ON DIFFERENTIABILITY PROPERTIES OF PLAYER CONVEX GENERALIZED NASH EQUILIBRIUM PROBLEMS[1]

Nadja Harms[2], Christian Kanzow[2], and Oliver Stein[3]

Preprint 307 March 2012

[2]University of Würzburg
Institute of Mathematics
Emil-Fischer-Str. 30
97074 Würzburg
Germany
e-mail: nadja.harms@mathematik.uni-wuerzburg.de
kanzow@mathematik.uni-wuerzburg.de

[3]Karlsruhe Institute of Technology
Institute of Operations Research
Kaiserstr. 12
76131 Karlsruhe
Germany
stein@kit.edu

March 21, 2012

[1]This research was partially supported by the DFG (Deutsche Forschungsgemeinschaft) under grants KA 1296/18-1 und STE 772/13-1

Abstract. This article studies differentiability properties for a reformulation of a player convex generalized Nash equilibrium problem as a constrained and possibly nonsmooth minimization problem. By using several results from parametric optimization we show that, apart from exceptional cases, all locally minimal points of the reformulation are differentiability points of the objective function. This justifies a numerical approach which basically ignores the possible nondifferentiabilities.

Key Words: Generalized Nash equilibrium problem, player convexity, Nikaido-Isoda function, Gâteaux differentiability, Fréchet differentiability, parametric optimization.

AMS subject classifications: 91A10, 90C31.

1 Introduction

We consider generalized Nash equilibrium problems (GNEPs) in which each player $\nu \in \{1,\ldots,N\}$ controls a decision variable $x^\nu \in \mathbb{R}^{n_\nu}$ and where $x = (x^1,\ldots,x^N) \in \mathbb{R}^n$ with $n = n_1 + \ldots + n_N$ describes the decision vector of all players. To emphasize the role of player ν's variable x^ν within the vector x, we often write $x = (x^\nu, x^{-\nu})$. For each player ν a cost function $\theta_\nu(\cdot, x^{-\nu})$ and a strategy space

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

are given, which both depend on the other players' decisions $x^{-\nu}$. All functions $\theta_\nu : \mathbb{R}^n \to \mathbb{R}$ and $g^\nu : \mathbb{R}^n \to \mathbb{R}^{m_\nu}$ are assumed to be at least continuous.

The difference to a classical Nash equilibrium problem (NEP) lies in the $x^{-\nu}$-dependence of the strategy spaces $X_\nu(x^{-\nu})$, that is, in a standard NEP, each player has a fixed strategy space X_ν. While NEPs were introduced in [17], GNEPs go back to [3, 1]. For a survey on theory, applications, and algorithms for the solution of GNEPs, we refer to [7, 10].

In a GNEP, each player $\nu \in \{1,\ldots,N\}$ wishes to minimize his cost function $\theta_\nu(\cdot, x^{-\nu})$ over his strategy space $X_\nu(x^{-\nu})$, that is, to solve the problem

$$Q_\nu(x^{-\nu}): \qquad \min_{x^\nu} \theta_\nu(x^\nu, x^{-\nu}) \quad \text{s.t.} \quad g^\nu(x^\nu, x^{-\nu}) \le 0,$$

where the other players' strategies enter as the parameter vector $x^{-\nu}$. If $S_\nu(x^{-\nu})$ denotes the set of optimal points of $Q_\nu(x^{-\nu})$, $\nu = 1,\ldots,N$, then the generalized Nash equilibrium problem can formally be stated as

$$\textit{GNEP:} \qquad \text{find some } x \in \mathbb{R}^n \text{ with } x^\nu \in S_\nu(x^{-\nu}),\; \nu = 1,\ldots,N.$$

A solution point $x^\star$ of *GNEP* is also called a *generalized Nash equilibrium*.

A natural assumption to make *GNEP* numerically tractable is the convexity of the problems $Q_\nu(x^{-\nu})$, $\nu = 1,\ldots,N$, in the respective player variable x^ν. Assumption 1.1 will be a standing assumption throughout this paper.

Assumption 1.1 (Player convexity) *For each $\nu \in \{1,\ldots,N\}$ and any given $x^{-\nu}$, the defining functions θ_ν, g_i^ν, $i = 1,\ldots,m_\nu$, of $Q_\nu(x^{-\nu})$ are convex with respect to the player variable x^ν.*

GNEPs satisfying Assumption 1.1 are called *player-convex.* Apart from very few exceptions, see [4, 19, 20], it is the most general form of a GNEP studied in the literature. A widely studied sub-class are the *jointly-convex* GNEPs where $g^1 = g^2 = \ldots = g^N =: g$ holds and the components of the constraint function g are convex in the whole vector $x = (x^1,\ldots,x^N)$. Although we will not study this problem class in detail in the present paper, Remark 4.14 below will summarize the implications of joint convexity for our approach.

The focus of this paper is on smoothness properties of a suitable optimization reformulation of GNEPs based on the Nikaido-Isoda function. The original reference [18] uses this mapping as a theoretical tool, however, later it was also explored algorithmically in [15, 25] where certain fixed-point algorithms are considered. A regularized version of the Nikaido-Isoda function was then introduced in [12] and later explored in [13, 5] in order to reformulate the jointly-convex GNEP either as a constrained or unconstrained optimization problem. This work was extended very recently to the larger class of player-convex GNEPs, see [6]. A major drawback of the corresponding optimization problems, however, is the fact that they typically have nonsmooth objective functions. The aim of this paper is therefore to have a closer look at the smoothness properties of these objective functions. Preliminary results of this kind, especially regarding the continuity and piecewise smoothness, can already be found in [6]. Here we show further structural properties, in particular, our main result indicates that, apart from some degenerate points, the objective functions are differentiable in (local or global) minima. Note that this result is also of some importance for jointly-convex GNEPs, although there exist differentiable optimization formulations of this class of problems, cf. [13]. However, the solutions of these differentiable formulations do not characterize the full solution set of jointly-convex GNEPs (only so-called *normalized* solutions can be obtained), whereas here we consider reformulations characterizing all solutions of both jointly-convex and player-convex GNEPs.

The paper is structured as follows. In Section 2 we recall from [6] the reformulation of a *GNEP* as a possibly nonsmooth constrained minimization problem and rewrite its objective function in a way which makes it more accessible for results from parametric optimization. Section 3 reviews from [6] a result on the continuity of the objective function and relates it to interior points of the domain of the objective function, where subsequently its differentiability properties can be studied. Our main result on differentiability of the objective function at locally minimal points of the reformulation (Theorem 4.10) is developed in Section 4. This result motivates the application of certain smooth optimization techniques, and Section 5 therefore presents some numerical results. Section 6 closes the article with final remarks. The main results are illustrated by accompanying examples.

2 Reformulation as a Constrained Optimization Problem

This section briefly reviews from [6] how *GNEP* can be equivalently replaced by a (possibly nonsmooth) constrained optimization problem. In the following, we will consider the optimal value function resulting from maximization of the so-called *Nikaido-Isoda function* ([18])

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

for parameter x over the variable

$$y \;\in\; \Omega(x) \;:=\; X_1(x^{-1}) \times \ldots \times X_N(x^{-N}).$$

As, under Assumption 1.1, all strategy spaces $X_\nu(x^{-\nu})$ are closed convex sets, so is their product space $\Omega(x)$ for any $x \in \mathbb{R}^n$. Moreover, the Nikaido-Isoda function is concave with respect to y.

To guarantee the existence of a maximal point of $\psi(x, \cdot)$ even on unbounded sets $\Omega(x)$, in the following we will replace ψ by the *regularized Nikaido-Isoda function* ([12])

$$\psi_\alpha(x, y) \;:=\; \sum_{\nu=1}^{N} \left[\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu})\right] - \frac{\alpha}{2}\|x - y\|_2^2$$

with $\alpha > 0$ (and $\psi_0 = \psi$). As ψ_α is strongly concave in y for $\alpha > 0$, a unique maximal point $y(x)$ exists and, thus, the optimal value function

$$V(x) \;:=\; \max_{y \in \Omega(x)} \psi_\alpha(x, y)$$

is real-valued exactly on the domain

$$\operatorname{dom}\Omega \;:=\; \{x \in \mathbb{R}^n |\; \Omega(x) \neq \emptyset\}$$

of the set-valued mapping $\Omega : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$. As $x \in \Omega(x)$ holds if and only if $g^\nu(x) \leq 0$ is satisfied for all $\nu = 1, \ldots, N$, it is natural to define the set

$$W \;:=\; \{x \in \mathbb{R}^n |\; g^\nu(x) \leq 0,\; \nu = 1, \ldots, N\}. \tag{1}$$

The central properties of V and W are summarized in the following result, whose proof may be found in [6].

Proposition 2.1 *The following statements hold:*

(a) $x \in \Omega(x)$ if and only if $x \in W$; in particular, we have $W \subset \operatorname{dom}\Omega$, so that V is real-valued on W.

(b) $V(x) \geq 0$ for all $x \in W$.

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

(d) $x^\star$ is a generalized Nash equilibrium if and only if $x^\star = y(x^\star)$ holds, that is, $x^\star$ is a fixed point of the mapping $x \mapsto y(x)$.

Proposition 2.1 (a)–(c) show that the computation of a generalized Nash equilibrium $x^\star$ is equivalent to finding an optimal point $x^\star$ of

$$P: \qquad \min V(x) \quad \text{s.t.} \quad x \in W \tag{2}$$

with $V(x^\star) = 0$. As the choice of the regularization parameter $\alpha > 0$ is irrelevant for this result, in the following we will not explicitly state the dependence on α of the optimal values $V(x)$, the optimal points $y(x)$, or the problem P. If the unique solvability of maximizing

$\psi_\alpha(x,\cdot)$ over $\Omega(x)$ is clear for other reasons, we will also allow the choice $\alpha = 0$, that is, employment of the original Nikaido-Isoda function.

Theoretical and numerical results on the solution of P obviously depend on the structures of the objective function V and of the feasible set W. While not much can be said about the set W, basic continuity and differentiability properties of V on W for player convex GNEPs were studied in [6]. The present paper complements this with additional interesting properties of V at its points of nondifferentiability.

