x^*)$. Using this special y and exploiting the convexity of θ_ν with respect to x^ν, we get

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

Dividing both sides by $(1-\lambda)$ and taking the limit $\lambda \uparrow 1$, we see

$$\theta_\nu(x^{*,\nu}, x^{*,-\nu}) \leq \theta_\nu(x^\nu, x^{*,-\nu}).$$

Since this is true for all $x^\nu \in X_\nu(x^{*,-\nu})$ and all players $\nu = 1, \dots, N$, x^* is a generalized Nash equilibrium.

(d) For $x \in \mathbb{R}^n$ with $\Omega(x) \neq \emptyset$, the closed and convex sets $X_\nu(x^{-\nu})$ are nonempty for all $\nu = 1, \ldots, N$. The statement therefore follows from the fact that the uniformly convex function (uniform with respect to y^ν)

$$\theta_\nu(y^\nu, x^{-\nu}) + \frac{\alpha}{2}\|x^\nu - y^\nu\|^2$$

attains a unique minimum on the nonempty, closed and convex set $X_\nu(x^{-\nu})$.

(e) First, let x^* be a generalized Nash equilibrium. Then $x^* \in \Omega(x^*)$ by definition, and $V_\alpha(x^*) = 0$ in view of part (c). Therefore, we obtain

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

where $y_\alpha(x^*) := \big(y_\alpha^1(x^*), \ldots, y_\alpha^N(x^*)\big)$. With $x^* \in \Omega(x^*)$, we get $x^* = y_\alpha(x^*)$ from part (d), taking into account the uniqueness of the maximizer $y_\alpha(x^*)$.

Conversely, let $x^* \in \mathbb{R}^n$ be such that $x^* = y_\alpha(x^*)$. Then $x^* \in \Omega(x^*)$ and, therefore, $x^* \in W$ in view of part (a). Moreover, we obtain

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

and this means that x^* is a generalized Nash equilibrium by part (c). $\square$

In our subsequent discussion, we frequently use the notation $y_\alpha(x)$ for the vector

$$y_\alpha(x) := \big(y_\alpha^1(x), \ldots, y_\alpha^N(x)\big),$$

where $y_\alpha^\nu(x)$ is defined by Theorem 2.2 (d).

Theorem 2.2 implies that finding a solution of the GNEP is equivalent to finding a minimum x^* of the constrained optimization problem

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

satisfying $V_\alpha(x^*) = 0$. The set W is nonempty (at least if the GNEP has at least one solution) and closed under Assumption 1.1, but might be nonconvex, cf. Remark 2.1. As for each $x \in W$ we have $x \in \Omega(x)$ by Theorem 2.2 (a), it follows that $\Omega(x)$ is nonempty and, thus, the objective function V_α is well-defined on W. However, V_α is nondifferentiable and might even be discontinuous. An example for the latter effect is given in [4].

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 Section 3. In the following, however, we modify the previous approach and present a new *unconstrained* optimization reformulation of the GNEP.

To this end, 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 $x \in W$. Since the set W is, in general, not convex, and since we will later compute suitable projections, we define the set

$$X := cl(conv(W)). \tag{5}$$

The following is a central assumption for our subsequent analysis.

Assumption 2.3 *W is nonempty and $\Omega(x)$ is nonempty for all $x \in X$, where W and X are defined by* (3) *and* (5)*, respectively.*

Assumption 2.3 implies that the set X is always nonempty, closed, and convex. In particular, the Euclidean projection onto this set is well-defined and unique. This observation will be exploited in order to derive our unconstrained optimization reformulation. To this end, we stress that the closure in (5) is really needed, cf. Remark 2.4 (d) below.

Note also that Assumption 2.3 restricts the class of GNEPs that we will be able to deal with in our unconstrained optimization reformulation. The following remark, however, shows that the class of GNEPs satisfying Assumption 2.3 contains the jointly convex GNEPs, and is, in fact, strictly larger, but that not all GNEPs satisfy Assumption 2.3.

Remark 2.4 (a) Consider a jointly convex GNEP. Then there is a common nonempty, closed, and convex set $Y \subseteq \mathbb{R}^n$ such that

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

for all players $\nu = 1, \ldots, N$. In this case, we obviously have $W = Y$ which, in turn, implies $X = cl(conv(W)) = Y$ since Y is already closed and convex. Moreover, $W \neq \emptyset$ and $\Omega(x) \neq \emptyset$ for all $x \in X = Y$ since x belongs to $\Omega(x)$ for all $x \in X$, cf. [14]. Hence Assumption 2.3 holds for jointly convex GNEPs.

(b) Remark (a) can be generalized in the following way: Consider a GNEP with strategy sets being given by

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

for $\nu = 1, \ldots, N$, where $g^\nu : \mathbb{R}^n \to \mathbb{R}^{m_\nu}$ is continuous and convex as a function of *all* variables x. Then, again, we have $X = W$ due to the assumed continuity and convexity of all g^ν. Moreover $\Omega(x) \neq \emptyset$ for all $x \in W = X$. If a solution exists, we have $W \neq \emptyset$. Hence Assumption 2.3 also holds in this situation.

(c) An explicit example of a non-jointly convex GNEP which obviously satisfies Assumption 2.3 is the one from Remark 2.1 (a). In particular, this shows that the class of GNEPs satisfying Assumption 2.3 strictly includes the class of jointly convex GNEPs.

(d) On the other hand, there exist GNEPs which do not satisfy Assumption 2.3. To see this, consider a GNEP with two players, player 1 controlling the single variable x_1, player 2 having the single variable x_2, and the strategy spaces of both players being defined by

$$X_1(x_2) = \{x_1 \in \mathbb{R} \mid g^1(x) := 1 - x_1 x_2 \leq 0\} = \begin{cases} (-\infty, 1/x_2], & \text{if} \quad x_2 < 0, \\ \emptyset, & \text{if} \quad x_2 = 0, \\ [1/x_2, \infty), & \text{if} \quad 0 < x_2, \end{cases}$$
$$X_2(x_1) = \{x_2 \in \mathbb{R} \mid g^2(x) := x_2 - 1 \leq 0\} = (-\infty, 1].$$

Then the functions g^ν are convex in x^ν for fixed $x^{-\nu}$ and all the $X_\nu(x^{-\nu})$ are closed and convex. Moreover, we have $W = \{x \in \mathbb{R}^2 \mid x_1 x_2 \geq 1, x_2 \leq 1\}$ which is not connected and, in particular, not convex. We further get $X = cl(conv(W)) = \mathbb{R} \times (-\infty, 1]$ and $\Omega((0,0)) = \emptyset$. Hence Assumption 2.3 is violated in this case.

(e) If we have a GNEP where Assumption 1.1 holds, but there exist $x \in X$ with $\Omega(x) = \emptyset$ then the set W is not convex, because otherwise, if $W = conv(W)$, Assumption 1.1 makes W a closed set, see Remark 2.1 (c), and hence $X = cl(conv(W)) = W$, so that $\Omega(x) \neq \emptyset$ for all $x \in X = W$ would follow from Theorem 2.2 (a). These GNEPs are the hard ones, since we do not get an unconstrained reformulation for them and even the constrained optimization problem is a non-convex and therefore difficult problem. $\Diamond$

In the following definition, we modify the objective function V_α in such a way that we obtain an unconstrained optimization reformulation.

Definition 2.5 *Consider GNEPs where Assumption 2.3 holds. For those we define for all* $x \in \mathbb{R}^n$ *and* $\alpha > 0$

$$\begin{aligned} \bar{y}_\alpha(x) &:= \arg\max\nolimits_{y \in \Omega(P_X[x])} \Psi_\alpha(x, y) \qquad \textit{and} \\ \bar{V}_\alpha(x) &:= \max_{y \in \Omega(P_X[x])} \Psi_\alpha(x, y) = \Psi_\alpha(x, \bar{y}_\alpha(x)). \end{aligned}$$

Given two parameters $0 < \alpha < \beta$ *and a constant* $c > 0$*, we then define*

$$\bar{V}_{\alpha\beta}(x) := \bar{V}_\alpha(x) - \bar{V}_\beta(x) + c\|x - P_X[x]\|^2$$

for all $x \in \mathbb{R}^n$*, where* $\bar{V}_\beta(x)$ *is defined similarly to* $\bar{V}_\alpha(x)$.

