ON A SMOOTH DUAL GAP FUNCTION FOR A CLASS OF PLAYER CONVEX GENERALIZED NASH EQUILIBRIUM PROBLEMS[1]

Nadja Harms, Tim Hoheisel, and Christian Kanzow

Preprint 322 March 2014

University of Würzburg
Institute of Mathematics
Emil-Fischer-Str. 30
97074 Würzburg
Germany

e-mail: nadja.harms@mathematik.uni-wuerzburg.de
hoheisel@mathematik.uni-wuerzburg.de
kanzow@mathematik.uni-wuerzburg.de

March 28, 2014

[1]This research was partially supported by the DFG (Deutsche Forschungsgemeinschaft) under grant KA 1296/18-1.

Abstract. We consider a class of generalized Nash equilibrium problems (GNEPs) where both the objective functions and the constraints are allowed to depend on the decision variables of the other players. It is well-known that this problem can be reformulated as a constrained optimization problem via the (regularized) Nikaido-Isoda-function, but this reformulation is usually nonsmooth. Here we observe that, under suitable conditions, this reformulation turns out to be the difference of two convex functions. This allows the application of the Toland-Singer duality theory in order to obtain a dual formulation which provides an unconstrained and continuously differentiable optimization reformulation of the GNEP. Moreover, based on a result from parametric optimization, the gradient of the unconstrained objective function is shown to be piecewise smooth. Some numerical results are presented to illustrate the theory.

Key Words: Generalized Nash equilibrium, DC optimization, conjugate function, dual gap function, nonconvex duality, optimal solution mapping, PC^1 function.

1 Introduction

This paper deals with the generalized Nash equilibrium problem (GNEP) where both the objective functions and, in contrast to the standard Nash equilibrium problem, also the constraints are allowed to depend on the decision variables of the other players. The precise definition of the GNEP considered here is given at the beginning of Section 3.

The GNEP has widespread applications and many solution methods exist in the meantime which work under different sets of assumptions. The interested reader is referred to [9] for a survey of applications, theory and algorithms up to the year 2010. There also exist quite a few newer contributions to this area, but the following discussion concentrates only on those which are particularly relevant for the approach followed in this paper where we want to obtain a suitable optimization reformulation of the GNEP.

One possibility to obtain such an optimization problem is to exploit the known equivalence of GNEPs to quasi-variational inequalities (QVIs), see [3, 16], and then to adapt the existing gap functions for QVIs like those discussed in [6, 13, 14, 19, 36] to the setting of GNEPs. This has been done, for example, in [2, 27]. However, except for some special cases, this yields a nondifferentiable optimization reformulation of the GNEP.

Another approach is to use the Nikaido-Isoda function that was originally introduced in [28] for theoretical purposes. Subsequently, it has also been exploited to derive some solution methods for certain classes of GNEPs, see [26, 39]. A variant is the regularized Nikaido-Isoda-function from [15] which might be used to obtain constrained and unconstrained optimization reformulations of the GNEP, see [7, 8, 18, 20]. While the optimization problems turn out to be smooth in the special case of jointly convex GNEPs, they only yield nondifferentiable reformulations in the player-convex case that will be discussed in this paper, see Section 3 for precise definitions.

The main motivation for this paper comes from the recent contribution [17] of the authors, where they extend an idea by Dietrich [5] and obtain a dual gap function for certain classes of QVIs based on the observation that one of the known regularized gap functions for QVIs can be viewed as a difference of convex functions, which then allows the application of the Toland-Singer duality theory [35, 37, 38] in order to get an unconstrained smooth optimization reformulation of certain QVIs.

The approach we follow here is similar to the one of the preceding paper [17]. The difference is that we use the regularized Nikaido-Isoda-function here, which seems to be better suited to GNEPs than any of the QVI-type gap functions. The main observation is again that also this regularized Nikaido-Isoda-function may be viewed, in a very natural way, as the difference of two convex functions, hence we can also apply the Toland-Singer duality theory in this setting in order to obtain a smooth and unconstrained optimization reformulation of a certain class of GNEPs. While the unconstrained objective function is usually not twice continuously differentiable, we show that its gradient is at least piecewise smooth under fairly mild conditions. This result is a consequence of a more general statement from parametric optimization which is also provided in this paper.

The organization is as follows: Section 2 first presents some background material from convex and variational analysis and then gives the above-mentioned smoothness result for a class of parametric optimization problems that fits into our framework. Section 3 then develops our smooth and unconstrained dual optimization reformulation of the GNEP. Section 4 applies the smoothness result from parametric optimization to our particular setting and therefore contains second-order properties of our unconstrained objective function. Section 5 presents some numerical results to illustrate our theory.

The notation used in this paper is pretty standard. The symbol $\|\cdot\|$ always denotes the Euclidean norm. Also, we put $\mathbb{R}_> := \{x \in \mathbb{R} \mid x > 0\}$. Moreover, for the support of a vector $x \in \mathbb{R}^n$, we set $\operatorname{supp} x := \{j \mid x_i \neq 0\} \subseteq \{1, \dots, n\}$. In addition to that, also given an index set $J \subseteq \{1, \dots, n\}$, we put $x_J = (x_j)_{j \in J} \in \mathbb{R}^{|J|}$. For a differentiable function $F : \mathbb{R}^n \to \mathbb{R}^m$, $DF(x)$ denotes its Jacobian at $x \in \mathbb{R}^n$, and for $m = 1$, $\nabla F(x) = DF(x)^T \in \mathbb{R}^n$ is its gradient. In the latter case, if F is twice differentiable, $\nabla^2 F(x)$ is the Hessian of F at x. The symbol $\exists!$ means that "there exists exactly one".

2 Preliminaries

2.1 Tools from Variational Analysis

In this subsection we review certain concepts from variational and convex analysis employed in the sequel. The notation and terminology is, in large parts, based on [32].

We first restate some definitions for set-valued mappings, see, e.g., [32, Chapter 5].

Definition 2.1 *Let $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a set-valued mapping. Then, Φ is called*

(a) outer semicontinuous (osc) *at $\bar{x} \in \mathbb{R}^n$ if for all sequences $\{x^k\} \subseteq \mathbb{R}^n$ with $x^k \to \bar{x}$ and all sequences $z^k \to \bar{z}$ with $z^k \in \Phi(x^k)$ for all $k \in \mathbb{N}$ sufficiently large we have $\bar{z} \in \Phi(\bar{x})$;*

(b) outer semicontinuous (osc) *on $\mathbb{R}^n$ if it is osc at every $x \in \mathbb{R}^n$;*

(c) graph-convex *if its graph $\operatorname{gph} \Phi = \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^m \mid z \in \Phi(x)\}$ is a convex set.*

The following properties of an osc and graph-convex set-valued mapping will be used in our subsequent analysis.

Lemma 2.2 *Let $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be an osc and graph-convex set-valued mapping. Then:*

(a) The sets $\Phi(x)$ are closed and convex (possibly empty).

(b) For all $x_1, x_2 \in \mathbb{R}^n$ with $\Phi(x_i) \neq \emptyset$ for $i = 1, 2$, and all $t \in [0, 1]$, we have

$$t\Phi(x_1) + (1-t)\Phi(x_2) \subseteq \Phi\big(tx_1 + (1-t)x_2\big),$$

in particular, the set on the right-hand side is nonempty.

(c) The set $\operatorname{gph} \Phi$ is closed and convex.

All statements are well known and easily verified; regarding assertion (b), see [32, p. 155].

We next introduce some important concepts for extended real-valued functions, more precisely, for functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Handy tools for the analysis of such a function are its *epigraph* $\operatorname{epi} f := \{(x, \gamma) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq \gamma\}$ and its *domain* $\operatorname{dom} f := \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$. Note that we call f *proper* if $\operatorname{dom} f \neq \emptyset$.