To study its structural properties, we rewrite the function V for any $x \in W$ as

$$\begin{aligned}
V(x) &= \max_{y\in\Omega(x)} \psi_\alpha(x,y) \\
&= \max_{y\in\Omega(x)} \left(\sum_{\nu=1}^{N} \left[\theta_\nu(x^\nu,x^{-\nu}) - \theta_\nu(y^\nu,x^{-\nu})\right] - \frac{\alpha}{2}\|x-y\|_2^2 \right) \\
&= \max_{y\in\Omega(x)} \sum_{\nu=1}^{N} \left(\theta_\nu(x^\nu,x^{-\nu}) - \theta_\nu(y^\nu,x^{-\nu}) - \frac{\alpha}{2}\|x^\nu-y^\nu\|_2^2 \right) \\
&= \sum_{\nu=1}^{N} \left[\theta_\nu(x^\nu,x^{-\nu}) - \min_{y^\nu\in X_\nu(x^{-\nu})} \left(\theta_\nu(y^\nu,x^{-\nu}) + \frac{\alpha}{2}\|x^\nu-y^\nu\|_2^2 \right) \right] \\
&= \sum_{\nu=1}^{N} \left[\theta_\nu(x) - \varphi_\nu(x)\right] \qquad (3)
\end{aligned}$$

with the optimal value functions

$$\varphi_\nu(x) := \min_{y^\nu\in X_\nu(x^{-\nu})} \left[\theta_\nu(y^\nu,x^{-\nu}) + \frac{\alpha}{2}\|x^\nu-y^\nu\|_2^2 \right] \qquad (4)$$

of the (convex and uniquely solvable) problems

$$Q_\nu(x): \qquad \min_{y^\nu} \theta_\nu(y^\nu,x^{-\nu}) + \frac{\alpha}{2}\|x^\nu-y^\nu\|_2^2 \quad \text{s.t.} \quad g^\nu(y^\nu,x^{-\nu}) \le 0$$

for $\nu = 1,\ldots,N$. Clearly, the structural properties of V heavily depend on the structural properties of the functions φ_ν.

In the following, $y^\nu(x)$ will denote the unique optimal point of $Q_\nu(x)$, $\nu = 1,\ldots,N$, for $x \in W$, hence we can rewrite the optimal value functions φ_ν as

$$\varphi_\nu(x) = \theta_\nu(y^\nu(x),x^{-\nu}) + \frac{\alpha}{2}\|x^\nu - y^\nu(x)\|_2^2.$$

It is easy to see that $(y^1(x),\ldots,y^N(x))$ coincides with the unique maximizer $y(x)$ of $\psi_\alpha(x,\cdot)$ on $\Omega(x)$.

3 Continuity and the Domain of V

Before we turn our attention to differentiability properties, let us briefly recall sufficient conditions for continuity of V on W.

Lemma 3.1 ([6, Lemma 3.4]) *Let Assumption 1.1 hold, and for $\bar{x} \in W$ as well as for all $\nu \in \{1, \dots, N\}$, let $X_\nu(\bar{x}^{-\nu})$ satisfy the Slater condition. Then the functions y^ν, φ_ν, $\nu = 1, \dots, N$, and V are continuous at $\bar{x}$.*

With the 'degenerate point set'

$$D_1 := \{x \in W | \text{ for some } \nu = 1, \dots, N \text{ the set } X_\nu(x^{-\nu}) \\ \text{violates the Slater condition}\}$$

Lemma 3.1 guarantees continuity of V on $W \setminus D_1$. As explained in [6] and illustrated in Example 3.2 below, one has to expect that D_1 is nonempty. This was the motivation to develop a weaker sufficient condition for continuity of V on W, for which the interested reader is referred to [6, Theorem 3.5].

The following example illustrates these continuity properties and will also serve to illustrate differentiability properties below.

Example 3.2 Consider a player convex GNEP with $N = 2$, $n_1 = n_2 = 1$, $\theta_1(x) = x_1$, $\theta_2(x) = x_2$, $g_1^1(x) = -2x_1 + x_2$, $g_1^2(x) = x_1^2 + x_2^2 - 1$, $g_2^2(x) = -x_1 - x_2$. Then for all $x \in W = \{x \in \mathbb{R}^2 | \ x_1^2 + x_2^2 \leq 1, \ -x_1 \leq x_2 \leq 2x_1\}$ (cf. Fig. 1) the problems $Q_1(x)$ and $Q_2(x)$ are easily seen to be uniquely solvable for $\alpha = 0$. Note that we have $x^1 = x^{-2} = x_1$ and $x^2 = x^{-1} = x_2$.

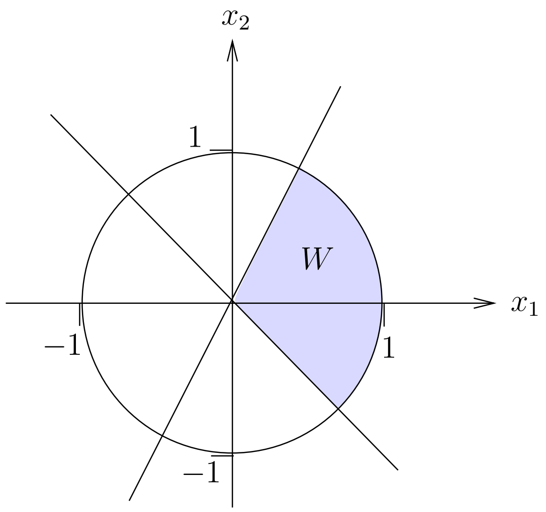

Figure 1: Illustration of the set W in Example 3.2

In fact, for $x \in W$ we obtain the strategy spaces

$$X_1(x_2) = \left[\frac{x_2}{2}, +\infty\right),$$
$$X_2(x_1) = \left[\max\left\{-x_1, -\sqrt{1 - x_1^2}\right\}, \sqrt{1 - x_1^2}\right],$$

so that the optimal points as well as optimal values of $Q_1(x)$ and $Q_2(x)$ are

$$\begin{aligned} y_1(x) &= \varphi_1(x) &=& \frac{x_2}{2}, \\ y_2(x) &= \varphi_2(x) &=& \max\left\{-x_1, -\sqrt{1-x_1^2}\right\}. \end{aligned}$$

Due to (3) this results in

$$\begin{aligned} V(x) &= \theta_1(x)+\theta_2(x)-\varphi_1(x)-\varphi_2(x) \\ &= x_1+x_2+\min\left\{x_1, \sqrt{1-x_1^2}\right\}-\frac{x_2}{2}. \end{aligned}$$

Note that, in spite of player convexity, V is a *concave* function as the minimum of two smooth concave functions (compare also Fig. 2 and Remark 3.10 below).

For all $x \in W$ the set $X_1(x_2)$ obviously satisfies the Slater condition. However, the set $X_2(x_1)$ satisfies the Slater condition only for values $x_1 \neq 1$, whereas $X_2(1) = \{0\}$ is a singleton. This results in $D_1 = \{(1,0)\}$ and shows that D_1 can easily be nonempty. Lemma 3.1 then yields continuity of V only on $W \setminus D_1$. On the other hand, direct inspection shows that V is continuous even on all of W where, however, V as 'infinite slope' at the point $x = (1,0)$. We remark that, for the present example, the improvement of Lemma 3.1 by [6, Theorem 3.5] also yields continuity of V at $(1,0)$.

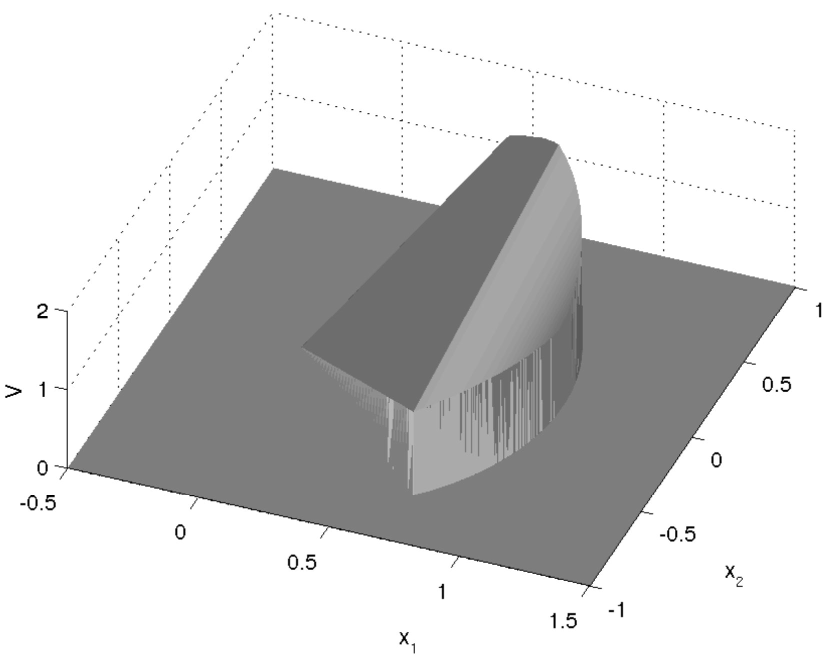

Figure 2: V on W for Example 3.2

It is not hard to see that the function V has a unique globally minimal point on W at the origin. Its optimal value is zero, so that the origin is the unique generalized Nash equilibrium in view of Proposition 2.1. However, one can also show that V has a locally minimal point on W at $(1,0)$ with value one. We point out that the latter point is the element of D_1, so that D_1 is not only nonempty, but, at least in the present example, also contains a 'structurally relevant' point. $\Diamond$

In the remainder of this section we will see that the set D_1 also plays a crucial role for differentiability properties of V. In fact, in Section 4 we shall study differentiability of V at points in the topological interior of the domain of V, where in Section 2 we have seen that the domain of V,

$$\operatorname{dom} V \;=\; \{x \in \mathbb{R}^n \mid V(x) \in \mathbb{R}\}$$

coincides with the domain of the set-valued mapping $\Omega(x) = X_1(x^{-1}) \times \ldots \times X_N(x^{-N})$,

$$\operatorname{dom} \Omega \;=\; \{x \in \mathbb{R}^n \mid \Omega(x) \neq \emptyset\}.$$

Hence, their topological interiors satisfy

$$\operatorname{int} \operatorname{dom} V \;=\; \operatorname{int} \operatorname{dom} \Omega. \tag{5}$$

The following example shows that, despite the fact that we have $W \subset \operatorname{dom} \Omega = \operatorname{dom} V$, we cannot expect the inclusion $W \subset \operatorname{int} \operatorname{dom} V$ to hold.