The difference between the definitions of V_α and $\bar{V}_\alpha$ is that we maximize over $\Omega(x)$ in the former case, whereas we maximize over the set $\Omega(P_X[x])$ in the latter case. This is important since $\Omega(x)$ might be empty for certain $x \in \mathbb{R}^n$, whereas the projection $P_X[x]$ always exists due to the fact that X is nonempty, closed, and convex as a consequence of Assumption 2.3 and, furthermore, the set $\Omega(P_X[x])$ is (closed, convex, and) nonempty again by Assumption 2.3. Consequently, $\bar{y}_\alpha(x)$ and therefore also $\bar{V}_\alpha(x)$ are well-defined for all $x \in \mathbb{R}^n$. This, in turn, implies that $\bar{V}_{\alpha\beta}$ is well-defined for all $x \in \mathbb{R}^n$. Therefore, Assumption 2.3 guarantees that our functions are well-defined. Note that, in [4], we introduced a similar function in order to get an unconstrained optimization reformulation of the jointly convex GNEP which, however, does not need the additional term $c\|x - P_X[x]\|^2$ in the definition of $\bar{V}_{\alpha\beta}$. In fact, this additional term is not needed for the case of jointly convex GNEPs, whereas in the more general player convex case considered here, it is strictly necessary to have this term in the definition of $\bar{V}_{\alpha\beta}$. An example is given at the end of this section.

Note also that we have

$$\bar{y}_\alpha(x) = y_\alpha(x) \quad \text{and therefore} \quad \bar{V}_\alpha(x) = V_\alpha(x) \quad \text{for all } x \in X, \tag{6}$$

hence these two functions coincide on the set X. This simple observation will be used fruitfully in our subsequent analysis.

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 2.6 *Let Assumption 2.3 hold. Then we have for all $x \in \mathbb{R}^n$ the following inequalities:*

$$\frac{\beta-\alpha}{2}\|x-\bar{y}_\beta(x)\|^2 + c\|x-P_X[x]\|^2 \le \bar{V}_{\alpha\beta}(x),$$
$$\frac{\beta-\alpha}{2}\|x-\bar{y}_\alpha(x)\|^2 + c\|x-P_X[x]\|^2 \ge \bar{V}_{\alpha\beta}(x).$$

Proof. Assumption 2.3 guarantees that all involved functions are well-defined. We have $\bar{y}_\alpha(x) \in \Omega(P_X[x])$ and $\bar{y}_\beta(x) \in \Omega(P_X[x])$. Therefore

$$\bar{V}_\beta(x) = \Psi_\beta(x, \bar{y}_\beta(x)) = \max_{y \in \Omega(P_X[x])} \Psi_\beta(x,y) \ge \Psi_\beta(x, \bar{y}_\alpha(x)), \tag{7}$$

$$\bar{V}_\alpha(x) = \Psi_\alpha(x, \bar{y}_\alpha(x)) = \max_{y \in \Omega(P_X[x])} \Psi_\alpha(x,y) \ge \Psi_\alpha(x, \bar{y}_\beta(x)). \tag{8}$$

On the one hand, this implies

$$\begin{aligned}
\bar{V}_{\alpha\beta}(x) &= \bar{V}_\alpha(x) - \bar{V}_\beta(x) + c\|x-P_X[x]\|^2 \\
&\overset{(7)}{\le} \Psi_\alpha(x,\bar{y}_\alpha(x)) - \Psi_\beta(x,\bar{y}_\alpha(x)) + c\|x-P_X[x]\|^2 \\
&= \frac{\beta-\alpha}{2}\|x-\bar{y}_\alpha(x)\|^2 + c\|x-P_X[x]\|^2
\end{aligned}$$

and, on the other hand, we obtain

$$\begin{aligned}
\bar{V}_{\alpha\beta}(x) &= \bar{V}_\alpha(x) - \bar{V}_\beta(x) + c\|x-P_X[x]\|^2 \\
&\overset{(8)}{\ge} \Psi_\alpha(x,\bar{y}_\beta(x)) - \Psi_\beta(x,\bar{y}_\beta(x)) + c\|x-P_X[x]\|^2 \\
&= \frac{\beta-\alpha}{2}\|x-\bar{y}_\beta(x)\|^2 + c\|x-P_X[x]\|^2
\end{aligned}$$

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 2.7 *Under Assumption 2.3 the following statements hold:*

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

(b) x^* *is a generalized Nash equilibrium if and only if* x^* *is a minimum of* $\bar{V}_{\alpha\beta}$ *satisfying* $\bar{V}_{\alpha\beta}(x^*) = 0$.

Proof. Again Assumption 2.3 is needed to guarantee that the functions are all well-defined. The first inequality in Lemma 2.6 immediately gives

$$\bar{V}_{\alpha\beta}(x) \ge \frac{\beta-\alpha}{2}\|x-\bar{y}_\beta(x)\|^2 + c\|x-P_X[x]\|^2 \ge 0$$

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 2.2 (a) therefore implies $x^* \in W \subseteq X$. This, in turn, gives $P_X[x^*] = x^*$, and together with the fixed point characterization of Theorem 2.2 (e), we get

$$x^* = y_\alpha(x^*) = \bar{y}_\alpha(x^*),$$

where the second equality follows from (6). The second inequality of Lemma 2.6 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 from the first inequality of Lemma 2.6 that

$$0 = \bar{V}_{\alpha\beta}(x^*) \geq \frac{\beta - \alpha}{2} \|x^* - \bar{y}_\beta(x^*)\|^2 + c\|x^* - P_X[x^*]\|^2 \geq 0.$$

Since $c > 0$, this means $P_X[x^*] = x^*$, i.e. $x^* \in X$, and

$$x^* = \bar{y}_\beta(x^*) = y_\beta(x^*),$$

where, once again, we used (6). Hence x^* is a generalized Nash equilibrium by the fixed point characterization from Theorem 2.2 (e). $\square$

In the jointly convex case, it is possible to show that the additional term $c\|x - P_X[x]\|^2$ is not necessary, so we can define $\bar{V}_{\alpha\beta}$ with $c := 0$ in this case. This observation is essentially the result from [4]. For the general case it is strictly necessary, as the next example shows.

Example 2.8 Consider the two player game defined via

$$\begin{array}{llll} \text{Player 1:} & \min\limits_{x_1} x_1^2 & \text{s.t.} & x_1^2 + x_2^2 \leq 1, \\ \text{Player 2:} & \min\limits_{x_2} (x_2 + 3)^2 & \text{s.t.} & -2 \leq x_2 \leq -1. \end{array}$$

Then we have $W = X = \{(0, -1)^T\}$. If we consider the point $\hat{x} = (0, -2)^T$, we have $P_X[\hat{x}] = (0, -1)^T$ and $\Omega((0, -1)^T) = \{0\} \times [-2, -1]$. Thus we get

$$\bar{y}_\gamma(\hat{x}) = (0, -2)^T = \hat{x}$$

for all $\gamma > 0$ and this implies for $0 < \alpha < \beta$

$$\bar{V}_\alpha(\hat{x}) - \bar{V}_\beta(\hat{x}) = \Psi_\alpha(\hat{x}, \bar{y}_\alpha(\hat{x})) - \Psi_\beta(\hat{x}, \bar{y}_\beta(\hat{x})) = 0.$$

But we have $\hat{x} \notin W$ and, therefore, $\hat{x}$ is not a solution of the GNEP. This shows that we cannot skip the additional term $c\|x - P_X[x]\|^2$ in the definition of $\bar{V}_{\alpha\beta}$. $\Diamond$

Theorem 2.7 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

$$\min \bar{V}_{\alpha\beta}(x), \quad x \in \mathbb{R}^n, \tag{9}$$

in order to find solutions of a GNEP. Note, however, that we obtain this unconstrained reformulation only for the class of GNEPs which satisfy Assumption 2.3, whereas the corresponding constrained reformulation (4) holds for an arbitrary player convex GNEP, not necessarily satisfying this condition.

Similar to the constrained reformulation, however, the objective function $\bar{V}_{\alpha\beta}$ is nondifferentiable in general and, even worse, might be discontinuous. The smoothness properties of $\bar{V}_{\alpha\beta}$ will be discussed in more detail in Section 4.

3 Smoothness Properties of the Constrained Reformulation

Here we consider the constrained reformulation (4) of the GNEP with the objective function V_α from (2). Since this objective function is nondifferentiable and possibly even discontinuous, we take a closer look at the smoothness properties of this mapping. Our aim is to show the following results:

- V_α is continuous provided that either $X_\nu(x^{-\nu})$ satisfies a Slater condition or consists of a single element;
- V_α is a PC^1 function provided that it is continuous, the functions g^ν and θ_ν are twice continuously differentiable and a constant rank constraint qualification holds.

The analysis is similar to the one given in [4] for the case of jointly convex GNEPs. For the sake of completeness and since some of the subsequent results need weaker assumptions than the corresponding ones in [4] even for jointly convex GNEPs, we give all the details here.

To verify the continuity of V_α, we first recall some terminology and results from set-valued analysis. The interested reader is referred to [16, 24] for further properties.