Definition 2.3 *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be proper.*

(a) f is called lower semicontinuous (lsc) *if $\operatorname{epi} f$ is closed.*

(b) f is called convex *if $\operatorname{epi} f$ is convex.*

(c) f is called strongly convex *with modulus $c > 0$ if $f - \frac{c}{2}\|\cdot\|^2$ is convex.*

(d) If f is convex and $\bar{x} \in \mathbb{R}^n$ then the (possibly empty) set

$$\partial f(\bar{x}) = \left\{ s \in \mathbb{R}^n \mid f(\bar{x}) + s^T(x - \bar{x}) \leq f(x) \quad \forall x \in \mathbb{R}^n \right\}$$

is called the subdifferential *of f at $\bar{x}$.*

(e) The conjugate *of f is the function $f^* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ defined by*

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \left[x^T y - f(x) \right] = \sup_{x \in \operatorname{dom} f} \left[x^T y - f(x) \right].$$

Note that, in view of its definition, an lsc function is often called *closed*. Further note that, for a proper and convex function f, the subdifferential $\partial f(\bar{x})$ is nonempty if x lies in the *(relative) interior* of $\operatorname{dom} f$, see [22].

Given $X \subseteq \mathbb{R}^n$, a popular extended real-valued function is the *indicator function* $\delta_X : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ defined by $\delta_X(x) := 0$ for $x \in X$ and $\delta_X(x) := +\infty$ for $x \notin X$. It is easily verified that δ_X is lsc if and only if X is closed, and convex if and only if X is convex.

The following result summarizes some well-known properties of the conjugate function.

Lemma 2.4 *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function. Then, the following statements hold:*

(a) The conjugate f^* of f is convex and lsc.

(b) The bi-conjugate function $f^{**} := (f^*)^*$ is convex and lsc.

(c) The inequality $f^{**}(x) \leq f(x)$ holds for all $x \in \mathbb{R}^n$.

(d) The equality $f^{**}(x) = f(x)$ holds for all $x \in \mathbb{R}^n$ if and only if f is a (convex and) lsc function.

(e) The Fenchel inequality $f(x) + f^*(y) \geq x^T y$ holds for all $x, y \in \mathbb{R}^n$.

(f) The equality $f(\bar{x}) + f^*(\bar{y}) = \bar{x}^T \bar{y}$ holds if and only if $\bar{y} \in \partial f(\bar{x})$.

All statements can entirely be found in [22, Chapter E]. Another useful observation on the conjugate function is restated in the following result, cf. [32, Prop. 12.60].

Lemma 2.5 *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper and lsc convex function. Then, f is strongly convex with modulus $c > 0$ if and only if f^* is differentiable with ∇f^* Lipschitz continuous with modulus $\frac{1}{c}$.*

2.2 A Piecewise Smoothness Result for a Parametric NLP

In this subsection, we analyze smoothness properties of the solution mapping for a class of strongly convex parametric optimization problems, where the parameter only occurs in the objective function. The main result in this section might be known, but we could not find an explicit reference. The difference to the existing literature is that we assume the objective function to be strongly convex, not just convex, which is, of course, a very restrictive assumption, but this assumption will be satisfied automatically in our applications. On the other hand, no Slater condition is required for the constraints.

It will be seen that the solution function of our parametric optimization problem is, under some standard assumptions, *piecewise smooth*. The analysis is carried out in the spirit of the results from [30] in combination with [23]. We commence by introducing the concept of piecewise smoothness, see [12, 33] for comprehensive accounts on the topic.

Definition 2.6 *A continuous function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is called* piecewise smooth *or* PC[1] *near $\bar{x} \in D$ if there exists an open neighborhood $U \subseteq D$ of $\bar{x}$ and a finite family of continuously differentiable functions $f_i : U \to \mathbb{R}^m$ $(i = 1, \dots, l)$ such that $f(x) \in \{f_1(x), \dots, f_l(x)\}$ for all $x \in U$.*

Now, for a parameter $v \in \mathbb{R}^n$, consider the optimization problem

$$\min_{u \in \mathbb{R}^m} \quad \phi(u, v) \quad \text{s.t.} \quad c_j(u) \leq 0 \quad (j = 1, \dots, p), \tag{$P(v)$}$$

where $\phi : (u, v) \in \mathbb{R}^m \times \mathbb{R}^n \mapsto \mathbb{R}$ is strongly convex in u for each fixed $v \in \mathbb{R}^n$ and continuous on $\mathbb{R}^m \times \mathbb{R}^n$, and the functions $c_j : \mathbb{R}^m \to \mathbb{R}$ $(j = 1, \dots, p)$ are convex and continuous. Let

$$\mathcal{F} := \{u \in \mathbb{R}^m \mid c_j(u) \leq 0 \ (j = 1, \dots, p)\}$$

denote the feasible set, which is, in particular, closed and convex, and is supposed to be nonempty. Under the assumptions from above, the next lemma shows that the solution mapping of the problem $P(v)$ is continuous.

Lemma 2.7 *The solution mapping $u^* : \mathbb{R}^n \to \mathbb{R}^m$ of the problem $P(v)$ given by*

$$u^*(v) = \operatorname*{argmin}_{u \in \mathcal{F}} \phi(u, v) \tag{1}$$

is well-defined and continuous.

Proof. Under the assumptions on ϕ and $c_j (j = 1, \ldots, p)$, the objective function is strongly convex in u and the feasible set $\mathcal{F}$ is nonempty, closed, and convex. Hence, the problem $P(v)$ is uniquely solvable for all $v \in \mathbb{R}^n$. Therefore, for each $v \in \mathbb{R}^n$, there exists a unique vector $u^*(v)$ solving (1). Therefore, the solution mapping u^* is well-defined. The continuity of the mapping u^* follows from [23, Corollaries 8.1 and 9.1]. □

Now, for $u \in \mathcal{F}$, we define the active set $J_0(u) := \{j \mid c_j(u) = 0\}$. Due to Lemma 2.7, for all $v \in \mathbb{R}^n$, the sets $J(v) := J_0(u^*(v))$ are well-defined.

For the sequel of this subsection, we assume that all functions defining $P(v)$ are, in addition to the convexity properties, twice continuously differentiable. As a reminder and a reference point, all of the demanded properties are summarized below.

Assumption 2.8 *The functions ϕ and c_j $(j = 1, \ldots, p)$ defining $P(v)$ are assumed to have the following properties:*

(a) The objective ϕ is strongly convex in the first variable (while fixing the second) and twice continuously differentiable in both components.

(b) The constraints c_j $(j = 1, \ldots, p)$ are convex and twice continuously differentiable.

(c) The feasible set $\mathcal{F} := \{u \in \mathbb{R}^m \mid c_j(u) \leq 0 \ (j = 1, \ldots, p)\}$ is nonempty.

For $v \in \mathbb{R}^n$ and a subset $J \subseteq J(v)$, we define $H^J(\cdot, v, \cdot) : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}^{m+p}$ by

$$H^J(u, v, \lambda) := \begin{pmatrix} \nabla_u \phi(u, v) + \sum_{j \in J} \lambda_j \nabla c_j(u) \\ c_J(u) \\ \lambda_{\hat{J}} \end{pmatrix},$$

where $\hat{J} := \{1, \ldots, p\} \setminus J$. Then, the following result is easily proven.

Lemma 2.9 *Let Assumption 2.8 hold, let $v \in \mathbb{R}^n$, and let $J \subseteq J(v)$ such that the vectors $\nabla c_j(u)$ $(j \in J)$ are linearly independent. Then, the Jacobian $D_{(u,\lambda)} H^J(u, v, \lambda)$ is nonsingular for all $\lambda_J \geq 0$.*

Proof. After reordering the components of λ accordingly, we get

$$D_{(u,\lambda)} H^J(u, v, \lambda) = \begin{pmatrix} \nabla^2_{uu} \phi(u, v) + \sum_{j \in J} \lambda_j \nabla^2 c_j(u) & Dc_J(u)^T & 0 \\ Dc_J(u) & 0 & 0 \\ 0 & 0 & I_{|\hat{J}|} \end{pmatrix}.$$

Since c_j $(j = 1, \ldots, p)$ is convex and ϕ is strongly convex in the first variable for each fixed $v \in \mathbb{R}^n$, the matrix $\nabla^2_{uu} \phi(u, v) + \sum_{j \in J} \lambda_j \nabla^2 c_j(u)$ is positive definite for all $\lambda_J \geq 0$. Hence, the assertion follows from the linear independence of the vectors $\nabla c_j(u)$ $(j \in J)$. □

We next introduce the *constant rank constraint qualification* due to [25], which occurs as a standard assumption in the context of parametric optimization and piecewise smoothness results, see, e.g., [4, 29, 30].