Example 3.3 In the situation of Example 3.2 we have $\operatorname{int} \operatorname{dom} V = \operatorname{int} \operatorname{dom} \Omega = (-1/\sqrt{2}, 1) \times \mathbb{R}$, so that

$$W \setminus \{(1,0)\} \;=\; W \cap \operatorname{int} \operatorname{dom} V.$$

Recall that $D_1 = \{(1,0)\}$ holds in this example, so that we arrive at $W \setminus D_1 = W \cap \operatorname{int} \operatorname{dom} V$. $\Diamond$

The next result guarantees that the inclusion $W \setminus D_1 \subset W \cap \operatorname{int} \operatorname{dom} V$ is also true in general.

Lemma 3.4 *Let Assumption 1.1 hold. Then we have*

$$W \setminus D_1 \;\subset\; W \cap \operatorname{int} \operatorname{dom} V. \tag{6}$$

Proof. In view of (5), the assertion is shown if we can prove the relation

$$W \setminus D_1 \;\subset W \cap \operatorname{int} \operatorname{dom} \Omega.$$

Choose $\bar{x} \in W \setminus D_1$. Then we have $\bar{x} \in W$, and for all $\nu = 1, \ldots, N$ there exists some $\bar{y}^\nu \in \mathbb{R}^{n_\nu}$ with $g^\nu(\bar{y}^\nu, \bar{x}^{-\nu}) < 0$. Due to the continuity of the functions g^ν, we can choose a neighborhood U of $\bar{x}$ such that for all $x \in U$ and $\nu = 1, \ldots, N$ also $g^\nu(\bar{y}^\nu, x^{-\nu}) < 0$ is satisfied. In particular, for all $x \in U$ each set $X_\nu(x^{-\nu})$, $\nu = 1, \ldots, N$, is nonempty, so that U is contained in $\operatorname{dom} \Omega$. This shows the assertion. $\square$

Lemma 3.4 will allow us to study differentiability properties of V on the set $W \setminus D_1$ in Section 4. While Lemma 3.4 does, of course, not exclude that also some elements of D_1 are contained in $\operatorname{int} \operatorname{dom} V$, we conjecture that, under mild additional assumptions, actually equality holds in (6). In the sequel, we will at least show this fact under certain conditions, including Assumption 3.5 below, though we believe that this assumption can be relaxed. Note, however, that we will use this assumption only in Theorem 3.9 below, but not in the remainder of this paper.

Assumption 3.5 (Joint constraint convexity)
All functions g_i^ν, $i = 1, \ldots, m_\nu$, $\nu = 1, \ldots, N$, are convex.

Clearly, under Assumption 3.5 the set W is convex. From now on we will also assume that all defining functions of *GNEP* are at least continuously differentiable.

Assumption 3.6 (Smoothness) *For each $\nu \in \{1, \ldots, N\}$ the functions θ_ν, g_i^ν, $i = 1, \ldots, m_\nu$, are continuously differentiable.*

In the sequel $\nabla \theta_\nu$ will stand for the column vector of partial derivatives of θ_ν, and $D\theta_\nu = \nabla^\intercal \theta_\nu$ will denote the corresponding row vector, etc.

Furthermore, for each $\nu = 1, \ldots, N$, we let $I^\nu = \{1, \ldots, m_\nu\}$ denote the index set of inequality constraints of player ν, we put

$$W_\nu \ := \ \{x \in \mathbb{R}^n |\ g_i^\nu(x) \leq 0,\ i \in I^\nu\} \tag{7}$$

and, for $x \in W_\nu$, we define the active index set

$$I_0^\nu(x) \ := \ \{i \in I^\nu |\ g_i^\nu(x) = 0\}. \tag{8}$$

Note that W_ν coincides with $\operatorname{gph} X_\nu$, the graph of the set-valued mapping X_ν, and that we obviously have $W = \bigcap_{\nu=1}^N W_\nu$.

Recall that for $\nu \in \{1, \ldots, N\}$ the Mangasarian Fromovitz constraint qualification (MFCQ) holds at $x \in W_\nu$ if there exists some vector $d \in \mathbb{R}^n$ (typically depending on the index ν) with

$$Dg_i^\nu(x)\, d \ < \ 0, \quad i \in I_0^\nu(x).$$

Assumption 3.7 (Joint MFCQ) *For each $\nu \in \{1, \ldots, N\}$ the MFCQ holds everywhere in W_ν.*

Note that, under Assumption 3.5, Assumption 3.7 is equivalent to the Slater condition for each W_ν, $\nu = 1, \ldots, N$.

Remark 3.8 We stress that Assumption 3.7 is unrelated to the assumption of MFCQ everywhere in the set W. In fact, on the one hand, assuming MFCQ at points in W does not allow conclusions about points in $W_\nu \setminus W$ for any ν. In particular, MFCQ may be violated at some $\bar{x} \in W_\nu \setminus W$, so that Assumption 3.7 does not hold. On the other hand, consider a two player game with $n_1 = n_2 = 1$ as well as $g^1(x) = (x_1 - 1)^2 + x_2^2 - 1$ and $g^2(x) = (x_1 + 1)^2 + x_2^2 - 1$. Then Assumption 3.7 holds, but MFCQ is violated in the set $W = \{0\}$. $\Diamond$

If, for player ν and a point $x \in W_\nu$, MFCQ holds at $x^\nu \in X_\nu(x^{-\nu})$, that is, there exists some vector $d^\nu \in \mathbb{R}^{n_\nu}$ with

$$D_{x^\nu} g_i^\nu(x^\nu, x^{-\nu})\, d^\nu \; < \; 0, \quad i \in I_0^\nu(x^\nu, x^{-\nu}),$$

we will refer to this as *player MFCQ* in the sequel. Note that, for given ν and $x^{-\nu}$, the active index set $I_0^\nu(x^\nu, x^{-\nu})$ of x^ν in $X_\nu(x^{-\nu})$ coincides with the active index set of $(x^\nu, x^{-\nu})$ in W_ν as defined in (8).

Theorem 3.9 *Let Assumptions 1.1, 3.5, 3.6, and 3.7 hold. Then we have*

$$W \setminus D_1 \; = \; W \cap \operatorname{int} \operatorname{dom} V.$$

Proof. In view of Lemma 3.4 and (5), the assertion is shown if we can prove the relation

$$W \cap \operatorname{int} \operatorname{dom} \Omega \; \subset \; W \setminus D_1.$$

Let $\bar{x} \in D_1$. We will show that then $\bar{x}$ lies in the set complement $(\operatorname{int} \operatorname{dom} \Omega)^c$. Due to $D_1 \subset W \subset \operatorname{dom} \Omega$ it suffices to guarantee that any neighborhood of $\bar{x}$ contains points from $(\operatorname{dom} \Omega)^c$.

Choose some $\nu \in \{1, \ldots, N\}$ such that $X_\nu(\bar{x}^{-\nu})$ violates the Slater condition. Then player MFCQ is violated at any element of $X_\nu(\bar{x}^{-\nu})$ and, in particular, at $\bar{x}^\nu$. By the Lemma of Gordan, there exist multipliers $\lambda_i \geq 0$, $i \in I_0^\nu(\bar{x})$, with $\sum_{i \in I_0^\nu(\bar{x})} \lambda_i = 1$ such that

$$d^\nu \; := \sum_{i \in I_0^\nu(\bar{x})} \lambda_i \nabla_{x^\nu} g_i^\nu(\bar{x}) \; = \; 0. \tag{9}$$

We use the same multipliers to define

$$d^{-\nu} \; := \sum_{i \in I_0^\nu(\bar{x})} \lambda_i \nabla_{x^{-\nu}} g_i^\nu(\bar{x})$$

as well as $d := (d^\nu, d^{-\nu})$. We claim that d is nonzero. In fact, if we had $d = 0$, we would obtain $\sum_{i \in I_0^\nu(\bar{x})} \lambda_i \nabla g_i^\nu(\bar{x}) = 0$, hence, noting that $\bar{x} \in W_\nu$, it would follow from Assumption 3.7 and the fact that MFCQ is equivalent to the positive linear independence of the corresponding vectors that $\lambda_i = 0$ for all $i \in I_0^\nu(\bar{x})$, a contradiction to $\sum_{i \in I_0^\nu(\bar{x})} \lambda_i = 1$. Consequently, $d \neq 0$ and, in view of (9), we then also know that $d^{-\nu}$ cannot vanish.

We now define the ray

$$x^{-\nu}(t) \; := \; \bar{x}^{-\nu} + t d^{-\nu}$$

and shall show that for all $t > 0$ the set $X_\nu(x^{-\nu}(t))$ is empty.

To this end, we note that Assumption 3.5 implies that, for all $x \in W_\nu$, we have

$$0 \;\geq\; Dg_i^\nu(\bar{x})(x - \bar{x}),\;\; i \in I_0^\nu(\bar{x}).$$

Taking the convex combination of the latter inequalities with the above coefficients λ_i, and using $d^\nu = 0$, implies that all $x \in W_\nu$ also satisfy

$$0 \;\geq\; d^\intercal(x - \bar{x}) \;=\; (d^{-\nu})^\intercal(x^{-\nu} - \bar{x}^{-\nu}). \tag{10}$$

Consequently for each $x^{-\nu} \in \operatorname{dom} X_\nu$ there exists some $x^\nu \in \mathbb{R}^{n_\nu}$ with $x \in \operatorname{gph} X_\nu = W_\nu$, hence (10) holds. As (10) does not depend on x^ν, this means

$$\operatorname{dom} X_\nu \;\subset\; \{x^{-\nu} \in \mathbb{R}^{n-n_\nu} |\; (d^{-\nu})^\intercal(x^{-\nu} - \bar{x}^{-\nu}) \leq 0\}.$$

On the other hand, for any $t > 0$ the point $x^{-\nu}(t)$ satisfies

$$(d^{-\nu})^\intercal(x^{-\nu}(t) - \bar{x}^{-\nu}) \;=\; t\|d^{-\nu}\|_2^2 \;>\; 0,$$

so that $x^{-\nu}(t) \notin \operatorname{dom} X_\nu$, that is, $X_\nu(x^{-\nu}(t)) = \emptyset$. This shows the assertion. □

Note that in Example 3.2 the assumptions of Theorem 3.9 are satisfied.