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 $\{y^k\} \subseteq Y$ with $y^k \to y^*$ and $y^k \in \Phi(x^k)$ for all $k \geq m$;*

(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 [16]) and "inner semicontinuous mapping" (see [24]). A useful result for our subsequent analysis is the following one which follows immediately from [16, 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 neighbourhood of $\bar{x}$ and lower semicontinuous in $\bar{x}$, and $\Phi(x)$ convex in a neighbourhood of $\bar{x}$. Define*

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

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 a sufficient condition for continuity of V_α.

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

Proof. First observe that 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 the product structure $\Omega(x) = X_1(x^{-1}) \times \ldots \times X_N(x^{-N})$ it is clear that $\Omega(x)$ is closed if and only if $X_\nu(x^{-\nu})$ is closed for all $\nu = 1, \ldots, N$. The point-to-set mappings $x^{-\nu} \mapsto X_\nu(x^{-\nu}), \nu = 1, \ldots, N$, are closed for all $x \in W$ since their graphs $\{(y^\nu, x^{-\nu}) \in \mathbb{R}^{n_\nu} \times \mathbb{R}^{n-n_\nu} \mid g^\nu(y^\nu, x^{-\nu}) \leq 0\}$ are closed sets due to the assumed continuity of g^ν, see [16, Theorem 2].

Theorem 2.2 (a) implies that $\Omega(x)$ is nonempty for all $x \in W$; moreover, these sets are also convex as a consequence of Assumption 1.1. Theorem 2.2 (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 $x \in W$, namely $y_\alpha(x)$. Lemma 3.2 therefore implies that $x \to \{y_\alpha(x)\}$, viewed as a point-to-set mapping, is lower semicontinuous and closed in $x^* \in W$. This implies that the single-valued function $x \mapsto y_\alpha(x)$ is continuous at $x^* \in W$. Hence, $V_\alpha(x) = \Psi_\alpha(x, y_\alpha(x))$, being a composition of continuous maps, is also continuous at $x^* \in W$. □

In view of Theorem 3.3, our next aim is to find a condition guaranteeing that $X_\nu(x^{-\nu}) = \{y^\nu \in \mathbb{R}^{n_\nu} \mid g^\nu(y^\nu, x^{-\nu}) \leq 0\}$ is lower semicontinuous for all $\nu = 1, \ldots, N$. The proof of the following lemma is based on the fact that lower semicontinuity can be obtained by the *Slater condition*, saying that for a given $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$ there exists a $y^\nu \in \mathbb{R}^{n_\nu}$ such that $g^\nu(y^\nu, x^{-\nu}) < 0$ for all $\nu = 1, \ldots, N$.

Lemma 3.4 *Suppose that Assumption 1.1 holds. Then the functions $y_\alpha^\nu, \nu = 1, \ldots, N$, and V_α are continuous in $x^* \in W$ provided the Slater condition holds at $X_\nu(x^{*,-\nu})$ for all $\nu = 1, \ldots, N$.*

Proof. Let $x^* \in W$ be given such that $X_\nu(x^{*,-\nu})$ satisfies the Slater condition for all $\nu = 1, \ldots, N$. By Assumption 1.1 we can apply [16, Theorem 12] and get lower semicontinuity of the point-to-set mapping $x^{-\nu} \mapsto X_\nu(x^{-\nu})$ at $x^{*,-\nu}$ for all $\nu = 1, \ldots, N$. Therefore, also the point-to-set mapping $x \mapsto \Omega(x)$ is lower semicontinuous at x^*. Hence continuity of y_α, in particular of all components $y_\alpha^\nu, \nu = 1, \ldots, N$, and of V_α at x^* follow from Theorem 3.3. $\square$

Unfortunately, it seems natural that many GNEPs possess points $x^* \in W$ at which the Slater condition is violated for some $\nu = 1, \ldots, N$. In the example from Remark 2.1 (a), e.g., for $x^* = (0, 0)$ the set $X_1(0)$ violates the Slater condition.

Whereas the latter example is degenerate, at least in the jointly convex case with a bounded common strategy space $Y = \{x \in \mathbb{R}^n \mid g(x) \leq 0\}$ the failure of the Slater condition at certain (boundary) points of $W = Y$ cannot be avoided: for any $\nu = 1, \ldots, N$ the domain of X_ν,

$$\operatorname{dom} X_\nu \; = \; \{x^{-\nu} \in \mathbb{R}^{n-n_\nu} \mid X_\nu(x^{-\nu}) \neq \emptyset\},$$

is closed and bounded as the orthogonal projection of Y to $\mathbb{R}^{n-n_\nu}$ and, by the continuity of g, at all boundary points $\bar{x}^{-\nu}$ of $\operatorname{dom} X_\nu$ the Slater condition has to be violated in $X_\nu(\bar{x}^{-\nu}) = \{y^\nu \in \mathbb{R}^{n_\nu} \mid g(y^\nu, \bar{x}^{-\nu}) \leq 0\}$. In view of $\bar{x}^{-\nu} \in \operatorname{dom} X_\nu$, on the other hand, there exists some $\bar{y}^\nu \in \mathbb{R}^{n_\nu}$ with $(\bar{y}^\nu, \bar{x}^{-\nu}) \in Y$ or, equivalently, $g(\bar{y}^\nu, \bar{x}^{-\nu}) \leq 0$. As $X_\nu(\bar{x}^{-\nu})$ violates the Slater condition, the latter inequality has to be satisfied with equality and, thus, under mild assumptions $(\bar{y}^\nu, \bar{x}^{-\nu})$ is a boundary point of $W = Y$ (e.g., if Y itself satisfies the Slater condition). Note that simple examples show that in general not all boundary points of W correspond to the violation of the Slater condition in some player's strategy space.

In the following we will prove continuity of y_α and V_α at points $x \in W$ also in the case that the Slater condition is violated in one or more strategy spaces $X_\nu(x^{-\nu})$, $\nu = 1, \ldots, N$, as long as the strategy spaces then collapse to singletons. In view of Theorem 2.2 (a), they then have to coincide with the corresponding set $\{x^\nu\}$.

Theorem 3.5 *Suppose Assumption 1.1 holds, and assume that for each $x^* \in W$ and all $\nu = 1, \ldots, N$ the set $X_\nu(x^{*,-\nu})$ either satisfies the Slater condition or coincides with the singleton $\{x^{*,\nu}\}$. Then the functions y_α and V_α are continuous on W.*

Proof. In view of Theorem 3.3 we have to show lower semicontinuity of $x \mapsto \Omega(x)$ at x^*, i.e. for all sequences $\{x^k\} \subseteq W$ with $\lim_{k\to\infty} x^k = x^*$ and all $y^* \in \Omega(x^*)$ we have to find a sequence $\{y^k\}$ converging to y^* with $y^k \in \Omega(x^k)$ for all $k \in \mathbb{N}$ sufficiently large. We will define the elements of y^k componentwise for each player $\nu = 1, \ldots, N$. For those $\nu \in \{1, \ldots, N\}$, where $X_\nu(x^{*,-\nu})$ satisfies the Slater condition, the mapping $x^{-\nu} \mapsto X_\nu(x^{-\nu})$ is lower semicontinuous at $x^{*,-\nu}$ by Lemma 3.4, and hence a sequence $\{y^{k,\nu}\}$ converging to $y^{*,\nu}$ with $y^{k,\nu} \in X_\nu(x^{k,-\nu})$ for all k sufficiently large exists. For all the other $\nu \in \{1, \ldots, N\}$ we have $X_\nu(x^{*,-\nu}) = \{x^{*,\nu}\} = \{y^{*,\nu}\}$ by assumption. Defining $y^{k,\nu} := x^{k,\nu}$ we get a sequence $\{y^{k,\nu}\}$ converging to $y^{*,\nu}$ with $y^{k,\nu} = x^{k,\nu} \in X_\nu(x^{k,-\nu})$ by Theorem 2.2 (a), since $x^k \in W$. Therefore $x \mapsto \Omega(x)$ is lower semicontinuous at x^*, since we have a sequence $\{y^k\}$ with $\lim_{k\to\infty} y^k = y^*$ and $y^{k,\nu} \in X_\nu(x^{k,-\nu})$ for all $\nu = 1, \ldots, N$, i.e. $y^k \in \Omega(x^k)$ for all k sufficiently large. $\square$

Hence the optimization reformulation (4) of the GNEP is at least a continuous problem under the assumptions of Theorem 3.5. This observation immediately gives the existence result from part (a) of the following note.

Remark 3.6 (a) If the set W is nonempty and bounded and for each $x^* \in W$ and all $\nu = 1, \ldots, N$, the set $X_\nu(x^{*,-\nu})$ either satisfies the Slater condition or coincides with the singleton $\{x^{*,\nu}\}$, it is an immediate consequence of the Weierstraß theorem that the optimization problem (4) possesses a globally minimal point.