Definition 2.10 *We say that the* constant rank constraint qualification (CRCQ) *holds at* $\bar{u} \in \mathcal{F}$ *(w.r.t.* $\mathcal{F}$*) if there exists a neighborhood* U *of* $\bar{u}$ *such that for every* $J \subseteq J_0(\bar{u})$ *the set* $\{\nabla c_j(u) \mid j \in J\}$ *has constant rank (depending on* J*) for all* $u \in U$*.*

Note that CRCQ is a local property of the feasible set $\mathcal{F}$ in the sense that if CRCQ holds at $\bar{u}$, it also holds at u for all $u \in \mathcal{F}$ sufficiently close to $\bar{u}$. — CRCQ allows us to prove the next main result on the piecewise smoothness of the solution mapping of the program $P(v)$.

Theorem 2.11 *Let* $\bar{v} \in \mathbb{R}^n$*, and suppose that Assumption 2.8 is fulfilled. Then, there exists a neighborhood* $\bar{V}$ *of* $\bar{v}$ *such that the function* $u^* : \mathbb{R}^n \to \mathbb{R}^m$ *defined in* (1) *is* PC^1 *on* $\bar{V}$*, provided CRCQ holds at* $\bar{u} := u^*(\bar{v}) \in \mathcal{F}$*.*

Proof. For $v \in \mathbb{R}^n$ we define

$$M(v) := \{\lambda \in \mathbb{R}^p \mid (u^*(v), \lambda) \text{ is a KKT point of } P(v)\}$$

as the set of KKT multipliers for $P(v)$ at $u^*(v)$. Since CRCQ at $\bar{u} = u^*(\bar{v})$ is inherited to a whole neighborhood and because u^* is continuous by Lemma 2.7, there exists a neighborhood V of $\bar{v}$ such that CRCQ holds at $u^*(v)$ for all $v \in V$. In particular, since CRCQ yields KKT multipliers at a local minimizer (see [25, Proposition 2.3]), we have $M(v) \neq \emptyset$ for all $v \in V$. It hence follows from [21, Lemma 3.2] that the set

$$B(v) := \left\{ J \subseteq J(v) \mid \nabla c_j(u^*(v)) \; (j \in J) \text{ linearly independent } \wedge \; \exists \lambda \in M(v) : \operatorname{supp} \lambda \subseteq J \right\}$$

is nonempty for all $v \in V$. Moreover, from [21, Lemma 3.3] it follows that

$$\forall v \in V, \; J \in B(v) \; \exists! \; \lambda^{*,J}(v) \in M(v) : \; H^J(u^*(v), v, \lambda^{*,J}(v)) = 0. \tag{2}$$

Note that, necessarily, $\operatorname{supp} \lambda^{*,J}(v) \subseteq J$, and that $\lambda^{*,J}(v)$ is nonnegative.

Now, let $J \in B(\bar{v})$ and $\bar{\lambda}^J := \lambda^{*,J}(\bar{v})$ such that $\bar{\lambda}^J \in M(\bar{v})$ and $H^J(\bar{u}, \bar{v}, \bar{\lambda}^J) = 0$, uniquely determined by (2). As $J \in B(\bar{v})$, the vectors $\nabla c_j(\bar{u})$ $(j \in J)$ are linearly independent, hence Lemma 2.9 together with $\bar{\lambda}^J \geq 0$ implies that $D_{(u,\lambda)} H^J(\bar{u}, \bar{v}, \bar{\lambda}^J)$ is nonsingular. Thus, the implicit function theorem yields neighborhoods V^J of $\bar{v}$ and N^J of $(\bar{u}, \bar{\lambda}^J)$, and a C^1 function $(u^J, \lambda^J) : V^J \to N^J$ such that

$$u^J(\bar{v}) = \bar{u}, \quad \lambda^J(\bar{v}) = \bar{\lambda}^J, \quad \text{and} \quad H^J(u^J(v), v, \lambda^J(v)) = 0 \quad \forall v \in V^J, \tag{3}$$

and for all $v \in V^J$ the vector $(u^J(v), \lambda^J(v))$ is the unique solution of

$$H(u, v, \lambda) \stackrel{!}{=} 0, \quad (u, \lambda) \in N^J.$$

Note that, w.l.o.g., we can assume that $V^J \subseteq V$.

Now, set $\bar{V} := \bigcap_{J \in B(\bar{v})} V^J \subseteq V$. Since $B(\bar{v})$ is finite, $\bar{V}$ is still a neighborhood of $\bar{v}$. Moreover, in view of [21, Lemma 3.5 (b)]), we can assume w.l.o.g. that $B(v) \subseteq B(\bar{v})$ for all $v \in \bar{V}$. We will now prove that, with a possibly smaller neighborhood of $\bar{v}$ which we still denote by $\bar{V}$, we have

$$u^*(v) \in \{u^J(v) \mid J \in B(\bar{v})\} \quad \forall v \in \bar{V}. \tag{4}$$

Then, it follows that $u^* : \bar{V} \to \mathbb{R}^m$ is in fact PC^1, as $\{u^J : \bar{V} \to \mathbb{R}^m \mid J \in B(\bar{v})\}$ is a finite family of C^1 functions, and u^* is continuous by Lemma 2.7. The desired inclusion in (4) follows immediately if we can show that

$$\forall v \in \bar{V},\ \forall J \in B(v):\ u^*(v) = u^J(v) \tag{5}$$

since $B(v) \subseteq B(\bar{v})$ for all $v \in \bar{V}$. Note that this does not imply that $u^* = u^J$ holds locally (which would imply u^* to be smooth) since the index set J also depends on v.

For these purposes, let $v \in \bar{V}(\subseteq V)$ and $J \in B(v)$. Due to (2), there exists a unique multiplier $\lambda^{*,J}(v) \in M(v)$ such that $H^J(u^*(v), v, \lambda^{*,J}(v)) = 0$. On the other hand, due to what was shown above, there exists a neighborhood V^J of $\bar{v}$ and a neighborhood N^J of $(\bar{u}, \bar{\lambda}^J)$ as well as a C^1 function $(u^J, \lambda^J) : V^J \to N^J$ such that (3) holds. Moreover, for all $v \in V^J$, the vector $(u^J(v), \lambda^J(v))$ is the unique solution of $H(u, v, \lambda) \stackrel{!}{=} 0,\ (u, \lambda) \in N^J$. Hence, in order to prove (5), it suffices to show that

$$\text{for all } v \in \bar{V} \text{ sufficiently close to } \bar{v},\ \forall J \in B(v):\ (u^*(v), \lambda^{*,J}(v)) \in N^J.$$

Suppose this were false: Then, there exists a convergent sequence $\{v^k \in \bar{V}\} \to \bar{v}$ and a sequence of index sets $\{J_k \in B(v^k)\}$ such that $(u^*(v^k), \lambda^{*,J_k}(v^k)) \notin N^{J_k}$ for all $k \in \mathbb{N}$. As $B(\bar{v})$ is finite and $B(v^k) \subseteq B(\bar{v})$ for all $k \in \mathbb{N}$, we can assume w.l.o.g. that $J_k = \bar{J}$ for all $k \in \mathbb{N}$. From (2) we infer that

$$0 = \nabla_u \phi(u^*(v^k), v^k) + \sum_{j \in \bar{J}} [\lambda^{*,\bar{J}}(v^k)]_j \nabla c_j(u^*(v^k)) \quad \forall k \in \mathbb{N}. \tag{6}$$

By continuity of all functions involved, the linear independence of the gradient vectors $\nabla c_j(u^*(v^k))$ ($j \in \bar{J}$) for all $k \in \mathbb{N}$ together with the assumed CRCQ condition and the fact that $\operatorname{supp} \lambda^{*,\bar{J}}(v^k) \subseteq \bar{J}$, we infer that $\{\lambda^{*,\bar{J}}(v^k)\}$ is convergent, i.e. there exists $\lambda^{*,\bar{J}}$ such that $\lambda^{*,\bar{J}}(v^k) \to \lambda^{*,\bar{J}}$ with $\operatorname{supp} \lambda^{*,\bar{J}} \subseteq \bar{J}$. Hence, passing to the limit in (6) yields

$$0 = \nabla_u \phi(\bar{u}, \bar{v}) + \sum_{j \in \bar{J}} [\lambda^{*,\bar{J}}]_j \nabla c_j(\bar{u}).$$