Remark 3.10 Under Assumption 3.5 and the additional assumption of (affine) linear functions θ_ν, $\nu = 1, \ldots, N$, along the lines of the proof of [22, Prop. 3.1.26] one can show that the functions φ_ν, $\nu = 1, \ldots, N$, are convex on W, and that V is concave on W. Example 3.2 illustrates this situation. ◊

4 Differentiability Properties

In this section we will study differentiability properties of V on $W \setminus D_1$, as motivated by Lemma 3.4 and Theorem 3.9. Assumptions 1.1 and 3.6 will be blanket assumptions for this section.

For the following lemma, recall that $S_\nu(x)$ denotes the set of optimal points of $Q_\nu(x)$, let

$$L_\nu(x, y^\nu, \gamma^\nu) \;:=\; \theta_\nu(y^\nu, x^{-\nu}) + \frac{\alpha}{2}\|x^\nu - y^\nu\|_2^2 + (\gamma^\nu)^\intercal g^\nu(y^\nu, x^{-\nu}) \tag{11}$$

denote the Lagrange function of $Q_\nu(x)$, and let

$$KKT_\nu(x) \;:=\; \{\gamma^\nu \in \mathbb{R}^{m_\nu} |\; \nabla_{y^\nu} L_\nu(x, y^\nu, \gamma^\nu) = 0,\; \gamma^\nu \geq 0,\; (\gamma^\nu)^\intercal g^\nu(y^\nu, x^{-\nu}) = 0\}$$

be the set of Karush-Kuhn-Tucker multipliers for $y^\nu \in S_\nu(x)$. Note that $KKT_\nu(x)$ does not depend on y^ν as $Q_\nu(x)$ is a convex problem ([11]). Before stating the next result, we recall that a given mapping f is called *directionally differentiable* at a point x if the limit

$$\lim_{t \searrow 0} \frac{f(x + td) - f(x)}{t}$$

exists for all directions d, whereas f is called *directionally differentiable in the Hadamard sense* or simply *Hadamard directionally differentiable* at x if the limit

$$\lim_{t\searrow 0, d'\to d} \frac{f(x+td')-f(x)}{t}$$

exists for all directions d. Note that Hadamard directional differentiability implies the usual directional differentiability, and that we denote the common limit by $f'(x,d)$.

Theorem 4.1 *Let $x \in W \setminus D_1$. Then V is Hadamard directionally differentiable at x with*

$$V'(x,d) \;=\; \sum_{\nu=1}^{N} \left[D\theta_\nu(x)\, d - \max_{\gamma^\nu \in KKT_\nu(x)} D_x L_\nu(x, y^\nu(x), \gamma^\nu)\, d \right]$$

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

Proof. In view of (3), we have

$$V(x) = \sum_{\nu=1}^{N} \left[\theta_\nu(x) - \varphi_\nu(x) \right]$$

with θ_ν being differentiable and the possibly nondifferentiable optimal value functions φ_ν from (4). Since $x \in W \setminus D_1$, the sets $X_\nu(x^{-\nu})$ satisfy the Slater condition for all $\nu = 1,\ldots,N$, hence, by a standard result (see, e.g., [11, 14, 21]), the functions φ_ν are Hadamard directionally differentiable, and their directional derivatives are given by

$$\varphi_\nu'(x,d) \;=\; \min_{y^\nu \in S_\nu(x)} \; \max_{\gamma^\nu \in KKT_\nu(x)} D_x L_\nu(x, y^\nu, \gamma^\nu)\, d \tag{12}$$

for all $d \in \mathbb{R}^n$. Taking into account that $S_\nu(x) = \{y^\nu(x)\}$ is actually a singleton in our case, the desired statement follows. $\square$

Obviously, the formula for the directional derivative of φ_ν from (12) simplifies further if not only $S_\nu(x)$, but also $KKT_\nu(x)$ is a singleton. For the following, recall that a function is called Gâteaux differentiable if it is directionally differentiable and if the directional derivative is a linear function of the direction.

Proposition 4.2 *For some $\nu \in \{1,\ldots,N\}$ and $x \in W_\nu$, let $X_\nu(x^{-\nu})$ satisfy the Slater condition, and let $KKT_\nu(x)$ be the singleton $\{\gamma^\nu(x)\}$. Then φ_ν is Gâteaux differentiable at x with*

$$\varphi_\nu'(x,d) \;=\; D_x L_\nu(x, y^\nu(x), \gamma^\nu(x))\, d$$

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

The previous result motivates to define a second 'degenerate point set',

$$D_2 \;=\; \{x \in W |\ \text{for some } \nu = 1,\ldots,N \text{ the set } KKT_\nu(x) \text{ is not a singleton}\}.$$

A *sufficient* condition for $x \in W$ to lie in D_2^c is that for all $\nu = 1, \ldots, N$ the linear independence constraint qualification holds at $y^\nu(x)$ in $X_\nu(x^{-\nu})$, that is, the gradients

$$\nabla_{x^\nu} g_i^\nu(y^\nu(x), x^{-\nu}),\ i \in I_0^\nu(y^\nu(x), x^{-\nu}),$$

are linearly independent. We will refer to this property as *player LICQ*. In fact, it is well known that player LICQ entails a unique KKT multiplier $\gamma^\nu(x)$ at the optimal point $y^\nu(x)$.

By a result from [16], at points $x \in W$ a *characterization* for $x \in D_2^c$ is given by the fact that for all $\nu = 1, \ldots, N$ the strict Mangasarian Fromovitz constraint qualification holds at $y^\nu(x)$ in $X_\nu(x^{-\nu})$ with a multiplier $\gamma^\nu \in KKT_\nu(x)$, that is, the gradients

$$\nabla_{x^\nu} g_i^\nu(y^\nu(x), x^{-\nu}),\ i \in I_{0+}^\nu(y^\nu(x), x^{-\nu}),$$

are linearly independent, and there exists some vector $d^\nu \in \mathbb{R}^{n_\nu}$ with

$$\begin{aligned} D_{x^\nu} g_i^\nu(y^\nu(x), x^{-\nu})\, d^\nu &= 0, \quad i \in I_{0+}^\nu(y^\nu(x), x^{-\nu}), \\ D_{x^\nu} g_i^\nu(y^\nu(x), x^{-\nu})\, d^\nu &< 0, \quad i \in I_{00}^\nu(y^\nu(x), x^{-\nu}), \end{aligned}$$

where

$$\begin{aligned} I_{0+}^\nu(x) &= \{i \in I_0^\nu(x) |\ \gamma_i^\nu > 0\}, \\ I_{00}^\nu(x) &= \{i \in I_0^\nu(x) |\ \gamma_i^\nu = 0\}. \end{aligned}$$

In the sequel, this property will be called *player SMFCQ*. This yields

$$\begin{aligned} D_2 = \{x \in W |\ &\text{for some } \nu = 1, \ldots, N \text{ player SMFCQ is violated} \\ &\text{at } y^\nu(x) \text{ in } X_\nu(x^{-\nu})\,\} \end{aligned} \tag{13}$$

and allows us to prove the following relation between D_1 and D_2.

Lemma 4.3 *The degenerate point sets satisfy $D_1 \subset D_2$.*

Proof. Choose any point $x \in D_2^c$. The assertion is trivial if $x \in W^c$. Otherwise, in view of (13) for all $\nu = 1, \ldots, N$ player SMFCQ holds at $y^\nu(x)$ in $X_\nu(x^{-\nu})$. Since player SMFCQ implies the ordinary player MFCQ at the latter point, $X_\nu(x^{-\nu})$ also satisfies the Slater condition for all $\nu = 1, \ldots, N$. This means that x lies in D_1^c and shows the assertion. □

In view of Lemma 4.3, in Proposition 4.2 the assumption that $X_\nu(x^{-\nu})$ satisfies the Slater condition may be dropped, and we arrive at the following theorem.

Theorem 4.4 *For $x \in W \setminus D_2$ let $\gamma^\nu(x)$, $\nu = 1, \ldots, N$, denote the unique KKT multipliers. Then V is Gâteaux differentiable at x with*

$$V'(x, d) = \left(\sum_{\nu=1}^N \left[D\theta_\nu(x) - D_x L_\nu(x, y^\nu(x), \gamma^\nu(x)) \right] \right) d$$

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

Clearly player LICQ, or even player SMFCQ, cannot be expected to hold at $y^\nu(x)$ in $X_\nu(x^{-\nu})$ for all $x \in W$, even if the set W itself satisfies LICQ everywhere, that is, the (full) gradients

$$\nabla g_i^\nu(x),\ i \in I_0^\nu(x),\ \nu = 1, \dots, N$$

are linearly independent at each $x \in W$. To begin with, the point $(y^\nu(x), x^{-\nu}) \in W_\nu$ does not even have to belong to W. However, violation of player SMFCQ is in some sense exceptional, as the following example illustrates.

Example 4.5 In the situation of Example 3.2, recall that for $x \in D_1 = \{(1,0)\}$ the set $X_2(x_1)$ violates the Slater condition and, thus, player SMFCQ is violated at its single element. In the following we check for points in $D_2 \setminus D_1$.

The Lagrangian of player 1 is

$$L_1(x, y_1, \gamma^1) \;=\; y_1 + \gamma_1^1(-2y_1 + x_2).$$

The optimal point $y_1(x) = \frac{x_2}{2}$ has the active index set $I_0^1(y_1(x), x_2) = \{1\}$, and we obtain the multiplier set

$$KKT_1(x) \;=\; \{\gamma^1 \in \mathbb{R} |\ 1 - 2\gamma_1^1 = 0,\ \gamma_1^1 \geq 0\} \;=\; \{1/2\}.$$

In particular, the multiplier $\gamma^1(x) = \frac{1}{2}$ is unique for any $x \in W$.