(b) In general, there are three possible situations which fully describe the relationship between the GNEP and the optimization problem (4):

- The GNEP has a solution, and therefore the optimization problem (4) also has a solution in view of Theorem 2.2 (with zero as optimal function value).
- The GNEP has no solution, but the optimization problem (4) has a solution (then, necessarily, with a positive optimal function value).
- Neither the GNEP nor the optimization problem (4) have a solution.

Under the assumption from part (a), the last case cannot occur. In this situation, the optimization problem (4) therefore characterizes the solvability of a GNEP: The existing minimum of (4) is a solution of the GNEP if and only if the optimal function value is zero.

(c) Here we give an instance for the second case mentioned in (b). In Example 2.8 we have a nonempty and bounded set $W = \{(0, -1)^T\}$, the single valued set $X_1(-1) = \{0\}$ and the set $X_2(0) = [-2, -1]$, which satisfies the Slater condition, but we do not have a solution, since for the only possible point $(0, -1)^T \in W$, we get $y_\alpha((0, -1)^T) = (0, -2)^T \neq (0, -1)^T$ for all $\alpha \leq 2$. A short calculation shows that $V_\alpha((0, -1)^T) = 3 - \alpha/2$ holds for all $\alpha \leq 2$, so the optimal value of the optimization problem (4) is strictly positive. $\Diamond$

In our subsequent analysis we will show that (4) has, in fact, a piecewise continuously differentiable objective function under some stronger assumptions. This additional smoothness property is highly important from a practical point of view since it implies that several algorithms for nonsmooth optimization problems can be applied to the problem (4).
To this end it will be useful to define the function

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

where

$$m := m_1 + \ldots + m_N$$

with m_ν being given by Assumption 1.1. This assumption also implies that all component functions h_i are convex as a function of y alone. Furthermore, the function h has the nice property that

$$y \in \Omega(x) \Longleftrightarrow h(x,y) \le 0 \tag{10}$$

for any given x.

Now we require some stronger smoothness properties of the defining functions θ_ν and g^ν.

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

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

$$\max_y \; \Psi_\alpha(x,y) \quad \text{s.t.} \quad h(x,y) \le 0. \tag{11}$$

Let

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

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

$$\max_y \; \Psi_\alpha(x,y) \quad \text{s.t.} \quad h_i(x,y) = 0 \;\; (i \in I). \tag{12}$$

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 (12). Then the KKT-system of this problem reads

$$\nabla_y L_\alpha^I(x,y,\lambda) = -\nabla_y \Psi_\alpha(x,y) + \sum_{i \in I} \lambda_i \nabla_y h_i(x,y) = 0, \quad h_i(x,y) = 0 \;\; \forall i \in I. \tag{13}$$

This can be written as a nonlinear system of equations

$$\Phi_\alpha^I(x,y,\lambda) = 0 \qquad \text{with} \qquad \Phi_\alpha^I(x,y,\lambda) := \begin{pmatrix} \nabla_y L_\alpha^I(x,y,\lambda) \\ h_I(x,y) \end{pmatrix}, \tag{14}$$

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 and therefore omitted.

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 positive definiteness assumption of the Hessian $\nabla^2_{yy} L^I(x, y, \lambda)$ can 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 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 $x^* \in W$ 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 $x \in N \cap W$.

Note that the previous definition requires the same rank only for those $x \in N$ which also belong to the common feasible set W; this is why we call this assumption the feasible CRCQ, although, in our subsequent discussion, we will simply speak of the CRCQ condition when we refer to Assumption 3.9. This feasible CRCQ has also been used before in [8], for example, where the authors simply call this condition the CRCQ.

The following result is motivated by [23] (see also [15]) 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 $x^* \in W$ be given, and suppose that the solution mapping $y_\alpha : W \to \mathbb{R}^n$ of (11) is continuous in a neighbourhood of x^* (see Theorem 3.5 for sufficient conditions). Then there exists a neighbourhood $\hat{N}$ of $x^* \in W$ such that y_α is a PC^1 function on $\hat{N} \cap W$ 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 $x^* \in W$ 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 $x \in N \cap W$. Furthermore, for any such $x \in N \cap W$, let us denote by

$$\mathcal{M}(x) := \{\lambda \in \mathbb{R}^m \mid (y_\alpha(x), \lambda) \text{ is a KKT point of (11)}\}$$

the set of all Lagrange multipliers of the optimization problem (11). Since CRCQ holds at x^*, it is easy to see that CRCQ also holds for all $x \in W$ sufficiently close to x^*. Without loss of generality, let us say that CRCQ holds for all $x \in N \cap W$ with the same neighbourhood N as before. Then it follows from a result in [17] that the set $\mathcal{M}(x)$ is nonempty for all $x \in N \cap W$. This, in turn, implies that the set

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

is also nonempty for all x in a sufficiently small neighbourhood of x^*, say, again, for all $x \in N \cap W$ (see [15] for a formal proof), where $\operatorname{supp}(\lambda)$ denotes the support of the nonnegative vector λ, i.e.,

$$\operatorname{supp}(\lambda) := \{ i \mid \lambda_i > 0 \}.$$

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., [23, 15].

Step 2: Here we show that, for every $x \in N \cap W$ 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 $x \in N \cap W$ 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 $\operatorname{supp}(\lambda_\alpha^I(x)) \subseteq I$. Since $(x, y_\alpha(x), \lambda_\alpha^I(x))$ satisfies the KKT conditions of the optimization problem (11), $[\lambda_\alpha^I(x)]_i = 0$ 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 $x^* \in W$ 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 $\big(y^I(\cdot), \lambda^I(\cdot)\big) : N^I(x^*) \to N^I(y_\alpha(x^*), \lambda^*)$ 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 $x^* \in W$ be given such that the CRCQ holds, choose $I \in \mathcal{B}(x^*)$ arbitrarily, and let $\lambda^* \in \mathcal{M}(x^*)$ with $\operatorname{supp}(\lambda^*) \subseteq I$ 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 (14). 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 $x^* \in W$ 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 $x \in \hat{N} \cap W$ arbitrarily. Step 2 shows that, for each $I \subseteq \mathcal{B}(x)\big(\subseteq \mathcal{B}(x^*)\big)$, there exist 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 $\big(y^I(\cdot), \lambda^I(\cdot)\big) : N^I(x^*) \to N^I(y_\alpha(x^*), \lambda^*)$ 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 $\Phi_\alpha^I(x, y, \lambda) = 0$. 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 $x \in \hat{N} \cap W$ 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 $\{x^k\} \subseteq W$ with $\{x^k\} \to x^*$ and a corresponding sequence of index sets $I_k \in \mathcal{B}(x^k)$ such that

$$\big(y_\alpha(x^k), \lambda_\alpha^{I_k}(x^k)\big) \notin N^{I_k}(y_\alpha(x^*), \lambda^*) \qquad \text{for all } k \in \mathbb{N}.$$

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 in x^*, 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

$$-\nabla_y \Psi_\alpha(x^k, y_\alpha(x^k)) + \sum_{i \in I} [\lambda_\alpha^I(x^k)]_i \nabla_y h_i(x^k, y_\alpha(x^k)) = 0 \tag{15}$$

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 (15) and using once again the continuity of the solution mapping $y_\alpha(\cdot)$ in x^* 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 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 for each $x^* \in W$ and all $\nu = 1, \dots, N$ the set $X_\nu(x^{*,-\nu})$ either satisfies the Slater condition or conicides with the singleton $\{x^{*,\nu}\}$, and that Assumption 3.9 holds in $x^* \in W$. Then y_α and V_α are PC^1 functions in a neighbourhood of x^* in W.*

Proof. From Corollary 3.5, 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^*. □

To close this section, we want to give an example showing that the (feasible) CRCQ from Assumption 3.9, without the Slater condition, does not even imply continuity.