On the other hand, also the vector $\bar{\lambda}^{\bar{J}}$ solves the above equation. Due to the linear independence of the gradients $\nabla c_j(\bar{u})$ ($j \in \bar{J}$), and the fact that $\operatorname{supp} \bar{\lambda}^{\bar{J}} \cup \operatorname{supp} \lambda^{*,\bar{J}} \subseteq \bar{J}$, we get $\lambda^{*,\bar{J}} = \bar{\lambda}^{\bar{J}}$. Therefore, we infer that $\lambda^{*,J_k}(v^k) \to \bar{\lambda}^{\bar{J}}$. In view of $u^*(v^k) \to \bar{u}$ by continuity, we get that $(u^*(v^k), \lambda^{*,J_k}(v^k)) \in N^{\bar{J}}$ for all k sufficiently large, in contradiction to what was assumed. Hence, the proof is complete. $\square$

3 The Smooth Dual Gap Function

The generalized Nash equilibrium problem (GNEP) considered in this paper consists of $N \in \mathbb{N}$ players which control the corresponding decision variables $x^\nu \in \mathbb{R}^{n_\nu}$ $(\nu = 1, \ldots, N)$ so that the vector $x = (x^1, \ldots, x^N) \in \mathbb{R}^n$ with $n = n_1 + \ldots + n_N$ describes the decision vector of all players. In order to emphasize the role of player ν's variable x^ν within the vector x, we often write $x = (x^\nu, x^{-\nu})$. Each player ν has a cost function $\theta_\nu : \mathbb{R}^n \to \mathbb{R}$ and a strategy space $X_\nu(x^{-\nu}) \subseteq \mathbb{R}^\nu$ defined by the set-valued mapping $X_\nu : \mathbb{R}^{n-n_\nu} \rightrightarrows \mathbb{R}^{n_\nu}$ where both the cost function and the strategy space depend on the other players' decisions $x^{-\nu}$. Then, the GNEP consists in finding a vector $\bar{x} = (\bar{x}^1, \ldots, \bar{x}^N) \in \Omega(\bar{x})$ such that, for each $\nu \in \{1, \ldots, N\}$, the vector $\bar{x}^\nu$ solves

$$\min_{x^\nu} \; \theta_\nu(x^\nu, \bar{x}^{-\nu}) \quad \text{s.t.} \quad x^\nu \in X_\nu(\bar{x}^{-\nu}) \tag{7}$$

where

$$\Omega(x) := X_1(x^{-1}) \times \ldots \times X_N(x^{-N}). \tag{8}$$

A solution point $\bar{x}$ of GNEP is also called a *generalized Nash equilibrium.*

In this section, we consider suitable optimization problems which are reformulations of the player convex GNEP, where the GNEP is called *player convex* if the following assumptions are satisfied.

Assumption 3.1 *(a) The cost functions θ_ν, $\nu = 1, \ldots, N$, are continuous on $\mathbb{R}^n$.*

(b) The cost functions $\theta_\nu(\cdot, x^{-\nu})$, $\nu = 1, \ldots, N$, are convex for each fixed $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$.

(c) The strategy spaces $X_\nu(x^{-\nu})$, $\nu = 1, \ldots, N$, are closed and convex.

Note that Assumption 3.1 (c) is satisfied if, e.g., the strategy spaces X_ν are defined by

$$X_\nu(x^{-\nu}) \; := \; \{x^\nu \in \mathbb{R}^{n_\nu} | \; g^\nu(x^\nu, x^{-\nu}) \leq 0\} \tag{9}$$

with functions $g^\nu : \mathbb{R}^n \to \mathbb{R}^{m_\nu}$, $\nu = 1, \ldots, N$, which are continuous on $\mathbb{R}^n$ and convex in x^ν for each fixed $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$.

First, we consider the so-called *Nikaido-Isoda function* ([28])

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

and the optimal value function

$$V(x) := \sup_{z \in \Omega(x)} \; \psi(z, x), \tag{10}$$

which takes the value $-\infty$ exactly for $x \notin \operatorname{dom}\Omega$, where

$$\operatorname{dom}\Omega := \{x \in \mathbb{R}^n |\ \Omega(x) \neq \emptyset\} \tag{11}$$

is the *domain* of the set-valued mapping Ω. Note that the supremum in (10) may be a nonuniquely attained maximum or not attained at all, because ψ is, in general, just concave in z for each fixed $x \in \mathbb{R}^n$. Furthermore, it is easily verified that V is nonnegative for all $x \in \Omega(x)$ and that $\bar{x}$ is a generalized Nash equilibrium of the player convex GNEP if and only if $\bar{x} \in \Omega(\bar{x})$ and $V(\bar{x}) = 0$. Let

$$W := \{x \in \mathbb{R}^n \mid x^\nu \in X_\nu(x^{-\nu})\ \forall \nu = 1, \ldots, N\} \tag{12}$$

be the *fixed point set* of the set-valued mapping Ω which is also called the *feasible set* of the corresponding GNEP. Since $x \in \Omega(x)$ if and only if $x \in W$, the player convex GNEP is equivalent to finding a solution of the constrained minimization problem

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

with zero as the optimal value.

In order to guarantee the existence of unique maximal points in (10) on $\operatorname{dom}\Omega$, we replace the function ψ by the *regularized Nikaido-Isoda function* ([15])

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

where $\alpha > 0$ denotes a given parameter. In view of Assumption 3.1 (b), the function ψ_α is strongly concave in z for each fixed $x \in \mathbb{R}^n$. Hence, for all $x \in \operatorname{dom}\Omega$, there exists a unique solution $z_\alpha(x)$ of the maximization problem

$$\max_z\ \psi_\alpha(z, x) \quad \text{s.t.} \quad z \in \Omega(x).$$

Therefore, the optimal value function

$$V_\alpha(x) := \sup_{z \in \Omega(x)}\ \psi_\alpha(z, x) = \sum_{\nu=1}^{N} \theta_\nu(x^\nu, x^{-\nu}) - \inf_{z \in \Omega(x)} \Big(\sum_{\nu=1}^{N} \theta_\nu(z^\nu, x^{-\nu}) + \frac{\alpha}{2}\|x - z\|^2\Big) \tag{13}$$

takes the value $-\infty$ exactly for $x \notin \operatorname{dom}\Omega$ and is real-valued for all $x \in \operatorname{dom}\Omega$. Further properties of V_α are given in the following result, the proof of which can be found in [8].

Lemma 3.2 *Under Assumption 3.1, the following statements hold:*

(a) $x \in \Omega(x)$ *if and only if* $x \in W$*; in particular, we have* $W \subseteq \operatorname{dom}\Omega$ *and* V_α *is real-valued on* W*;*

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

(c) $\bar{x}$ *is a generalized Nash equilibrium if and only if* $\bar{x} \in W$ *and* $V_\alpha(\bar{x}) = 0$*;*

(d) For all $x \in \operatorname{dom}\Omega$*, there exists a unique vector* $z_\alpha(x)$ *such that*

$$z_\alpha(x) = \underset{z \in \Omega(x)}{\operatorname{argmin}} \Big(\sum_{\nu=1}^{N} \theta_\nu(z^\nu, x^{-\nu}) + \frac{\alpha}{2}\|x - z\|^2 \Big); \tag{14}$$

(e) $\bar{x}$ *is a generalized Nash equilibrium if and only if* $\bar{x} = z_\alpha(\bar{x})$*.*

It follows from Lemma 3.2 (a)–(c) that the player convex GNEP is equivalent to finding a solution of the constrained minimization problem

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

or, alternatively, using the indicator function δ_W of W, to solving the unconstrained minimization problem

$$\min_{x \in \mathbb{R}^n} \; [V_\alpha(x) + \delta_W(x)] \tag{15}$$

with zero optimal value in both reformulations and with the convention $\eta + \infty = +\infty$ for all $\eta \in \mathbb{R} \cup \{\pm\infty\}$ in the unconstrained reformulation. This convention makes sense since the objective function from (15) should take the function value $+\infty$ on the complement W^c of the set W and, in particular, on $(\operatorname{dom}\Omega)^c \subseteq W^c$. Note that, in general, the optimal value function $V_\alpha + \delta_W$ is nonconvex and nondifferentiable as the functions V and V_α.