For player 2 the Lagrangian is

$$L_2(x, y_2, \gamma^2) \;=\; y_2 + \gamma_1^2(x_1^2 + y_2^2 - 1) + \gamma_2^2(-x_1 - y_2).$$

For $x \in W \setminus D_1$ with $x_1 > 1/\sqrt{2}$ the optimal point is $y_2(x) = -\sqrt{1 - x_1^2}$ with active index set $I_0^2(x_1, y_2(x)) = \{1\}$, and we obtain the multiplier set

$$\begin{aligned} KKT_2(x) \;&=\; \{\gamma^2 \in \mathbb{R}^2 |\ 1 + 2\gamma_1^2 y_2(x) - \gamma_2^2 = 0,\ \gamma_1^2 \geq 0,\ \gamma_2^2 = 0\} \\ &=\; \left\{ \left(\left(2\sqrt{1 - x_1^2} \right)^{-1}, 0 \right) \right\}. \end{aligned}$$

For $x \in W$ with $x_1 < 1/\sqrt{2}$ the optimal point is $y_2(x) = -x_1$ with active index set $I_0^2(x_1, y_2(x)) = \{2\}$, and we obtain the multiplier set

$$KKT_2(x) \;=\; \{\gamma^2 \in \mathbb{R}^2 |\ 1 + 2\gamma_1^2 y_2(x) - \gamma_2^2 = 0,\ \gamma_1^2 = 0,\ \gamma_2^2 \geq 0\} \;=\; \{(0,1)\}.$$

Altogether, for all $x \in W \setminus D_1$ with $x_1 \neq 1/\sqrt{2}$ the multiplier $\gamma^2(x)$ is unique.

On the other hand, for $x \in W$ with $x_1 = 1/\sqrt{2}$, the active index set of $y_2(x) = -1/\sqrt{2}$ is $I_0^2(x_1, y_2(x)) = \{1, 2\}$, and we obtain the nonunique multiplier set

$$\begin{aligned} KKT_2(x) \;&=\; \{\gamma^2 \in \mathbb{R}^2 |\ 1 + 2\gamma_1^2 y_2(x) - \gamma_2^2 = 0,\ \gamma^2 \geq 0\} \\ &=\; \left\{ \gamma^2 \geq 0 \middle|\ \gamma_2^2 = 1 - \sqrt{2}\gamma_1^2 \right\} \end{aligned}$$

$$= \left\{ \left(t, 1-\sqrt{2}t\right) \,\middle|\, t \in \left[0, 1/\sqrt{2}\right] \right\}.$$

Thus, we arrive at

$$D_2 \;=\; D_1 \cup \{x \in W |\; x_1 = 1/\sqrt{2}\}$$

and, in view of Theorem 4.4, V is Gâteaux differentiable on $W \setminus D_2$.

Finally we check for differentiability properties of V on $D_2 \setminus D_1$. Recall that we cannot study differentiability of V on $D_1 = \{(1,0)\}$, as the point $(1,0)$ is not an interior point of $\operatorname{dom} V$. In fact, as mentioned already in Example 3.2, V actually has 'infinite slope' at $(1,0)$.

For all $x \in D_2 \setminus D_1$, that is, $x \in W$ with $x_1 = 1/\sqrt{2}$, Theorem 4.1 yields directional differentiability of V with

$$\begin{aligned}
V'(x,d) \;&=\; D\theta_1(x)\,d + D\theta_2(x)\,d - D_xL_1(x, y^1(x), \gamma^1(x))\,d \\
&\quad - \max_{\gamma^2 \in KKT_2(x)} D_xL_2(x, y^2(x), \gamma^2(x))\,d \\
&=\; d_1 + d_2 - \frac{1}{2}d_2 - \max_{t \in \left[0, 1/\sqrt{2}\right]} (2t/\sqrt{2} - (1-\sqrt{2}t))\,d_1 \\
&=\; d_1 + \frac{1}{2}d_2 + \min_{t \in \left[0, 1/\sqrt{2}\right]} (1 - 2\sqrt{2}t)\,d_1 \\
&=\; \begin{cases} 2d_1 + \frac{1}{2}d_2, & d_1 < 0 \\ \frac{1}{2}d_2, & d_1 \geq 0 \end{cases}
\end{aligned}$$

for all $d \in \mathbb{R}^2$. This corresponds to the 'concave kink' in the graph of V which is illustrated in Figure 2. $\Diamond$

The observed differentiability properties in Example 4.5 guarantee that any local minimizer $\bar{x}$ of V on W either lies in D_1, or the function V is Gâteaux differentiable at $\bar{x}$.

In the sequel we will show that, under mild assumptions, this also holds in the general case. The essential implications for the design of numerical methods to solve P in (2) are apparent (cf. Sec. 5). We begin with a preliminary result which gives a representation for the gradient of Lagrangian L_ν from (11).

Lemma 4.6 *Let L_ν be the Lagrangian of $Q_\nu(x^{-\nu})$. Then the gradient with respect to all variables $x = (x^1, \ldots, x^N)$, evaluated at a point $(\bar{x}, \bar{y}^\nu, \gamma^\nu)$ with $\bar{x} \in W$, $\bar{y}^\nu := y^\nu(\bar{x})$, has the representation*

$$\nabla_x L_\nu(\bar{x}, \bar{y}^\nu, \gamma^\nu) \;=\; \nabla\theta_\nu(\bar{y}^\nu, \bar{x}^{-\nu}) + \sum_{i \in I_0^\nu(\bar{y}^\nu, \bar{x}^{-\nu})} \gamma_i^\nu \nabla g_i^\nu(\bar{y}^\nu, \bar{x}^{-\nu}).$$

Proof. The definitions of $L_\nu(x, y^\nu, \gamma)$ and $KKT_\nu(x)$ immediately imply that, for all $\mu \in \{1, \ldots, N\} \setminus \{\nu\}$, we have

$$\nabla_{x^\mu} L_\nu(\bar{x}, \bar{y}^\nu, \gamma^\nu) \;=\; \nabla_{x^\mu}\theta_\nu(\bar{y}^\nu, \bar{x}^{-\nu}) + \sum_{i \in I_0^\nu(\bar{y}^\nu, \bar{x}^{-\nu})} \gamma_i^\nu \nabla_{x^\mu} g^\nu(\bar{y}^\nu, \bar{x}^{-\nu}). \tag{14}$$

Moreover, the combination of

$$\nabla_{x^\nu} L_\nu(\bar{x}, \bar{y}^\nu, \gamma^\nu) \;=\; \alpha(\bar{x}^\nu - \bar{y}^\nu)$$

and, by the definition of $KKT_\nu(\bar{x})$,

$$\begin{aligned} 0 \;&=\; \nabla_{y^\nu} L_\nu(\bar{x}, \bar{y}^\nu, \gamma^\nu) \\ &=\; \nabla_{x^\nu} \theta_\nu(\bar{y}^\nu, \bar{x}^{-\nu}) - \alpha(\bar{x}^\nu - \bar{y}^\nu) + \sum_{i \in I_0^\nu(\bar{y}^\nu, \bar{x}^{-\nu})} \gamma_i^\nu \nabla_{x^\nu} g^\nu(\bar{y}^\nu, \bar{x}^{-\nu}) \end{aligned}$$

shows that (14) also holds for $\mu = \nu$ (independent of α). □

The following result will be used in order to show that at certain points there exist feasible descent directions for V on W. To this end, we exclude some degenerate points from D_2 for proving the result of Theorem 4.10. In fact, we define D_3 to be the set of points in D_2 such that whenever player SMFCQ is violated at $y^\mu(x)$ in $X_\mu(x^{-\mu})$ for some $\mu = 1, \ldots, N$, we have

$$\begin{aligned} &\operatorname{span}\left\{\nabla g_i^\nu(x),\ i \in I_0^\nu(x),\ \nu = 1, \ldots, N\right\} \,\cap \\ &\operatorname{span}\left\{\nabla g_i^\mu(y^\mu(x), x^{-\mu}),\ i \in I_0^\mu(y^\mu(x), x^{-\mu})\right\} \;\neq\; \{0\}, \end{aligned} \tag{15}$$

that is,

$$\begin{aligned} D_3 \;:=\; \big\{&x \in D_2 \mid (15) \text{ holds for all } \mu = 1, \ldots, N \\ &\text{where SMFCQ is violated at } y^\mu(x) \in X_\mu(x^{-\mu})\big\}. \end{aligned}$$

Note that at least the *unconstrained* points $x \in D_2$, that is, the ones with $I_0^\nu(x) = \emptyset$, $\nu = 1, \ldots, N$, do not lie in D_3, as $\operatorname{span} \emptyset = \{0\}$. More generally, in the case $\sum_{\nu=1}^N |I_0^\nu(x)| + |I_0^\mu(y^\mu(x), x^{-\mu})| \leq n$ one may expect that x does not lie in D_3, as the involved gradients are evaluated at different arguments. On the other hand, for $\sum_{\nu=1}^N |I_0^\nu(x)| + |I_0^\mu(y^\mu(x), x^{-\mu})| > n$ and linearly independent gradients, x will definitely lie in D_3.

Also, if a generalized Nash equilibrium x happens to lie in D_2, under mild conditions it necessarily is an element of D_3. In fact, due to $x \in D_2$, for some ν the SMFCQ is violated at $(y^\nu(x), x^{-\nu})$ in $X_\nu(x^{-\nu})$ so that, in particular, $I_0^\nu(y^\nu(x), x^{-\nu})$ is nonempty. Moreover, by Proposition 2.1(d), in a generalized Nash equilibrium the points $y^\nu(x)$ and x^ν coincide. This means that in (15) the vectors $\nabla g_i^\nu(x),\ i \in I_0^\nu(x)$, appear in both spans, where $I_0^\nu(x)$ is nonempty. Thus, except for the case where all these vectors vanish, the intersection of the two spans is strictly larger than $\{0\}$.

With regard to linear independence, we will actually need the following assumption, which strengthens Assumption 3.7.