Example 3.12 Consider the two player game where each player ν controls a single variable $x_\nu, \nu = 1, 2$. The cost functions are defined by

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

and the strategy sets are given by

$$X_1(x_2) := \{y_1 \in \mathbb{R} \mid y_1^2 x_2^2 \le 0\} = \begin{cases} \mathbb{R}, & \text{if } x_2 = 0, \\ \{0\}, & \text{if } x_2 \neq 0, \end{cases}$$
$$X_2(x_1) := \{y_2 \in \mathbb{R} \mid y_2^2 x_1^2 \le 0\} = \begin{cases} \mathbb{R}, & \text{if } x_1 = 0, \\ \{0\}, & \text{if } x_1 \neq 0, \end{cases}$$

i.e., the strategy sets are defined by the functions $g^1(x) := g^2(x) := x_1^2 x_2^2$ which satisfy the requirements from Assumptions 1.1 and 3.7 (whereas the Slater condition is violated). We therefore have $W = \{x \in \mathbb{R}^2 \mid x_1 = 0 \text{ or } x_2 = 0\}$ and $X = cl(conv(W)) = \mathbb{R}^2$. An easy calculation shows that

$$y_\alpha(x) = \frac{1}{1+\alpha} \begin{cases} (1, 1), & \text{if } x = (0, 0), \\ (1 + \alpha x_1, 0), & \text{if } x_1 \neq 0, x_2 = 0, \\ (0, 1 + \alpha x_2), & \text{if } x_1 = 0, x_2 \neq 0, \\ (0, 0), & \text{if } x_1 \neq 0, x_2 \neq 0. \end{cases}$$

Using this expression for $y_\alpha(x)$ and Theorem 2.2 (e), we deduce that the GNEP has two solutions given by $(0, 1)^T$ and $(1, 0)^T$. With the function $h : \mathbb{R}^4 \to \mathbb{R}^2$, $h(x, y) = \begin{pmatrix} y_1^2 x_2^2 \\ y_2^2 x_1^2 \end{pmatrix}$, we have

$$\nabla_y h(x, y_\alpha(x)) = \begin{pmatrix} 2y_\alpha^1(x) x_2^2 & 0 \\ 0 & 2y_\alpha^2(x) x_1^2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

for all $x \in X = \mathbb{R}^2$ which shows, in particular, that the (feasible) CRCQ condition from Assumption 3.9 holds everywhere. But, obviously, the function y_α is not continuous in any point of W except $(-\frac{1}{\alpha}, 0)^T$ and $(0, -\frac{1}{\alpha})^T$, in particular, it is discontinuous in the two solutions. Moreover, this function is discontinuous in $(0, 0)^T$ even if we view it as a mapping on W only. ◊

4 Smoothness Properties of the Unconstrained Reformulation

In contrast to the previous section where a general player convex GNEP was considered, this section deals with GNEPs satisfying Assumption 2.3. Recall, however, that this is still a rather large class of GNEPs including, in particular, the jointly convex ones, cf. the observations from Remark 2.4. Under Assumption 2.3, we have the unconstrained optimization reformulation (9) with the objective function $\bar{V}_{\alpha\beta}$ from Definition 2.5. Furthermore, throughout this section we denote by X the set defined in (5).

We will show that the PC^1 property, that was shown for the constrained reformulation in the previous section, also holds for the unconstrained reformulation and, with a further assumption, also the continuity property transfers to the unconstrained reformulation. The proofs of these smoothness properties are similar (though not identical) to the proofs given in the previous section, so that we concentrate on the differences in the proofs without recapitulating all the details. We also refer to [4] for similar considerations in the context of jointly convex GNEPs, although the additional term $c\|x - P_X[x]\|^2$ from the definition of $\bar{V}_{\alpha\beta}$ does not occur in [4].

Our first aim is to obtain a continuity result for $\bar{V}_{\alpha\beta}$. Since the projection mapping is continuous the additional term $c\|x - P_X[x]\|^2$ is continuous, hence we only need continuity of $\bar{y}_\alpha$ for arbitrary $\alpha > 0$ to get this property for $\bar{V}_\alpha$ and $\bar{V}_{\alpha\beta}$. In Theorem 3.5 we used the property $x \in \Omega(x)$ for all $x \in W$. The problem occuring here is that the corresponding property $P_X[x] \in \Omega(P_X[x])$ is only valid for $x \in W$ but not necessarily for $x \in X$. To prove a continuity result for the unconstrained reformulation we need the additional assumption of uniform continuity of the functions $g_i^\nu(y^\nu, .) : \mathbb{R}^{n-n_\nu} \to \mathbb{R}, \nu = 1, \dots, N, i = 1, \dots, m_\nu$ for all $y^\nu \in \mathbb{R}^{n_\nu}$.

Theorem 4.1 *Suppose Assumptions 1.1 and 2.3 hold and further assume that the functions* $g_i^\nu(y^\nu, .) : \mathbb{R}^{n-n_\nu} \to \mathbb{R}, \nu = 1, \dots, N, i = 1, \dots, m_\nu$ *are uniformly continuous for all* $y^\nu \in \mathbb{R}^{n_\nu}$*. Then* $\bar{V}_{\alpha\beta}$ *is continuous in* $x^* \in \mathbb{R}^n$ *provided the sets* $X_\nu(P_X[x^*]^{-\nu})$ *are either single-valued or satisfy the Slater condition.*

Proof. As in the proof of the corresponding Theorem 3.5 for the constrained formulation, it is enough to show that the point-to-set maps $x \mapsto X_\nu(P_X[x]^{-\nu}), \nu = 1, \dots, N$ are closed on $\mathbb{R}^n$ and lower semicontinuous in x^*. The proof of closedness is analogous to the constrained formulation and so is the proof of lower semicontinuity in the case where $X_\nu(P_X[x^*]^{-\nu})$ satisfies the Slater condition.

Hence it remains to show lower semicontinuity when $X_\nu(P_X[x^*]^{-\nu})$ is single valued. Therefore let an $x^* \in \mathbb{R}^n$ and an arbitrary but fixed $\nu \in \{1, \dots, N\}$ be given such that we have a single valued set $X_\nu(P_X[x^*]^{-\nu}) = \{y^{*,\nu}\}$. For a given sequence $\{x^k\} \subseteq \mathbb{R}^n$ with $x^k \to x^*$ we have to show the existence of a sequence $\{y^{k,\nu}\} \subseteq \mathbb{R}^{n_\nu}$ with $y^{k,\nu} \to y^{*,\nu}$ and $y^{k,\nu} \in X_\nu(P_X[x^k]^{-\nu})$ for all k sufficiently large.

Define the function $\bar{g}(y^\nu, x^{-\nu}) := \max_{i=1,\dots,m_\nu} g_i^\nu(y^\nu, x^{-\nu})$ which is still uniformly continuous and convex in y^ν. Further define $K := \{k \in \mathbb{N} \mid y^{*,\nu} \notin X_\nu(P_X[x^k]^{-\nu})\}$. For $k \notin K$ we simply define $y^{k,\nu} := y^{*,\nu}$. If K is a finite set the proof is already complete. Otherwise consider the subsequence $\{y^{k,\nu}\}_{k\in K}$. We have $\bar{g}(y^{*,\nu}, P_X[x^k]^{-\nu}) > 0$ and, since $X_\nu(P_X[x^k]^{-\nu})$ is nonempty by Assumption 2.3, there exists a $w^{k,\nu} \in \mathbb{R}^{n_\nu}$ such that $\bar{g}(w^{k,\nu}, P_X[x^k]^{-\nu}) \leq 0$. Continuity of $\bar{g}$ implies the existence of an $y^{k,\nu} \in X_\nu(P_X[x^k]^{-\nu})$ on the line segment from $w^{k,\nu}$ to $y^{*,\nu}$ with $\bar{g}(y^{k,\nu}, P_X[x^k]^{-\nu}) = 0$. It remains to show that $y^{k,\nu}$ converges to $y^{*,\nu}$.

First of all we have $\lim_{k\in K} \bar{g}(y^{k,\nu}, P_X[x^k]^{-\nu}) = 0$. The uniform continuity of $\bar{g}$ together with the continuity of the projection map imply

$$\lim_{k\in K} |\bar{g}(y^{k,\nu}, P_X[x^k]^{-\nu}) - \bar{g}(y^{k,\nu}, P_X[x^*]^{-\nu})| = 0,$$

and thus we obtain

$$\lim_{k\in K} \bar{g}(y^{k,\nu}, P_X[x^*]^{-\nu}) = 0. \tag{16}$$

We have $\bar{g}(z^\nu, P_X[x^*]^{-\nu}) > 0$ for all $z^\nu \neq y^{*,\nu}$, because $X_\nu(P_X[x^*]^{-\nu}) = \{y^{*,\nu}\}$. Therefore $y^{*,\nu}$ is a strict global minimum of the convex function $\bar{g}(., P_X[x^*]^{-\nu})$ which implies

$$\varepsilon := \min_{a^\nu \in bd(B_1(y^{*,\nu}))} \bar{g}(a^\nu, P_X[x^*]^{-\nu}) > 0,$$

where $bd(B_1(y^{*,\nu})) := \{a^\nu \mid \|a^\nu - y^{*,\nu}\| = 1\}$ is the boundary of the ball $B_1(y^{*,\nu})$ with centre $y^{*,\nu}$ and radius 1. With convexity of $\bar{g}(\cdot, P_X[x^*]^{-\nu})$ we get

$$\bar{g}(y^\nu, P_X[x^*]^{-\nu}) \geq \varepsilon \quad \text{for all} \quad y^\nu \notin B_1(y^{*,\nu}).$$