Similar to the approach from [17] for quasi-variational inequalities, it is possible to get a smooth reformulation of certain GNEPs in case that the optimal value function from (15) can be rewritten as a difference of two strongly convex and lsc functions. A class of GNEPs satisfying the next assumption has this property.

Assumption 3.3 *(a) The feasible set* W *of the GNEP* (7) *defined in* (12) *is nonempty.*

(b) The cost functions θ_ν*,* $\nu = 1, \ldots, N$*, are continuous and convex on* $\mathbb{R}^n$*.*

(c) The set-valued mappings X_ν*,* $\nu = 1, \ldots, N$*, are graph-convex and osc on* $\mathbb{R}^{n-n_\nu}$*.*

Since these assumptions play a central role within our subsequent analysis, we would like to add a few comments. Assumption 3.3 (a) is rather natural since otherwise the corresponding GNEP is not solvable. Assumption 3.3 (c) is satisfied if, e.g., the set-valued mappings X_ν, $\nu = 1, \ldots, N$, are defined by (9) with functions $g^\nu : \mathbb{R}^n \to \mathbb{R}^{m_\nu}$, $\nu = 1, \ldots, N$, which are convex in the whole variable $x = (x^\nu, x^{-\nu})$ on $\mathbb{R}^n$. In particular, Assumption 3.3 (c) therefore holds for the class of jointly convex GNEPs, where $g^1 = g^2 = \ldots = g^N =: g$ and g is convex in all variables, cf. [9] for more details. Finally, Assumption 3.3 (b) is probably the most restrictive condition since it requires all cost functions to be convex in the entire vector x. However, we will see later that this assumption can be relaxed considerably, see the discussion following Lemma 3.9. In order to avoid any technical discussion, it is convenient to assume this condition to formulate and prove the subsequent results. — Assumption 3.3 (c) yields the following result.

Lemma 3.4 *Let Assumption 3.3 (c) hold. Then, the set-valued mapping* $\Omega : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *defined by* (8) *is graph-convex and osc on* $\mathbb{R}^n$.

Proof. First, we verify the graph-convexity of Ω. By the definition of Ω in (8) and Assumption 3.3 (c), it holds that

$$\begin{aligned}\operatorname{gph}\Omega &= \{(x,z) \in \mathbb{R}^n \times \mathbb{R}^n \mid z^\nu \in X_\nu(x^{-\nu})\ \forall \nu = 1,\dots,N\} \\ &= \bigcap_{\nu=1}^{N} \{(x,z) \in \mathbb{R}^n \times \mathbb{R}^n \mid (x^{-\nu}, z^\nu) \in \operatorname{gph} X_\nu\} =: \bigcap_{\nu=1}^{N} C_\nu\end{aligned}$$

with the convex sets $\operatorname{gph} X_\nu = \{(x^{-\nu}, z^\nu) \in \mathbb{R}^{n-n_\nu} \times \mathbb{R}^{n_\nu} \mid z^\nu \in X_\nu(x^{-\nu})\}$ for all $\nu \in \{1,\dots,N\}$. Furthermore, the sets C_ν, $\nu = 1,\dots,N$, are convex as suitable Cartesian products of the convex sets $\operatorname{gph} X_\nu$ and $\mathbb{R}^{n_\nu} \times \mathbb{R}^{n-n_\nu}$. Then, the set $\operatorname{gph}\Omega$ is convex as an intersection of convex sets. Therefore, the set-valued mapping Ω is graph-convex.

Next, we show that the set-valued mapping Ω is osc on $\mathbb{R}^n$. Let $\bar{x} \in \mathbb{R}^n$ be arbitrarily given. Since the set-valued mappings $X_\nu : \mathbb{R}^{n-n_\nu} \rightrightarrows \mathbb{R}^{n_\nu}$, $\nu = 1,\dots,N$, are osc on $\mathbb{R}^{n-n_\nu}$, for all sequences $\{x^{k,-\nu}\} \subseteq \mathbb{R}^{n-n_\nu}$ with $x^{k,-\nu} \to \bar{x}^{-\nu}$ and all sequences $z^{k,\nu} \to \bar{z}^\nu$ with $z^{k,\nu} \in X_\nu(x^{k,-\nu})$ for all $k \in \mathbb{N}$ sufficiently large we have $\bar{z}^\nu \in X_\nu(\bar{x}^{-\nu})$. Then, for all sequences $\{x^k\} \subseteq \mathbb{R}^n$ with $x^k \to \bar{x}$ and $z^k \to \bar{z}$ with $z^k \in X_1(x^{k,-1}) \times \ldots \times X_N(x^{k,-N}) = \Omega(x^k)$ for all $k \in \mathbb{N}$ sufficiently large we have $\bar{z} \in X_1(\bar{x}^{-1}) \times \ldots \times X_N(\bar{x}^{-N}) = \Omega(\bar{x})$. Consequently, Ω is osc at $\bar{x}$. Since $\bar{x} \in \mathbb{R}^n$ was arbitrarily chosen, the set-valued mapping Ω is osc on $\mathbb{R}^n$. □

Lemma 3.5 *Let Assumptions 3.3 (a) and (c) hold. Then:*

(a) the feasible set W *of the GNEP* (7) *defined in* (12) *is nonempty, closed, and convex.*

(b) the domain $\operatorname{dom}\Omega$ *from* (11) *of the set-valued mapping* Ω *is nonempty and convex.*

Proof. (a) In view of Assumption 3.3 (a), the set W is nonempty. Furthermore, let $\{x^k\} \subseteq W$ be an arbitrary convergent sequence with a limit $\bar{x} \in \mathbb{R}^n$. Then $x^k \in \Omega(x^k)$ for all $k \in \mathbb{N}$. Since the set-valued mapping $\Omega : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is osc by Lemma 3.4, it follows that $\bar{x} \in \Omega(\bar{x})$. Therefore, $\bar{x} \in W$ so that the set W is closed.

Next, we show that W is convex. To this end, let $x_1, x_2 \in W$ and $t \in [0,1]$ be arbitrarily given. Then $x_1 \in \Omega(x_1)$ and $x_2 \in \Omega(x_2)$. By Lemmas 3.4 and 2.2 (b), it follows that Ω is graph-convex and $tx_1 + (1-t)x_2 \in \Omega\big(tx_1 + (1-t)x_2\big)$, i.e., $tx_1 + (1-t)x_2 \in W$. Hence, the set W is convex.

(b) The set $\operatorname{dom}\Omega$ is nonempty since $\operatorname{dom}\Omega$ contains the nonempty set W in view of Lemma 3.2 (a). The convexity of $\operatorname{dom}\Omega$ follows immediately from graph-convexity of Ω, see Lemma 2.2 (b). □

The subsequent example illustrates that even for a graph-convex and osc set-valued mapping Ω, its domain $\operatorname{dom}\Omega$ is not necessarily closed.

Example 3.6 Consider a GNEP with two players having arbitrary cost functions and each controlling a single variable which, for simplicity of notation, we call x_1 and x_2, respectively. Furthermore, let $\Omega : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ be given by $\Omega(x) = X_1(x_2) \times X_2(x_1)$ with the set-valued mappings X_1, $X_2 : \mathbb{R} \rightrightarrows \mathbb{R}$ defined by

$$X_1(x_2) := \begin{cases} \{x_1 \in \mathbb{R} \mid x_1 \geq \frac{1}{x_2}\} & \text{if } x_2 > 0, \\ \emptyset & \text{if } x_2 \leq 0, \end{cases} \quad \text{and} \quad X_2(x_1) := [0, \infty[.$$

These set-valued mappings X_1 and X_2 are obviously graph-convex. Furthermore, X_1 and X_2 are osc on $\mathbb{R}$, since, if $x_2^k \downarrow 0$, all sequences $\{z_1^k\}$ with $z_1^k \in X_1(x_2^k)$ are divergent, and all other cases are unproblematic. In view of Lemma 3.4, the set-valued mapping Ω is also graph-convex and osc on $\mathbb{R}^2$. On the other hand, $\operatorname{dom}\Omega = \mathbb{R} \times \mathbb{R}_>$ is not closed. $\diamondsuit$

In order to get a differentiable reformulation of GNEPs satisfying Assumption 3.3, we rewrite the unconstrained objective function from (15) as a difference of two strongly convex and lsc functions and apply the duality theory by Toland [37] and Singer [35] to this DC minimization problem. Note that a problem is called *DC minimization problem* if it consists of the minimization of a difference of two convex functions. For a survey of DC programming, we refer to [24].