Assumption 4.7 (Joint LICQ) *For each $\nu = 1, \ldots, N$, LICQ holds everywhere in W_ν.*

Proposition 4.8 *Let $\bar{x} \in D_2 \setminus (D_1 \cup D_3)$ and Assumption 4.7 hold. Then there exists a vector $d \in \mathbb{R}^n$ solving the system*

$$V'(\bar{x}, d) < 0, \quad Dg_i^\nu(\bar{x})\, d \leq 0,\ i \in I_0^\nu(\bar{x}),\ \nu = 1, \ldots, N. \tag{16}$$

Proof. Assume that (16) does not possess a solution $d \in \mathbb{R}^n$. By Theorem 4.1, the directional derivative of V is

$$V'(\bar{x}, d) \;=\; \sum_{\nu=1}^{N} \left[D\theta_\nu(\bar{x})\, d - \max_{\gamma^\nu \in KKT_\nu(\bar{x})} D_x L_\nu(\bar{x}, \bar{y}^\nu, \gamma^\nu)\, d \right]$$

for all $d \in \mathbb{R}^n$, where we put $\bar{y}^\nu = y^\nu(\bar{x})$ for $\nu = 1, \ldots, N$. Hence, the inconsistency of (16) implies that for any choice

$$\gamma := (\gamma^1, \ldots, \gamma^N) \in KKT_1(\bar{x}) \times \ldots \times KKT_N(\bar{x}) \tag{17}$$

also the system

$$\begin{aligned} \left(\sum_{\nu=1}^{N} \left[D\theta_\nu(\bar{x}) - D_x L_\nu(\bar{x}, \bar{y}^\nu, \gamma^\nu) \right] \right) d \;&<\; 0, \\ Dg_i^\nu(\bar{x})\, d \;&\leq\; 0, \; i \in I_0^\nu(\bar{x}), \; \nu = 1, \ldots, N \end{aligned}$$

is inconsistent. By the Lemma of Farkas, the latter holds if and only if there exist scalars $\lambda_i^\nu(\gamma) \geq 0$, $i \in I_0^\nu(\bar{x})$, $\nu = 1, \ldots, N$, with

$$\sum_{\nu=1}^{N} \left[\nabla\theta_\nu(\bar{x}) - \nabla_x L_\nu(\bar{x}, \bar{y}^\nu, \gamma^\nu) \right] + \sum_{\nu=1}^{N} \sum_{i \in I_0^\nu(\bar{x})} \lambda_i^\nu(\gamma) \nabla g_i^\nu(\bar{x}) \;=\; 0.$$

After rearranging terms, we find that for any choice γ with (17) there exist multipliers $\lambda_i^\nu(\gamma) \geq 0$ with

$$\sum_{\nu=1}^{N} \Big(\nabla\theta_\nu(\bar{x}) + \sum_{i \in I_0^\nu(\bar{x})} \lambda_i^\nu(\gamma) \nabla g_i^\nu(\bar{x}) \Big) \;=\; \sum_{\nu=1}^{N} \nabla_x L_\nu(\bar{x}, \bar{y}^\nu, \gamma^\nu). \tag{18}$$

Using Lemma 4.6 to replace the expression for the gradient on the right-hand side, we conclude that for any choice γ with (17) there exist multipliers $\lambda_i^\nu(\gamma) \geq 0$ with

$$\begin{aligned} &\sum_{\nu=1}^{N} \Big(\nabla\theta_\nu(\bar{x}) + \sum_{i \in I_0^\nu(\bar{x})} \lambda_i^\nu(\gamma) \nabla g_i^\nu(\bar{x}) \Big) \\ &\quad = \; \sum_{\nu=1}^{N} \Big(\nabla\theta_\nu(\bar{y}^\nu, \bar{x}^{-\nu}) + \sum_{i \in I_0^\nu(\bar{y}^\nu, \bar{x}^{-\nu})} \gamma_i^\nu \nabla g_i^\nu(\bar{y}^\nu, \bar{x}^{-\nu}) \Big). \end{aligned} \tag{19}$$

Next we use that $\bar{x}$ was chosen from D_2 so that at least for one $\mu \in \{1, \ldots, N\}$ the player SMFCQ is violated at $\bar{y}^\mu$, say for $\mu = 1$. Then $KKT_1(\bar{x})$ contains two different multipliers $\hat{\gamma}^1 \neq \tilde{\gamma}^1$. For $\nu = 2, \ldots, N$ we choose any $\gamma^\nu \in KKT_\nu(\bar{x})$ and put $\hat{\gamma} := (\hat{\gamma}^1, \gamma^2, \ldots, \gamma^N)$ as well as $\tilde{\gamma} := (\tilde{\gamma}^1, \gamma^2, \dots, \gamma^N)$. Equation (19) then holds with $\gamma = \hat{\gamma}$ as well as with $\gamma = \tilde{\gamma}$. Subtracting these two equations leads to

$$\sum_{\nu=1}^{N} \sum_{i \in I_0^\nu(\bar{x})} \big(\lambda_i^\nu(\hat{\gamma}) - \lambda_i^\nu(\tilde{\gamma})\big) \nabla g_i^\nu(\bar{x}) \;=\; \sum_{i \in I_0^1(\bar{y}^1, \bar{x}^{-1})} (\hat{\gamma}_i^1 - \tilde{\gamma}_i^1) \nabla g_i^1(\bar{y}^1, \bar{x}^{-1}),$$

where the left hand side is some element of

$$\operatorname{span} \{\nabla g_i^\nu(\bar{x}),\ i \in I_0^\nu(\bar{x}),\ \nu = 1, \dots, N\},$$

and the right hand side is some element of

$$\operatorname{span} \left\{\nabla g_i^1(\bar{y}^1, \bar{x}^{-1}),\ i \in I_0^1(\bar{y}^1, \bar{x}^{-1})\right\},$$

which, in addition, cannot vanish due to $\hat{\gamma}^1 \neq \tilde{\gamma}^1$, $(\bar{y}^1, \bar{x}^{-1}) \in W_1$ and Assumption 4.7. Therefore, the intersection of these two spans is nontrivial. However, since $\bar{x}$ was taken from $D_2 \setminus D_3$, this is a contradiction. Consequently, contrary to our assumption, (16) must be consistent. This shows the assertion. □

Before we state the main result of this section, we need one more assumption. To this end, we first recall that the tangent (or contingent or Bouligand) cone to W at $\bar{x}$ is defined by

$$\mathcal{T}_W(\bar{x}) \;:=\; \{d \in \mathbb{R}^n |\ \exists t_k \searrow 0,\ d^k \to d :\ \bar{x} + t_k d^k \in W \text{ for all } k \in \mathbb{N}\},$$

and that the linearization cone to W at $\bar{x}$ is given by

$$\mathcal{L}_W(\bar{x}) \;:=\; \{d \in \mathbb{R}^n |\ Dg_i^\nu(\bar{x})\, d \leq 0,\ i \in I_0^\nu(\bar{x}),\ \nu = 1, \dots, N\}.$$

The inclusion $\mathcal{T}_W(\bar{x}) \subset \mathcal{L}_W(\bar{x})$ always holds (cf., e.g., [23]). The *Abadie constraint qualification* (ACQ) is said to hold at $\bar{x} \in W$, if both cones actually coincide.

Assumption 4.9 (Joint ACQ in W) *The ACQ holds everywhere in W.*

The ACQ is typically considered a very weak constraint qualification. Nevertheless, we point out that the example from Remark 3.8 shows that neither Assumption 3.7 nor Assumption 4.7 imply Assumption 4.9.

The following is the main result this section.

Theorem 4.10 *Let Assumptions 1.1, 3.6, 4.7, and 4.9 hold. Then any local minimizer $\bar{x}$ of V on W either lies in $D_1 \cup D_3$, or the function V is Gâteaux differentiable at $\bar{x}$.*

Proof. By Theorem 4.4, V is Gâteaux differentiable everywhere on $W \setminus D_2$. Choose any $\bar{x} \in D_2 \setminus (D_1 \cup D_3)$. The assertion is shown if we can prove that $\bar{x}$ is not a local minimizer of V on W. The main idea of the proof is to show this by guaranteeing the existence of a first order feasible descent direction for V on W in $\bar{x}$.

In view of Proposition 4.8, there exists a vector $d \in \mathbb{R}^n$ solving the system (16). In particular, d belongs to the linearization cone $\mathcal{L}_W(\bar{x})$. In view of Assumption 4.9, d also lies in the tangent cone $\mathcal{T}_W(\bar{x})$, that is, there exist sequences $t_k \searrow 0$ and $d^k \to d$ with $\bar{x} + t_k d^k \in W$ for all $k \in \mathbb{N}$.

Assume that $\bar{x}$ is a local minimizer of V on W. This implies $V(\bar{x} + t_k d^k) \geq V(\bar{x})$ and, thus,

$$\frac{V(\bar{x} + t_k d^k) - V(\bar{x})}{t_k} \; \geq \; 0 \tag{20}$$

for all sufficiently large $k \in \mathbb{N}$. By Theorem 4.1, V is Hadamard directionally differentiable at $\bar{x}$ so that the limit of the left-hand side in (20) exists and equals $V'(\bar{x}, d)$ (note that, here, we exploit the fact that V is actually Hadamard directionally differentiable and not just directionally differentiable in the ordinary sense). However, since the implication $V'(\bar{x}, d) \geq 0$ contradicts (16), $\bar{x}$ cannot be a local minimizer of V on W. □

It is well-known (cf., e.g., [23, Prop. 3.2]) that ACQ holds everywhere in W in the case where all constraints g_i^ν are linear. This immediately leads to the following result.

Corollary 4.11 *Let Assumptions 1.1, 3.6, 4.7 hold, and assume that all constraint functions g_i^ν are linear. Then any local minimizer $\bar{x}$ of V on W either lies in $D_1 \cup D_3$, or the function V is Gâteaux differentiable at $\bar{x}$.*

We remark that in Example 4.5 the two constrained elements of D_2 belong to D_3, but first order feasible descent directions for V on W still exist in these points. This indicates that it should be possible to weaken the assumptions of Theorem 4.10.

With respect to Remark 3.10 we note that, although V is not necessarily concave under the general assumptions of Theorem 4.10, coarsely speaking ‘a concavity property prevails in the kinks of V’.

For completeness, we emphasize that in Theorem 4.4 the Gâteaux differentiability of V can be replaced by Fréchet differentiability if the partial derivatives of V are continuous. The next corollary formulates this fact more explicitly.

Corollary 4.12 *For $\bar{x} \in W \setminus D_2$ and each $\nu = 1, \ldots, N$, let the function y^ν be continuous at $\bar{x}$, let the unique multiplier γ^ν be continuous at $\bar{x}$, and put $\bar{y}^\nu := y^\nu(\bar{x})$ as well as $\bar{\gamma}^\nu := \gamma^\nu(\bar{x})$. Then V is Fréchet differentiable at $\bar{x}$ with*

$$\nabla V(\bar{x}) \; = \; \sum_{\nu=1}^{N} \big[\nabla \theta_\nu(\bar{x}) - \nabla_x L_\nu(\bar{x}, \bar{y}^\nu, \bar{\gamma}^\nu)\big].$$

Under the assumptions of Corollary 4.12, also in Theorem 4.10 Gâteaux differentiability can be replaced by Fréchet differentiability.