This together with (16) shows that $\{y^{k,\nu}\} \in B_1(y^{*,\nu})$ for all $k \in K$ sufficiently large. But this implies boundedness of the whole sequence $\{y^{k,\nu}\}$ and thus the existence of an accumulation point $\hat{y}^\nu$. Closedness of the point-to-set mapping $x \mapsto X_\nu(P_X[x]^{-\nu})$ therefore shows $\hat{y}^\nu \in X_\nu(P_X[x^*]^{-\nu}) = \{y^{*,\nu}\}$. Since this is true for all accumulation points, we have convergence of the sequence $\{y^{k,\nu}\}$ to $y^{*,\nu}$, which completes the proof. □

Our next aim is to show that the function $\bar{V}_{\alpha\beta}$ is a PC^1 mapping under suitably adopted assumptions. To this end, we first define the function

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

This function will play the role of the mapping h from Section 3. In particular, it has the corresponding property that, for any given x,

$$y \in \Omega(P_X[x]) \Longleftrightarrow \bar{h}(x, y) \leq 0. \tag{17}$$

This implies that $\bar{y}_\alpha(x)$ is the unique solution of

$$\max_y \ \Psi_\alpha(x, y) \quad \text{s.t.} \quad \bar{h}(x, y) \leq 0. \tag{18}$$

Note, however, that (in contrast to the function h) the function $\bar{h}$ is not differentiable in general (even if all g^ν are differentiable) due to the projection term inside the definition of $\bar{h}$. This causes some technical difficulties in generalizing the PC^1 property to the unconstrained reformulation. However $\bar{h}$ is a PC^1 mapping if all g^ν are smooth and the projection mapping is PC^1. The latter holds in view of [22] under the smoothness conditions of Assumption 3.7 and a constant rank constraint qualification in a version that we define next.

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 $x \in N$.

Assumption 4.2 is slightly stronger than the feasible CRCQ from Assumption 3.9 since we consider a full-dimensional neighbourhood N of x^*, whereas in Assumption 3.9 we only consider a feasible neighbourhood of x^*.

Consider the optimization problem (18) once again. Let

$$\bar{I}(x) := \{i \in \{1, \dots, m\} \mid \bar{h}_i(x, \bar{y}_\alpha(x)) = 0\}$$

be the set of active inequality constraints. Consider, for a fixed subset $I \subseteq \bar{I}(x)$, the equality constrained problem

$$\max_y \Psi(x, y) \quad \text{s.t.} \quad \bar{h}_i(x, y) = 0 \;\; (i \in I). \tag{19}$$

Let

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

be the corresponding Lagrangian. Then the KKT conditions of (19) are equivalent to the nonlinear system of equations

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

In the proof of Theorem 3.10 (Step 3), we applied the implicit function theorem to the corresponding mapping Φ_α^I from the previous section. In contrast to Φ_α^I, however, $\bar{\Phi}_\alpha^I$ is not differentiable everywhere, hence the standard implicit function theorem cannot be used in the current situation. However, under suitable assumptions including the CRCQ condition, the projection map and, therefore, also the function $\bar{\Phi}_\alpha^I$ is a PC^1 mapping. Hence we need an implicit function theorem for PC^1 equations. The following is such a result that was obtained in [4] as a consequence of an inverse function theorem for PC^1 mappings from [8].

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 $x \in U$.*

Now we are in a position to generalize Theorem 3.10 to the unconstrained optimization reformulation.

Theorem 4.4 *Suppose that Assumptions 1.1, 2.3, and 3.7 hold. Furthermore, suppose that $x^* \in \mathbb{R}^n$ is such that the CRCQ condition from Assumption 4.2 is satisfied at x^* and the solution mapping $\bar{y}_\alpha : \mathbb{R}^n \to \mathbb{R}^n$ of (18) is continuous in a neighbourhood of x^* (see Theorem 4.1 for a sufficient condition). Then $\bar{y}_\alpha$ is a PC^1 function in a neighbourhood of 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 Section 3, let us introduce the sets

$$\bar{\mathcal{M}}(x) := \{\lambda \in \mathbb{R}^m \mid (\bar{y}_\alpha(x), \lambda) \text{ is a KKT point of (18)}\}$$

and

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

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: Using the notation of (18), $\bar{L}_\alpha^I$ and $\bar{\Phi}_\alpha^I$, 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 $\lambda_\alpha^I(x) \in \bar{\mathcal{M}}(x)$ such that $\bar{\Phi}_\alpha^I(x, \bar{y}_\alpha(x), \lambda_\alpha^I(x)) = 0$, 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 $\bar{\Phi}_\alpha^I$ defined in Step 2 is only a PC^1 function, but not continuously differentiable (in contrast to the mapping Φ_α^I 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 $I \in \mathcal{B}(x^*)$ with corresponding multiplier λ^* be given. Since $\bar{\Phi}_\alpha^I(x, y, \lambda)$ is continuously differentiable with respect to y and λ, it follows that $\pi_{(y,\lambda)} \partial \Phi_\alpha^I(x^*, \bar{y}_\alpha(x^*), \lambda^*)$ 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 $\big(y^I(\cdot), \lambda^I(\cdot)\big) : N^I(x^*) \to N^I(\bar{y}_\alpha(x^*), \lambda^*)$ such that $y^I(x^*) = \bar{y}_\alpha(x^*), \lambda^I(x^*) = \lambda^*$ and $\bar{\Phi}_\alpha^I(x, y^I(x), \lambda^I(x)) = 0$ 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 neighbourhood of any x^* satisfying the CRCQ condition from Assumption 4.2. $\square$

Altogether, we get the following corollary.

Corollary 4.5 *Suppose that Assumptions 1.1, 2.3 and 3.7 hold. Moreover, suppose that the functions $g_i^\nu(y^\nu, .) : \mathbb{R}^{n-n_\nu} \to \mathbb{R}$ are uniformly continuous for all $y^\nu \in \mathbb{R}^{n_\nu}$, Assumption 4.2 holds in $x^* \in \mathbb{R}^n$ and that the sets $X_\nu(P_X[x]^{-\nu}), \nu = 1, \dots, N$ either satisfy the Slater condition or coincide with a singleton for all x sufficiently close to x^*. Then $\bar{V}_{\alpha\beta}$ is a PC^1 function in a neighbourhood of x^*.*

Proof. Since the projection mapping has PC^1 property, the additional term $c\|x - P_X[x]\|^2$ also has. From Theorem 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 $\bar{V}_{\alpha\beta} = \bar{V}_\alpha(x) - \bar{V}_\beta(x) + c\|x - P_X[x]\|^2$ are PC^1 mappings in a neighborhood of x^*. $\square$

Being a PC^1 mapping, it follows that $\bar{V}_{\alpha\beta}$ is, in particular, directionally differentiable, locally Lipschitz continuous and semismooth, cf. [2].

Unfortunately we were not able to give simple sufficient conditions for stationary points of $\bar{V}_{\alpha\beta}$ to be global minima of the function. As the following example shows this is probably a difficult task.

Example 4.6 Consider the jointly convex 2-player game defined via

$$\theta_1(x) := \frac{1}{2}(x_1 + 2)^2 \quad \text{and} \quad \theta_2(x) := \frac{1}{2}(x_2 + 2)^2 \quad \text{and}$$

$$X := \{x \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 2, x_2 - x_1 \leq 0, x_1 - x_2 - 1 \leq 0\}.$$

A simple calculation shows that, for all $\alpha \in (0, 1]$ and all $x \in X$, we have

$$\bar{y}_\alpha(x) = (\max\{0, x_2\}, x_1 - 1).$$

Moreover, it is not difficult to see that the only solution of this GNEP is given by $(0, -1)$. Taking $0 < \alpha < \beta \leq 1$, we obtain from the previous observation that $\bar{y}_\alpha(x) = \bar{y}_\beta(x)$ and, therefore,

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

Thus we have for all $x_2 > 0$

$$\bar{V}_{\alpha\beta}(x) = \frac{\beta - \alpha}{2}\left((x_1 - x_2)^2 + (x_2 - x_1 + 1)^2\right) = (\beta - \alpha)\left(\left(x_2 - x_1 + \frac{1}{2}\right)^2 + \frac{1}{4}\right).$$

Hence we see, that for all $x \in X$ with $x_2 > 0$ and $x_2 - x_1 = -\frac{1}{2}$, the function $\bar{V}_{\alpha\beta}$ has local minima and hence stationary points with function value $\frac{\beta-\alpha}{4} > 0$. Therefore these points are not global minima and hence not solutions of the GNEP. $\Diamond$

5 Numerical Results

As in the recent paper [4] we used the robust gradient sampling algorithm from [1] to test our unconstrained optimization reformulation using the objective function $\bar{V}_{\alpha\beta}$. The MATLAB® implementation 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`. Since the algorithm computes random sampling gradients, it may generate different solutions even if the same starting point is used. With probability one every limiting point of a sequence generated by this method is a Clarke stationary point. The algorithm stops if the norm of the vector with the smallest Euclidian norm in the convex hull of the sampled gradients is less than 10^{-6}. Apart from using standard parameter settings, we use the three values $\alpha = 0.02$, $\beta = 0.05$ and $c = 10^3$ which define our objective function. For every function evaluation we have to solve optimization problems to obtain $\bar{y}_\alpha(x)$ and $\bar{y}_\beta(x)$. This is done by using the `fmincon` solver from the MATLAB® Optimization Toolbox. Further if $x \notin X$ we have to compute the projection onto the convex set X for which we either used the `quadprog` (for polyhedral X) or again the `fmincon` solver (for nonpolyhedral X) from the MATLAB® Optimization Toolbox.