Using the reformulation of the optimal value function V_α in (13), we first get

$$V_\alpha(x) + \delta_W(x) = \varrho(x) - \varphi_\alpha(x) \tag{16}$$

with the functions $\varrho : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $\varphi_\alpha : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\varrho(x) := \sum_{\nu=1}^{N} \theta_\nu(x) + \delta_W(x) \quad \text{and} \quad \varphi_\alpha(x) := \inf_{z \in \Omega(x)} \Big(\sum_{\nu=1}^{N} \theta_\nu(z^\nu, x^{-\nu}) + \frac{\alpha}{2}\|x - z\|^2 \Big) \tag{17}$$

where the infimum is uniquely attained at $z_\alpha(x)$ defined in (14) for all $x \in \operatorname{dom}\Omega$, and takes the value $+\infty$ for all $x \notin \operatorname{dom}\Omega$. The functions ϱ and φ_α are lsc and convex. Lower semicontinuity and convexity of ϱ is easily verified since θ_ν, $\nu = 1, \ldots, N$, is continuous and convex on $\mathbb{R}^n$ by Assumption 3.3 (b) and the set W is closed and convex by Lemma 3.5 (a). For the proof of lower semicontinuity and convexity of φ_α we need the following result, which was proven in [17, Lemma 3.3].

Lemma 3.7 *Let $\Omega : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be graph-convex and osc on $\mathbb{R}^n$. Then, the function $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $\Phi(z, x) := \delta_{\Omega(x)}(z)$ is lsc and convex in (z, x).*

Using Lemma 3.7, we are now in position to verify lower semicontinuity and convexity of the function φ_α defined in (17).

Lemma 3.8 *Let Assumption 3.3 hold. Then, the function φ_α is proper, lsc, and convex.*

Proof. In view of (17), we rewrite φ_α as $\varphi_\alpha(x) = \inf_{z\in\mathbb{R}^n} \tau_\alpha(z,x)$ with

$$\tau_\alpha : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\},\ \tau_\alpha(z,x) := \sum_{\nu=1}^{N} \theta_\nu(z^\nu, x^{-\nu}) + \frac{\alpha}{2}\|x-z\|^2 + \delta_{\Omega(x)}(z).$$

By Assumption 3.3 and Lemma 3.7, each summand of τ_α is convex and (at least) lsc on $\mathbb{R}^n$. Hence, the function τ_α is lsc and convex. Furthermore, the function τ_α is proper, since $\operatorname{dom}\tau_\alpha = \operatorname{dom}\Omega \neq \emptyset$. Moreover, it holds that $\operatorname{argmin}_{z\in\mathbb{R}^n}\ \tau_\alpha(z,x) = \{z_\alpha(x)\}$ for all $x \in \operatorname{dom}\Omega$ is single-valued. Therefore, the assertions follow from [32, Corollary 3.32]. □

Since the functions ϱ and φ_α are lsc and convex, the representation in (16) is a lsc DC formulation of the unconstrained objective function from (15). For the purpose of a differentiable dual reformulation of GNEPs satisfying Assumption 3.3, we add to both functions ϱ and φ_α the same strongly convex quadratic term. This alteration leads to the following DC decomposition of the optimal value function from (15):

$$V_\alpha(x) + \delta_W(x) = f_\alpha(x) - h_\alpha(x)$$

with two functions $f_\alpha, h_\alpha : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ defined by

$$f_\alpha(x) := \frac{\alpha}{2}\|x\|^2 + \sum_{\nu=1}^{N} \theta_\nu(x) + \delta_W(x) = \frac{\alpha}{2}\|x\|^2 + \varrho(x), \tag{18}$$

$$h_\alpha(x) := \frac{\alpha}{2}\|x\|^2 + \inf_{z\in\Omega(x)} \Big(\sum_{\nu=1}^{N} \theta_\nu(z^\nu, x^{-\nu}) + \frac{\alpha}{2}\|x-z\|^2 \Big) = \frac{\alpha}{2}\|x\|^2 + \varphi_\alpha(x). \tag{19}$$

In principle, we could have used a different parameter for the quadratic term. Furthermore, this quadratic term could be replaced by any strongly convex function without really changing the subsequent theory.

Some elementary properties of the above DC decomposition are summarized in the following result.

Lemma 3.9 *Let Assumption 3.3 hold, and let f_α and h_α be defined as in* (18) *and* (19), *respectively. Then, the following statements hold:*

(a) The function f_α is lsc and strongly convex on $\mathbb{R}^n$ and has the domain W.

(b) The function h_α is lsc and strongly convex on $\mathbb{R}^n$ and has the domain $\operatorname{dom}\Omega$.

(c) $\bar{x}$ is a solution of the GNEP if and only if it is a solution of the unconstrained optimization problem

$$\min_{x\in\mathbb{R}^n} [f_\alpha(x) - h_\alpha(x)]$$

with optimal function value equal to zero.

Note that the previous result still holds for certain classes of nonconvex cost functions θ_ν, i.e. for functions not satisfying Assumption 3.3 (b). This follows directly from the definitions of f_α and h_α since these functions may become strongly convex even for nonconvex θ_ν by adding a suitable strongly convex term. For example, for quadratic cost functions θ_ν, it is possible by adding of the strongly convex quadratic term $\frac{\alpha}{2}\|x\|^2$ with a sufficiently large parameter α. This observation will be exploited in our numerical section in order to compute a suitable parameter α.

Before we apply the duality theory by Toland and Singer to this DC decomposition, we consider the required conjugate functions of f_α and h_α in the next two results.

Lemma 3.10 *Let Assumption 3.3 hold. Then, the following statements hold for the conjugate f_α^* of f_α:*

(a) f_α^* *is given by*

$$f_\alpha^*(y) = x_\alpha^{f^*}(y)^T y - \frac{\alpha}{2}\left\|x_\alpha^{f^*}(y)\right\|^2 - \sum_{\nu=1}^N \theta_\nu\big(x_\alpha^{f^*}(y)\big)$$

where $x_\alpha^{f^*}(y)$ *denotes the unique solution of the maximization problem*

$$\max_x \left[x^T y - \frac{\alpha}{2}\|x\|^2 - \sum_{\nu=1}^N \theta_\nu(x)\right] \quad s.t. \quad x \in W.$$

(b) f_α^* *has the domain* $\operatorname{dom} f_\alpha^* = \mathbb{R}^n$.

(c) f_α^* *is differentiable with Lipschitz gradient given by* $\nabla f_\alpha^*(y) = x_\alpha^{f^*}(y)$.

Proof. Application of Definition 2.3 (e) leads to

$$f_\alpha^*(y) = \sup_{x\in W}\left[x^T y - \frac{\alpha}{2}\|x\|^2 - \sum_{\nu=1}^N \theta_\nu(x)\right] =: \sup_{x\in W} F_\alpha(x,y). \tag{20}$$

The function F_α is continuous on $\mathbb{R}^n \times \mathbb{R}^n$ and strongly concave in x for each fixed $y \in \mathbb{R}^n$. Since the set W is nonempty, closed, and convex by Lemma 3.5 (a), the maximization problem in (20) has a unique solution $x_\alpha^{f^*}(y)$ for each fixed $y \in \mathbb{R}^n$, so that $\operatorname{dom} f_\alpha^* = \mathbb{R}^n$. This proves statements (a) and (b).