Remark 4.13 Under the assumptions of Corollary 4.12 and using (13), a point $x \in W$ is not an element of D_2 if for all $\nu = 1, \ldots, N$ at $y^\nu(x)$ in $X_\nu(x^{-\nu})$ the gradients

$$\nabla_{x^\nu} g_i^\nu(y^\nu(x), x^{-\nu}), \; i \in I_{0+}^\nu(y^\nu(x), x^{-\nu}),$$

are linearly independent, and there exists some vector $d^\nu \in \mathbb{R}^{n_\nu}$ with

$$\begin{array}{rcll} D_{x^\nu} g_i^\nu(y^\nu(x), x^{-\nu})\, d^\nu &=& 0, & i \in I_{0+}^\nu(y^\nu(x), x^{-\nu}) \\ D_{x^\nu} g_i^\nu(y^\nu(x), x^{-\nu})\, d^\nu &<& 0, & i \in I_{00}^\nu(y^\nu(x), x^{-\nu}), \end{array}$$

where

$$\begin{array}{rcl} I_{0+}^\nu(x) &=& \{i \in I_0^\nu(x) \mid \gamma_i^\nu(x) > 0\}, \\ I_{00}^\nu(x) &=& \{i \in I_0^\nu(x) \mid \gamma_i^\nu(x) = 0\}. \end{array}$$

Hence, if additionally for each $x \in W \setminus D_2$ the set $I_{00}^\nu(x)$ remains constant under small perturbations of x (e.g., due to $I_{00}^\nu(x) = \emptyset$, i.e., strict complementary slackness), then continuity arguments show that SMFCQ is stable at $y^\nu(x)$ for all $\nu = 1, \ldots, N$ under sufficiently small perturbations of x. Consequently, then the set D_2 is closed and, under the additional assumptions of Theorem 4.10, V is not only Fréchet differentiable at each local minimizer of V on W, outside of $D_1 \cup D_3$, but also on a whole neighborhood of such a local minimizer. $\Diamond$

Remark 4.14 Recall that in the jointly convex case one assumes identical constraints for all players, $g^1 = g^2 = \ldots = g^N =: g$, and that the components of g are convex in the whole vector $x = (x^1, \ldots, x^N)$. The strategy spaces thus have the representation

$$X_\nu(x^{-\nu}) \;=\; \{x^\nu \in \mathbb{R}^{n_\nu} \mid g(x^\nu, x^{-\nu}) \le 0\}, \quad \nu = 1, \ldots, N,$$

so that $x \in \Omega(x)$ holds if and only if x lies in the set

$$\widetilde{W} \;:=\; \{x \in \mathbb{R}^n \mid g(x) \le 0\}. \tag{21}$$

Note that, in contrast to the player convex case, $\widetilde{W}$ is necessarily convex (and Assumption 3.5 automatically holds). An important observation is that the definition of $\widetilde{W}$ is slightly different from the definition of W in (1). In fact, while the *geometries* of both sets coincide, their *functional descriptions* are different as, formally, W is described by N identical inequalities $g(x) \le 0$ in the jointly convex case. With regard to constraint qualifications like LICQ, the latter description of W is necessarily degenerate at boundary points, while the description of $\widetilde{W}$ from (21) may be expected to enjoy nondegeneracy properties.

In particular, all sets W_ν, $\nu = 1, \ldots, N$, from (7) coincide with $\widetilde{W}$, and Assumption 3.7 coincides with the assumption of MFCQ everywhere in $\widetilde{W}$. Of course, the latter is equivalent to the Slater condition for $\widetilde{W}$ and implies the Abadie constraint qualification for $\widetilde{W}$. Moreover, Assumption 4.7 on LICQ at each point of each set W_ν, $\nu = 1, \ldots, N$, can be replaced by the assumption of LICQ at each point of $\widetilde{W}$ which, in turn, implies the Slater condition for $\widetilde{W}$. In any case, Assumption 4.9 is superfluous in Theorem 4.10 for jointly convex problems. Hence, in contrast to the player convex case, our assumptions on constraint qualifications are highly interrelated in the jointly convex case. We also remark that, in the jointly convex case, D_3 is defined to be the set of points in D_2 such that whenever player SMFCQ is violated at $y^\mu(x)$ in $X_\mu(x^{-\mu})$ for some $\mu = 1, \ldots, N$, we have

$$\operatorname{span}\{\nabla g_i(x),\ i \in I_0(x)\} \ \cap\ \operatorname{span}\left\{\nabla g_i(y^\mu(x), x^{-\mu}),\ i \in I_0(y^\mu(x), x^{-\mu})\right\} \ \neq\ \{0\}. \tag{22}$$

We emphasize that so-called shared constraints as in the jointly convex case lead to numerical difficulties in all established numerical methods for the solution of GNEPs. In fact, while repeating identical constraints can be expected to lead to degeneracies in any numerical approach, dropping redundant constraints basically leads to underdetermined systems and, thus, alternative numerical problems in all approaches which we are aware of. In contrast to this, dropping redundant constraints in the present approach by switching from W to $\widetilde{W}$ does not introduce numerical problems. $\Diamond$

The following example illustrates for the jointly convex case that a generalized Nash equilibrium can fall in any of the three categories mentioned in Theorem 4.10.

Example 4.15 We slightly modify Example 3.2 by setting $N = 2$, $n_1 = n_2 = 1$, $\theta_1(x) = -x_1$, $\theta_2(x) = x_2$, $g_1^1(x) = g_1^2(x) = -2x_1 + x_2$, $g_2^1(x) = g_2^2(x) = x_1^2 + x_2^2 - 1$, $g_3^1(x) = g_3^2(x) = -x_1 - x_2$. Then for all $x \in \widetilde{W} = \{x \in \mathbb{R}^2 |\ x_1^2 + x_2^2 \leq 1,\ -x_1 \leq x_2 \leq 2x_1\}$ (cf. Fig. 1) the problems $Q_1(x)$ and $Q_2(x)$ are easily seen to be uniquely solvable for $\alpha = 0$. Moreover, it is not hard to see that the assumptions of Theorem 4.10 with the modifications from Remark 4.14 are satisfied.

As to be expected for problems with shared constraints, the set of generalized Nash equilibria is not a singleton, but it is formed by the (closed) boundary arc of $\widetilde{W}$ connecting the points $\bar{x}^1 = 1/\sqrt{2}\,(1, -1)$ and $\bar{x}^2 = (1, 0)$. Both $\bar{x}^1$ and $\bar{x}^2$ are elements of D_1, and $\bar{x}^1$ also lies in D_3 since SMFCQ is violated at $y_2(\bar{x}^1) = -1/\sqrt{2}$ in $X_2(1/\sqrt{2})$ and the intersection of spans in (22) is $\mathbb{R}^2$.

On the other hand, by direct inspection or as shown in Theorem 4.10, the resulting function $V(x) = -x_1 + x_2 + \sqrt{1 - x_2^2} + \min\{x_1, \sqrt{1 - x_1^2}\}$ is (even Fréchet) differentiable at all generalized Nash equilibria except for $\bar{x}^1$ and $\bar{x}^2$. $\Diamond$

5 Numerical Results

In view of Proposition 2.1, we know that the computation of a generalized Nash equilibrium is equivalent to solving the constrained optimization problem P from (2). The objective function V of this optimization problem is, in general, nondifferentiable. However, our previous results indicate that, on the one hand, the set of nondifferentiable points is exceptional and, on the other hand (and more importantly), we may expect differentiability of V at any solution of the GNEP. Hence, we may view problem P essentially as a smooth optimization problem. The objective function V, however, may not be defined outside the feasible set W, hence any suitable algorithm applied to problem P should guarantee that all iterates stay feasible. We therefore decided to apply a feasible direction-type method to problem P.

The class of feasible direction methods was introduced by Zoutendijk [26]. A variant is due to Topkins and Veinott [24] which, in turn, is the basis of the method presented by Birge et al. in [2]. The latter method uses a convex quadratic program at each iteration and will be used in order to solve our problem P. For the sake of completeness, we restate this method here.

Algorithm 5.1 *(Feasible direction-type method from [2])*

(S.0) Choose $x^0 \in W$, $H_0 \in \mathbb{R}^{n \times n}$ symmetric positive definite, $\beta, \sigma \in (0,1), c^0_{\nu,i} > 0$ for all $i = 1, \dots, m_\nu, \nu = 1, \dots, N$, $c^0_V > 0$, and set $k := 0$.

(S.1) If a suitable termination criterion holds: STOP.

(S.2) Compute a solution $(d^k, \delta^k) \in \mathbb{R}^n \times \mathbb{R}$ of

$$\begin{array}{ll} \min & \delta + \frac{1}{2} d^T H_k d \\ s.t. & \nabla V(x^k)^T d \leq c^k_V \delta, \\ & g^\nu_i(x^k) + \nabla g^\nu_i(x^k)^T d \leq c^k_{\nu,i} \delta \quad \forall i = 1, \dots, m_\nu, \nu = 1, \dots, N. \end{array}$$

If $(d^k, \delta^k) = (0, 0)$: STOP. Otherwise go to (S.3).

(S.3) Compute a stepsize $t_k = \max\{\beta^l \mid l = 0, 1, 2, \dots\}$ such that the following conditions hold:

$$V(x^k + t_k d^k) \leq V(x^k) + \sigma t_k \nabla V(x^k)^T d^k$$

and

$$g^\nu_i(x^k + t_k d^k) \leq 0 \quad \forall i = 1, \dots, m_\nu, \nu = 1, \dots, N.$$

(S.4) Choose $c^{k+1}_{\nu,i} > 0$ $(i = 1, \dots, m_\nu, \nu = 1, \dots, N), c^{k+1}_V > 0, H_{k+1} \in \mathbb{R}^{n \times n}$ symmetric positive definite, set $x^{k+1} := x^k + t_k d^k, k \leftarrow k + 1$, and go to (S.1).

The main termination criterion used in (S.1) is

$$V(x^k) \leq N \cdot \varepsilon \quad \text{with} \quad \varepsilon := 10^{-5}.$$

The factor N in front of ε comes from the fact that V is the sum of N terms, see (3), and the basic idea is that each term should be less than ε, hence our termination criterion is, in some way, independent of the number of players. Moreover, the parameter ε should not be taken too small since the feasible directon method used here is not a locally fast convergent method.