Since we already presented numerical results for jointly convex GNEPs in the previous paper [4], we now concentrate on player convex GNEPs. First we present four small examples with 2 players, each contolling a single variable, where the solution sets can be computed analytically. These examples were tested with 100 randomly choosen starting vectors and the distribution of the solutions is shown in scatter plots in Figure 2. There one can see that the solutions obtained by the algorithm spread over the entire solution set.

Example 5.1 This two player game is defined via the following problems:

$$\begin{array}{llll} \text{Player 1:} & \min\limits_{x_1} (x_1 - 2)^2 & \text{s.t.} & x_1 + x_2 \leq 1, \\ \text{Player 2:} & \min\limits_{x_2} (x_2 - 2)^2 & \text{s.t.} & x_1 + x_2 \leq 1,\ x_2 - x_1 \leq 0. \end{array}$$

An elementary calculation shows that the solution set is given by

$$\left\{ (\lambda, 1 - \lambda) \mid \lambda \in \left[\frac{1}{2}, 2 \right] \right\}.$$

Our starting points were chosen randomly in $[-10, 10]^2$. $\Diamond$

Example 5.2 Here we consider once again a GNEP with two players, whose optimization problems are given by

$$\begin{array}{llll} \text{Player 1:} & \min\limits_{x_1} \frac{1}{2}x_1^2 + 3x_1x_2 & \text{s.t.} & x_1 - 3x_2 \leq 2,\ -3x_1 + x_2 \leq 2,\ x_1 + x_2 \leq 1, \\ \text{Player 2:} & \min\limits_{x_2} \frac{1}{2}x_2^2 + 3x_1x_2 & \text{s.t.} & x_1 - 3x_2 \leq 2,\ -3x_1 + x_2 \leq 2. \end{array}$$

Based on some elementary considerations, it is possible to verify that the corresponding solution set is given by

$$\{(0,0),(-1,-1)\} \cup \left\{(\lambda, -\frac{2}{3} + \frac{1}{3}\lambda) \mid \lambda \in \left[\frac{1}{5}, 1\right]\right\} \cup \left\{(\lambda, 2 + 3\lambda) \mid \lambda \in \left[-\frac{3}{5}, -\frac{1}{3}\right]\right\}.$$

The starting points were taken randomly from $[-2,2]^2$. ◊

Example 5.3 This game is defined by

$$\begin{array}{llll} \text{Player 1:} & \min\limits_{x_1} (x_1 - 1)^2 & \text{s.t.} & x_1^2 - x_2^2 \leq 0, \\ \text{Player 2:} & \min\limits_{x_2} (x_2 - x_1)^2 & \text{s.t.} & x_1^2 + x_2^2 \leq 1. \end{array}$$

In this example the set $W := \{x \in \mathbb{R}^2 \mid x_1^2 - x_2^2 \leq 0, x_1^2 + x_2^2 \leq 1\}$ is not convex, but the set $X = cl(conv(W)) = \{x \in \mathbb{R}^2 \mid -\frac{1}{2} \leq x_1 \leq \frac{1}{2}, x_1^2 + x_2^2 \leq 1\}$ can be computed easily, cf. Figure 1. The solution set of the game is

$$\left\{(\lambda, \lambda) \mid \lambda \in \left[0, \frac{1}{\sqrt{2}}\right]\right\}$$

and the starting vector was choosen randomly in $[-2,2]^2$. ◊

Example 5.4 Here we have the following situation for the two players:

$$\begin{array}{llll} \text{Player 1:} & \min\limits_{x_1} (x_1 - 1)^2 & \text{s.t.} & x_1^2 - x_2^2 \leq 0,\ x_1^2 + x_2^2 \leq 1, \\ \text{Player 2:} & \min\limits_{x_2} (x_2 - \frac{1}{2}x_1 - 2)^2 & \text{s.t.} & x_1^2 + x_2^2 \leq 1. \end{array}$$

The sets W and X are the same as in the previous Example 5.3. The solution set is

$$\left\{(\lambda, \sqrt{1-\lambda^2}) \mid \lambda \in \left[0, \frac{1}{\sqrt{2}}\right]\right\}$$

and the starting vector was choosen randomly in $[-2,2]^2$. ◊

The further test examples we used are all taken from the appendix of the paper [7]. We only report results for the test runs with the starting vectors given therein. Recall, however, that two runs with the same starting vector can give different solutions, since the algorithm uses a random sampling strategy. Table 5 shows the results, where the column x^0 contains the equal value of all variables of the starting vector, column (It.) is the number of iterations, x^* the computed solution and $\bar{V}_{\alpha\beta}(x^*)$ the corresponding objective function value at the solution. Since the MATLAB® function `fmincon` computes a solution at a certain precision it is possible that for $\bar{y}_\alpha(x) \approx \bar{y}_\beta(x)$ the objective function values $\bar{V}_{\alpha\beta}(x)$ get slightly negative, although theoretically we have $\bar{V}_{\alpha\beta}(x) \geq 0$ for all $x \in \mathbb{R}^n$.

6 Final Remarks

This paper gives a constrained reformulation of all solutions of a player convex GNEP and therefore generalizes the results of [4] where a reformulation is given only for jointly convex GNEPs. Further we introduce a new unconstrained reformulation for general player convex GNEPs where the sets $\Omega(x)$ are nonempty for certain x. Both reformulations characterize all solutions of general player convex GNEPs as solutions of optimization problems. These problems are continuous if a Slater condition is satisfied on the sets $\Omega(x)$ or if all degenerate strategy spaces are singletons. If we additionally suppose a constant rank constrained qualification we get a PC^1 objective function, which allows the application of nonsmooth optimization software for finding a solution of a GNEP.

References

[1] J.V. Burke, A.S. Lewis, and M.L. Overton: *A robust gradient sampling algorithm for nonsmooth, nonconvex optimization.* SIAM Journal on Optimization 15, 2005, pp. 751–779.

[2] R.W. Chaney: *Piecewise C^k functions in nonsmooth analysis.* Nonlinear Analysis: Theory, Methods & Applications 15, 1990, pp. 649–660.

[3] F.H. Clarke: *Optimization and Nonsmooth Analysis.* SIAM, Philadelphia, 1990.

[4] A. Dreves and C. Kanzow: *Nonsmooth optimization reformulations characterizing all solutions of jointly convex generalized Nash equilibrium problems.* Computational Optimization and Applications, to appear.

[5] F. Facchinei, A. Fischer, and V. Piccialli: *Generalized Nash equilibrium problems and Newton methods.* Mathematical Programming 117, 2009, pp. 163–194.

[6] F. Facchinei and C. Kanzow: *Generalized Nash equilibrium problems.* Annals of Operations Research 175, 2010, pp. 177–211.

[7] F. Facchinei and C. Kanzow: *Penalty methods for the solution of generalized Nash equilibrium problems.* SIAM Journal on Optimization 20, 2010, pp. 2228–2253.

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

[9] F. Facchinei and J.-S. Pang: *Exact penalty functions for generalized Nash problems.* In G. Di Pillo and M. Roma (eds): *Large Scale Nonlinear Optimization.* Springer, Berlin, 2006, pp. 115–126.

[10] F. Facchinei and J.-S. Pang: *Nash equilibria: the variational approach.* In: Convex Optimization in Signal Processing and Communications, D.P. Palomar and Y.C. Eldar, editors, Cambridge University Press, 2010, pp. 443–493.

[11] F. Facchinei and S. Sagratella: *On the computation of all solutions of jointly convex generalized Nash equilibrium problems.* Optimization Letters, to appear.

[12] M. Fukushima: *Restricted generalized Nash equilibria and controlled penalty algorithm.* Computational Management Science, to appear.

[13] G. Gürkan and J.-S. Pang: *Approximations of Nash equilibria.* Mathematical Programming 117, 2009, pp. 223–253.

[14] A. von Heusinger and C. Kanzow: *Optimization reformulations of the generalized Nash equilibrium problem using Nikaido-Isoda-Type Functions.* Computational Optimization and Applications 43, 2009, pp. 353–377.