Furthermore, the function F_α is continuously differentiable in the second variable for each fixed $x \in \mathbb{R}^n$, and the mapping $y \mapsto x_\alpha^{f^*}(y)$ is continuous on $\mathbb{R}^n$ by [23, Corollaries 8.1 and 9.1]. Due to Danskin's Theorem (see, e.g., [1, Chapter 4, Theorem 1.7]), the function f_α^* is continuously differentiable with $\nabla f_\alpha^*(y) = \nabla_y F_\alpha(x,y)\big|_{x=x_\alpha^{f^*}(y)} = x_\alpha^{f^*}(y)$. In view of Lemma 2.5, this gradient ∇f_α^* is even Lipschitz. This completes the proof. □

In a similar way as for the function f_α, we consider the conjugate function of h_α.

Lemma 3.11 *Let Assumption 3.3 hold. Then, the following statements hold for the conjugate h_α^* of h_α:*

(a) $h_\alpha^*(y)$ is given by

$$h_\alpha^*(y) = x_\alpha^{h^*}(y)^T y - \frac{\alpha}{2}\big\|x_\alpha^{h^*}(y)\big\|^2 - \sum_{\nu=1}^N \theta_\nu\big(z_\alpha^{h^*}(y)^\nu, x_\alpha^{h^*}(y)^{-\nu}\big) - \frac{\alpha}{2}\big\|x_\alpha^{h^*}(y) - z_\alpha^{h^*}(y)\big\|^2$$

where $\big(x_\alpha^{h^*}, z_\alpha^{h^*}\big)(y)$ is the unique solution of the maximization problem

$$\max_{(x,z)} \Big[x^T y - \frac{\alpha}{2}\|x\|^2 - \sum_{\nu=1}^N \theta_\nu(z^\nu, x^{-\nu}) - \frac{\alpha}{2}\|x-z\|^2\Big] \quad s.t. \quad (x,z) \in \operatorname{gph}\Omega.$$

(b) $h_\alpha^*(y)$ has the domain $\operatorname{dom} h_\alpha^* = \mathbb{R}^n$.

(c) $h_\alpha^*(y)$ is differentiable with Lipschitz gradient given by $\nabla h_\alpha^*(y) = x_\alpha^{h^*}(y)$.

Proof. Due to Definition 2.3 (e), we obtain

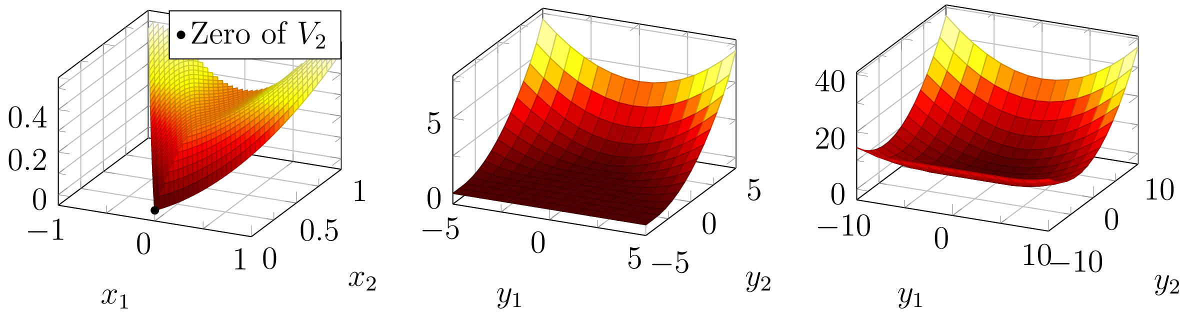
$$\begin{aligned} h_\alpha^*(y) &= \sup_{x\in\mathbb{R}^n} \Big[x^T y - \frac{\alpha}{2}\|x\|^2 - \inf_{z\in\Omega(x)} \Big(\sum_{\nu=1}^N \theta_\nu(z^\nu, x^{-\nu}) + \frac{\alpha}{2}\|x-z\|^2\Big)\Big] \\ &= \sup_{(x,z)\in\operatorname{gph}\Omega} \Big[x^T y - \frac{\alpha}{2}\|x\|^2 - \sum_{\nu=1}^N \theta_\nu(z^\nu, x^{-\nu}) - \frac{\alpha}{2}\|x-z\|^2\Big] \\ &=: \sup_{(x,z)\in\operatorname{gph}\Omega} H_\alpha(x,z,y). \end{aligned} \tag{21}$$

The function H_α is continuous on $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$, strongly concave in (x,z) for each fixed $y \in \mathbb{R}^n$, and continuously differentiable in the third variable for each fixed $(x,z) \in \mathbb{R}^n \times \mathbb{R}^n$. Since $\operatorname{gph}\Omega$ is nonempty, closed and convex by Assumption 3.3 (a) and (c), the proof of all statements of Lemma 3.11 is analogous to the proof of Lemma 3.10. □

The following simple example illustrates the two previous results.

Example 3.12 Consider a GNEP satisfying Assumption 3.3 with $N = 2$, $n_1 = n_2 = 1$, x_1 and x_2 the variables controlled by player 1 and 2, respectively, $\theta_1(x) := x_1^2$, $\theta_2(x) := x_2$ and the constraints $g_1^2(x) := x_1 - x_2 \le 0$ and $g_2^2(x) := -x_1 - x_2 \le 0$ for the second player and without constraints for the first player for simplicity. Then, we have $W = \{x \in \mathbb{R}^2 \mid |x_1| - x_2 \le 0\}$. For $\alpha = 2$, we get

$$\begin{aligned} V_2(x) &= x_1^2 + x_2 - \min_{z_1\in\mathbb{R}} \big[z_1^2 + (x_1 - z_1)^2\big] - \min_{z_2\in[|x_1|,+\infty[} \big[z_2 + (x_2 - z_2)^2\big] \\ &= \begin{cases} \frac{1}{2}x_1^2 + \frac{1}{4}, & \text{if } x_2 - \frac{1}{2} \ge |x_1|, \\ \frac{1}{2}x_1^2 + x_2 - |x_1| - (x_2 - |x_1|)^2, & \text{else.} \end{cases} \end{aligned}$$



(a) The function V_2 on W (b) The function f_2^* (c) The function h_2^*

Figure 1: Illustrations for Example 3.12

The optimal value function V_2 is nondifferentiable at $x \in \mathbb{R}^2$ with $x_1 = 0$ and $x_2 < \frac{1}{2}$ (see 'kinks' in Figure 1a) which, in particular, includes the unique solution $\bar{x} = 0$ of the considered GNEP. This solution can be verified using Lemma 3.2 (e). On the other hand, for the functions

$$f_2(x) = 2x_1^2 + x_2^2 + x_2 + \delta_W(x)$$

and

$$h_2(x) = \frac{3}{2}x_1^2 + x_2^2 + \begin{cases} x_2 - \frac{1}{4}, & \text{if } x_2 - \frac{1}{2} \geq |x_1|, \\ |x_1| + (x_2 - |x_1|)^2, & \text{else,} \end{cases}$$

we get the following continuously differentiable conjugates; see Figure 1b and 1c:

$$f_2^*(y) = \begin{cases} \frac{1}{8}\left(y_1^2 + 2(y_2 - 1)^2\right), & \text{if } y_2 > 1 + \frac{1}{2}|y_1|, \\ \frac{1}{12}(1 - |y_1| - y_2)^2, & \text{if } 1 - |y_1| < y_2 \leq 1 + \frac{1}{2}|y_1|, \\ 0, & \text{if } y_2 \leq 1 - |y_1|, \end{cases}$$

and

$$h_2^*(y) = \begin{cases} \frac{1}{12}\left(2y_1^2 + 3(y_2 - 1)^2 + 3\right), & \text{if } y_2 > 2 + \frac{2}{3}|y_1|, \\ \frac{1}{32}\left((2 - 2|y_1| - y_2)^2 + 4y_2^2\right), & \text{if } 2 - 2|y_1| < y_2 \leq 2 + \frac{2}{3}|y_1|, \\ \frac{1}{8}y_2^2, & \text{if } y_2 \leq 2 - 2|y_1|. \end{cases}$$

The continuous differentiability of both functions can be shown by simple calculations or follows directly from Lemmas 3.10 and 3.11, respectively. ◊

Finally, we obtain the main result of this section in the next theorem, by applying the duality theory by Toland and Singer [38, Theorem 2.2]. An essential finding of this duality theory is the following statement: It holds that

$$\inf_{x \in \mathbb{R}^n} [f(x) - h(x)] = \inf_{y \in \mathbb{R}^n} [h^*(y) - f^*(y)]$$

for all functions $f, h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ with h convex and lower semicontinuous.