The computation of the matrix H_k was done in the following way: We begin with $H_0 := I$ and compute H_{k+1} as the BFGS-update of H_k whenever this gives a symmetric positive definite matrix, whereas we simply take $H_{k+1} := I$ otherwise. Furthermore, the parameters $c^k_{\nu,i}, c^k_V$ are chosen in the following way: We always use $c^k_{\nu,i} := 1$ for all i, ν and for all iterations k (including $k = 0$), whereas we take $c^0_V := 10$ in Step (S.0) and update

Ex.	N	n	x^0	k	V_{opt}
A1	10	10	0.01	6	6.4089e-05
			0.1	2	4.9819e-06
			1	2	4.9819e-06
A3	3	7	0	14	2.6397e-05
			1	14	2.9342e-06
			10	29	2.3998e-06
A4	3	7	10	62	0
A5	3	7	0	13	1.5470e-05
			1	29	2.5578e-05
			10	55	1.9534e-05
A6	3	7	0	92	2.1231e-05
			1	33	2.4189e-05
			10	89	2.7651e-05
A7	4	20	0	77	2.2178e-05
			1	65	2.4832e-05
			10	153	2.5712e-05
A8	3	3	0	15	1.1013e-05
			1	13	1.2011e-05
			10	13	1.2010e-05

Ex.	N	n	x^0	k	V_{opt}
A9a	7	56	0	119	5.4673e-05
A9b	7	112	0	395	4.1144e-05
A11	2	2	0	6	1.0419e-05
A12	2	2	(2,0)	15	1.4158e-05
A13	3	3	0	10	8.1729e-06
A14	10	10	0	3	6.2243e-05
A15	3	6	0	24	1.0849e-05
A16a	5	5	10	22	0
A16b	5	5	10	11	0
A16c	5	5	10	10	0
A16d	5	5	10	13	2.3575e-05
A17	2	3	0	23	1.0266e-05
A18	2	12	0	44	9.1178e-06
			1	57	2.8294e-06
			10	57	1.6509e-05
Ex. 3.2	2	2	1	8	0
			0.5	8	1.7113e-05
			(0.5,0)	6	1.8224e-05
			(0.9,0)	9	1.3397e-05

Table 1: Table with numerical results using different problems from the collection in [9] as well as Example 3.2

this parameter in Step (S.4) by $c_V^{k+1} := 5 \cdot c_V^k$ whenever $t_k < 1$ had to be chosen in (S.3); otherwise we set $c_V^{k+1} := c_V^k$. We also note that for Algorithm 5.1, the stepsize $t_k = 1$ is not necessarily a natural choice; therefore, we also allow a larger stepsize whenever this is possible, i.e., when $t_k = 1$ satisfies the criteria from (S.3), we test $t_k = 1/\beta$ and so on, until one of the conditions is violated for the first time. Finally, the values $\beta = 0.5$ and $\sigma = 10^{-4}$ were chosen for all test runs.

The numerical results obtained in this way are summarized in Table 1. All results are based on the choice $\alpha = 10^{-2}$ for the regularization parameter α in the definition of V. The test examples called A1–A18 are those taken from the report-version [9] of the paper [8], where all details are given and suitable references are provided. We choose the same starting point as in [9] which, however, does not necessarily belong to the feasible set W. Hence we first project this starting point onto W and then begin our iteration with this projected starting point. In some cases, this projection was already the solution of the underlying GNEP, and we therefore do not present results for these GNEPs; for example, the reader therefore does not find problem A2 in Table 1.

For each test example, Table 1 contains the following data: The name of the example, the number of players N, the total number of variables n, the starting point x^0 (all components of this starting point are equal to the number given here, unless more than one number is provided), the number of iterations k needed until convergence, and the final value of the objective function V in column V_{opt}.

The results from Table 1 may be difficult to interpret, however, it can be noted that

Algorithm 5.1 solves all test examples, whereas, for example, the penalty method from [9] has two failures on this set, namely on Examples A7 and A8, when using the third starting point. Furthermore, the number of iterations is quite reasonable and typically better than the corresponding number of iterations reported in [6] for an unconstrained optimization reformulation, especially because each function evaluation of the objective function in this unconstained optimization reformulation comes with the cost of the solution of two maximization problems, whereas in our case we only need to solve one maximization problem in order to compute $V(x)$.

Moreover, we would like to draw the attention to the results obtained for Example 3.2. Table 1 shows that we always converge to the unique solution of the GNEP and, hence, to the global minimum of problem P. This is particularly interesting since one can verify, similar to Example 3.2 where $\alpha = 0$ was chosen, that the objective function V still has a strict local minimum at $(1, 0)$ with a positive function value $V(1, 0) = 1 - \frac{\alpha}{2}$ (at least for all $\alpha < 0.5$), so that this minimum does not correspond to a solution of the GNEP. We even converge to the global minimum when the starting point is chosen close to the local minimum. In this respect, please recall that we cannot start exactly at the local minimum since V is not differentiable in this point.

Finally, we stress that we had to use a method for the solution of problem P which generates feasible iterates since otherwise V might not be well-defined. On the other hand, this also has the advantage that we may apply our method to problems where the objective functions θ_ν of the players ν are not defined outside of W due to some logarithmic terms, for example. This is in contrast to other existing methods which assume that the functions θ_ν are defined on the whole space $\mathbb{R}^n$.

6 Final Remarks

We investigated some structural properties of a constrained optimization reformulation of the player-convex generalized Nash equilibrium problem. In particular, we proved that, apart from some exceptional cases, the objective function of the optimization problem is differentiable at every minimum point. Hence the optimization problem is essentially differentiable and therefore allows the application of suitable algorithms for smooth optimization problems.

On the other hand, for jointly-convex GNEPs, one can characterize certain (normalized) solutions as the minima of a smooth optimization problem, cf. [13], whose objective function is once but not twice continuously differentiable. We believe that a similar analysis can be carried out in order to verify twice continuous differentiability of this function under suitable assumptions. The details are left as a future research topic.

Acknowledgements: The third author wishes to thank Christian Reger and Frieder Scharr for fruitful discussions on topics of this paper.

References

[1] K.J. Arrow, G. Debreu, *Existence of an equilibrium for a competitive economy,* Econometrica, Vol. 22 (1954), 265–290.

[2] J.R. Birge, L. Qi, Z. Wei, *A variant of the Topkins-Veinott method for solving inequality constrained optimization,* Journal of Applied Mathematics and Optimization, Vol. 41 (1997), 309–330.

[3] G. Debreu, *A social equilibrium existence theorem,* Proceedings of the National Academy of Sciences, Vol. 38 (1952), 886–893.

[4] D. Dorsch, H. Th. Jongen, V. Shikhman, *On structure and computation of generalized Nash equilibria,* Preprint No. 143, Department of Mathematics - C, RWTH Aachen University, 2011.

[5] A. Dreves, C. Kanzow, *Nonsmooth optimization reformulations characterizing all solutions of jointly convex generalized Nash equilibrium problems,* Computational Optimization and Applications, Vol. 50 (2011), 23–48.

[6] A. Dreves, C. Kanzow, O. Stein, *Nonsmooth optimization reformulations of player convex generalized Nash equilibrium problems,* Journal of Global Optimization, to appear.

[7] F. Facchinei, C. Kanzow, *Generalized Nash equilibrium problems,* Annals of Operations Research, Vol. 175 (2010), 177–211.

[8] F. Facchinei, C. Kanzow, *Penalty methods for the solution of generalized Nash equilibrium problems,* SIAM Journal on Optimization, Vol. 20 (2010), 2228–2253.

[9] F. Facchinei, C. Kanzow, *Penalty methods for the solution of generalized Nash equilibrium problems (with complete test problems,* Technical Report, Institute of Mathematics, University of Würzburg, Würzburg, Germany, February 2009.

[10] F. Facchinei, J.-S. Pang, *Nash equilibria: the variational approach,* in D.P. Palomar and Y.C. Eldar (eds.), *Convex Optimization in Signal Processing and Communications,* Cambridge University Press, 2009, 443–493.

[11] E.G. Gol'stein, *Theory of Convex Programming,* Translations of Mathematical Monographs, Vol. 36, American Mathematical Society, Providence, Rhode Island, 1972.

[12] G. Gürkan, J.-S. Pang, *Approximations of Nash equilibria,* Mathematical Programming, Vol. 117 (2009), 223–253.

[13] A. von Heusinger, C. Kanzow, *Optimization reformulations of the generalized Nash equilibrium problem using Nikaido-Isoda-type functions,* Computational Optimization and Applications, Vol. 43 (2009), 353–377.

[14] W.W. Hogan, *Directional derivatives for extremal-value functions with applications to the completely convex case,* Operations Research, Vol. 21 (1973), 188–209.

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

[16] J. Kyparisis, *On uniqueness of Kuhn-Tucker multipliers in nonlinear programming,* Mathematical Programming, Vol. 32 (1985), 242–246.

[17] J. Nash, *Non-cooperative games,* Annals of Mathematics, Vol. 54 (1951), 286–295.

[18] H. Nikaido, K. Isoda, *Note on noncooperative convex games,* Pacific Journal of Mathematics, Vol. 5 (1955), 807–815.

[19] J.-S. Pang, G. Scutari, *Nonconvex cognitive radio games — Part I: Quasi-Nash equilibria,* Technical Report, Department of Industrial and Enterprise Systems Engineering, University of Illinois, Urbana, USA, November 2010.

[20] J.-S. Pang, G. Scutari, *Nonconvex games with side constraints,* SIAM Journal on Optimization, Vol. 21 (2011), 1491–1522.

[21] R.T. Rockafellar, *Directional differentiability of the optimal value function in a nonlinear programming problem,* Mathematical Programming Study, Vol. 21 (1984), 213–226.

[22] O. Stein, *Bi-level Strategies in Semi-infinite Programming,* Kluwer, Boston, 2003.

[23] O. Stein, *On constraint qualifications in non-smooth optimization,* Journal of Optimization Theory and Applications, Vol. 121 (2004), 647–671.

[24] D.M. Topkins, A.F. Veinott, *On the convergence of some feasible direction algorithms for nonlinear programming,* SIAM Journal on Control, Vol. 5 (1967), 268–279.

[25] S. Uryasev, R.Y. Rubinstein, *On relaxation algorithms in computation of noncooperative equilibria,* IEEE Transactions on Automatic Control, Vol. 39 (1994), 1263–1267.

[26] G. Zoutendijk, *Methods of Feasible Directions,* Elsevier, Amsterdam, 1960.