[15] A. von Heusinger, C. Kanzow, and M. Fukushima: *Newton's method for computing a normalized equilibrium in the generalized Nash game through fixed point formulation.* Mathematical Programming, to appear.

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

[17] R. Janin: *Directional derivative of the marginal function in nonlinear programming.* Mathematical Programming Study 21, 1984, pp. 110–126.

[18] 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.

[19] K. Nabetani, P. Tseng, and M. Fukushima: *Parametrized variational inequality approaches to generalized Nash equilibrium problems with shared constraints.* Computational Optimization and Applications, to appear.

[20] J.-S. Pang: *Computing generalized Nash equilibria.* Technical Report, Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, MD, October 2002.

[21] 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.

[22] J.-S. Pang and D. Ralph: *Piecewise smoothness, local invertibility, and parametric analysis of normal maps.* Mathematics of Operations Research 21, 1996, pp. 401–426.

[23] D. Ralph and S. Dempe: *Directional derivatives of the solution of a parametric nonlinear program.* Mathematical Programming 70, 1995, pp. 159–172.

[24] R.T Rockafellar and R.J.-B. Wets: *Variational Analysis.* A Series of Comprehensive Studies in Mathematics, Vol. 317, Springer, Berlin, Heidelberg, 1998.

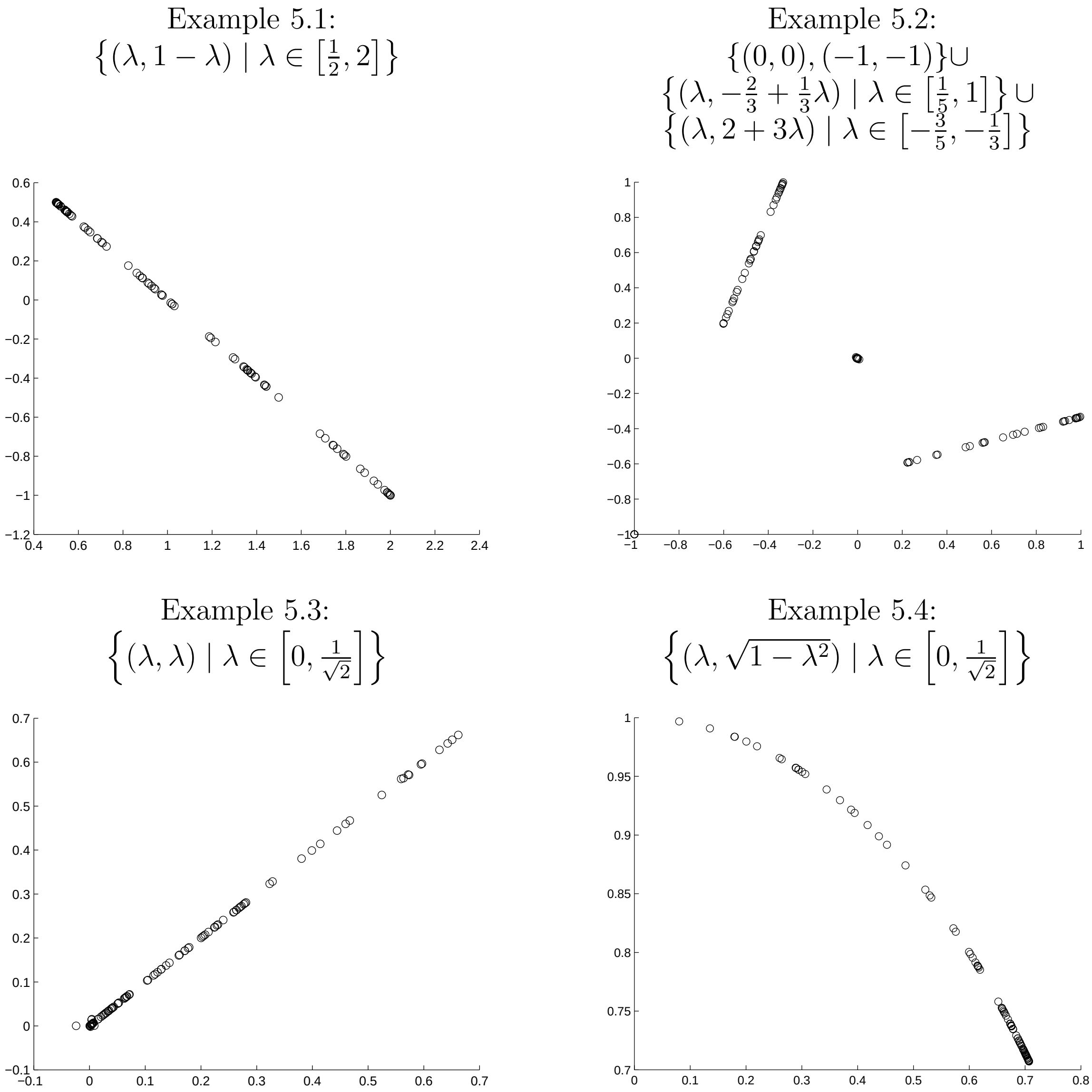

Figure 2: Scatter plots of the solutions of the 2-dimensional GNEPs

Example	x^0	It.	x^*	$\bar{V}_{\alpha\beta}(x^*)$
A.1	0.01	78	$(0.3004, 0.0692, 0.0694, 0.0694, 0.0692,$ $0.0698, 0.0696, 0.0690, 0.0690, 0.0700)$	-4.8e-7
	0.1	49	$(0.3000, 0.0694, 0.0693, 0.0694, 0.0695,$ $0.0694, 0.0695, 0.0694, 0.0695, 0.0694)$	3.2e-9
	1	53	$(0.3000, 0.0696, 0.0694, 0.0691, 0.0697,$ $0.0691, 0.0695, 0.0693, 0.0699, 0.0691)$	-4.8e-7
A.2	0.01	89	$(0.3008, 0.0247, 0.0316, 0.0294, 0.2444,$ $0.3180, 0.0120, 0.0118, 0.0121, 0.0110)$	-2.4e-6
	0.1	66	$(0.3026, 0.0364, 0.0250, 0.0374, 0.2425,$ $0.2944, 0.0134, 0.0122, 0.0178, 0.0110)$	-6.5e-6
	1	103	$(0.3004, 0.0326, 0.0280, 0.0251, 0.2233,$ $0.3342, 0.0116, 0.0115, 0.0115, 0.0118)$	9.1e-8
A.3	0	84	$(-0.3812, -0.1265, -0.9972, 0.3900, 1.1653, 0.0502, 0.0186)$	-3.3e-7
	1	64	$(-0.3730, -0.1259, -0.9764, 0.3806, 1.1357, 0.0489, 0.0200)$	-3.8e-7
	10	81	$(-0.3874, -0.1236, -1.0119, 0.3988, 1.1876, 0.0521, 0.0171)$	-2.3e-7
A.4	0	45	$(1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000)$	1.5e-12
	1	33	$(1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000)$	1.3e-11
	10	38	$(1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000)$	5.9e-12
A.5	0	56	$(0.0000, 0.2028, 0.0000, 0.0000, 0.0722, 0.0255, 0.0001)$	-5.0e-7
	1	41	$(0.0001, 0.2024, 0.0001, 0.0001, 0.0714, 0.0253, 0.0003)$	-4.3e-7
	10	61	$(0.0008, 0.1970, 0.0010, 0.0019, 0.0577, 0.0300, 0.0039)$	-3.7e-8
A.6	0	49	$(1.0000, 1.0000, 1.0000, 1.4167, 1.0000, 1.0000, 1.0000)$	3.8e-11
	1	38	$(1.0000, 1.0000, 1.0000, 1.4167, 1.0000, 1.0000, 1.0000)$	6.6e-11
	10	35	$(1.0000, 1.0000, 1.0000, 1.4167, 1.0000, 1.0000, 1.0000)$	2.6e-11
A.7	0	115	$(1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000,$ $1.0000, 1.0000, 1.0000, 1.0000, 1.8427, 1.0000, 1.0000,$ $1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000)$	1.8e-11
	1	120	$(1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000,$ $1.0000, 1.0000, 1.0000, 1.0000, 1.8427, 1.0000, 1.0000,$ $1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000)$	1.7e-11
	10	149	$(1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000,$ $1.0000, 1.0000, 1.0000, 1.0000, 1.8427, 1.0000, 1.0000,$ $1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000)$	1.5e-11
A.8	0	23	$(0.6005, 0.3995, 0.9009)$	0
	1	20	$(0.5550, 0.4447, 0.8322)$	7.7e-9
	10	18	$(0.6284, 0.3716, 0.9421)$	1.2e-8
A.9 (a)	0	79	available on request	3.7e-9
A.9 (b)	0	157	available on request	5.0e-6

Table 1: Numerical results for test problems from [7]