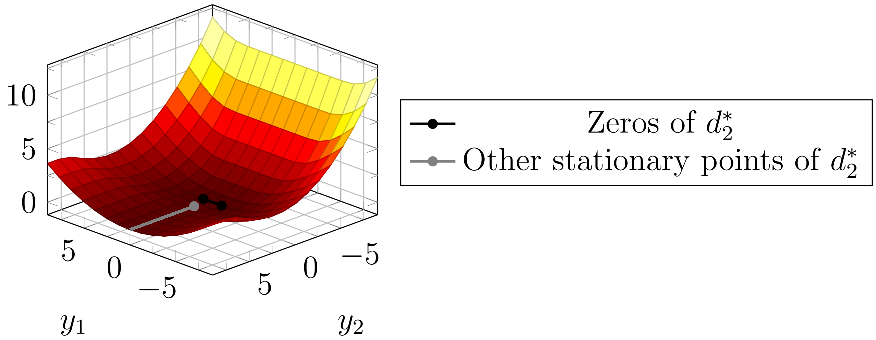


Figure 2: The dual gap function d_2^* from Example 3.14

Theorem 3.13 *Let Assumption 3.3 hold, and define the dual gap function $d_\alpha^* := h_\alpha^* - f_\alpha^*$ with the functions f_α^* and h_α^* given by Lemmas 3.10 and 3.11, respectively. Then:*

(a) The function d_α^* is continuously differentiable on $\mathbb{R}^n$.

(b) If $\bar{y}$ is a solution of the unconstrained minimization problem

$$\min \; d_\alpha^*(y), \quad y \in \mathbb{R}^n, \tag{22}$$

with $d_\alpha^*(\bar{y}) = 0$, then $\bar{x} := \nabla f_\alpha^*(\bar{y})$ is a solution of the GNEP.

(c) Conversely, if $\bar{x}$ is a solution of the GNEP and $\bar{y} \in \partial h_\alpha(\bar{x})$, then $\bar{y}$ is a solution of (22) with $d_\alpha^*(\bar{y}) = 0$.

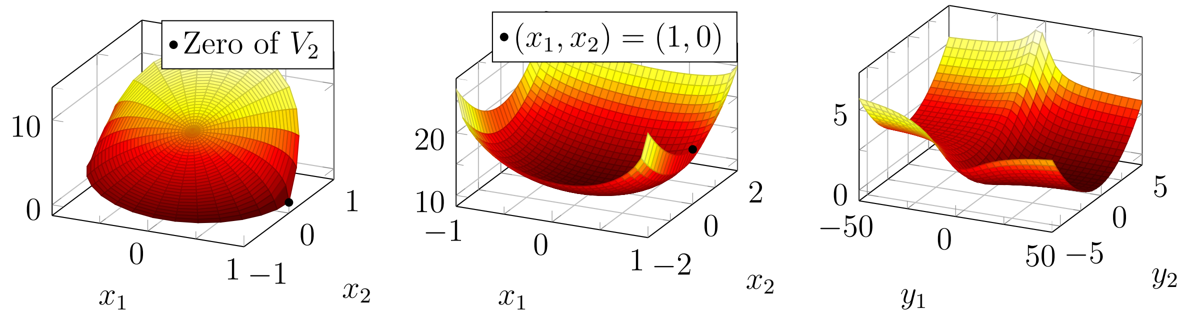
Proof. This result follows directly from the duality theory by Toland [37, 38] and Singer [35], and the details of the proof are similar to those in [17, Theorem 3.1]. □

Example 3.14 Consider the GNEP with the unique solution $\bar{x} = 0$ from Example 3.12. Since

$$h_2(x) = \frac{3}{2}x_1^2 + x_2^2 + \begin{cases} x_2 - \frac{1}{4}, & \text{if } x_2 - \frac{1}{2} \geq |x_1|, \\ x_1^2 + x_2^2 + (1 - 2x_2)|x_1|, & \text{else,} \end{cases}$$

we have $\partial h_2(0,0) = \{s \in \mathbb{R}^2 \mid s_1 \in [-1,1],\ s_2 = 0\}$. Due to Theorem 3.13 (c), all vectors $\bar{y} \in \partial h_2(0,0)$ are solutions of the dual minimization problem (22) with zero as the optimal value. Simple calculations of global minima of the dual gap function $d_2^* = h_2^* - f_2^*$ confirm this assertion; see Figure 2. Furthermore, Theorem 3.13 (b) states that $\bar{x} = \nabla f_2^*(\bar{y}) = 0$ is a solution of the GNEP. This fact was already mentioned in Example 3.12. ◊

Note that the points $\bar{y} \in \mathbb{R}^2$ with $\bar{y}_1 = 0$ and $\bar{y}_2 \geq 2$ are stationary points or local minima of the dual gap function d_2^* in Example 3.14 which are not solutions of the corresponding GNEP; see Figure 2. This example, which has fairly nice properties, points to the fact that it might be difficult to find sufficient conditions for optimality of stationary points. The following proposition is only a partial result in this direction.



(a) The function V_2 on W (b) The function h_2 (c) The dual gap function d_2^*

Figure 3: Illustrations for Example 3.16

Proposition 3.15 *Let Assumption 3.3 hold, let $d_\alpha^* = h_\alpha^* - f_\alpha^*$ be the dual gap function, and let $x_\alpha^{f^*}(y)$ and $x_\alpha^{h^*}(y), z_\alpha^{h^*}(y)$ denote the vectors defined in Lemmas 3.10 and 3.11, respectively. Then $d_\alpha^*(\bar{y}) = 0$ if and only if $x_\alpha^{f^*}(\bar{y}) = x_\alpha^{h^*}(\bar{y}) = z_\alpha^{h^*}(\bar{y})$.*

Proof. The proof is analogous to the proof of [17, Proposition 3.1]. □

Proposition 3.15 states that $x_\alpha^{f^*}(\bar{y}) = x_\alpha^{h^*}(\bar{y}) = z_\alpha^{h^*}(\bar{y}) =: \bar{x}$ implies $d_\alpha^*(\bar{y}) = 0$ and, consequently, that $\bar{x}$ is a solution of the GNEP. Since it is not difficult to see that $x_\alpha^{f^*}(\bar{y}) = x_\alpha^{h^*}(\bar{y})$ holds at any stationary point of d_α^*, it remains to provide conditions under which these two vectors are equal to $z_\alpha^{h^*}(\bar{y})$. However, we leave this question open and, therefore, also the question in which cases stationary points of the dual gap function d_α^* provide solutions of a GNEP. On the other hand, we know the optimal value of d_α^*, so this disadvantage might not be that strong, since the function value itself tells us whether we are in a solution or not. Note that, in Example 3.14, we have $x_\alpha^{f^*}(\bar{y}) = x_\alpha^{h^*}(\bar{y}) \neq z_\alpha^{h^*}(\bar{y})$ for all stationary points $\bar{y} \in \mathbb{R}^2$ with $\bar{y}_1 = 0$ and $\bar{y}_2 \geq 2$ as well as $d_\alpha^*(\bar{y}) = \frac{1}{4} \neq 0$. This function value alone shows us that none of these stationary points provides a solution of the corresponding GNEP.

Theorem 3.13 treats the relation between the solutions of the GNEP and the global minima of the dual gap function d_α^*. More precisely, it shows that every solution of the optimization problem (22) provides a solution of the GNEP, but the converse is not necessarily true, because statement (c) of Theorem 3.13 assumes (implicitly) that the subdifferential $\partial h_\alpha(\bar{x})$ is nonempty. In fact, this subdifferential could be empty, and the global minimum of the function d_α^* could be non-existent although the corresponding GNEP is solvable. The next example illustrates this assertion.

Example 3.16 Consider a GNEP satisfying Assumption 3.3 with $N = 2$, $n_1 = n_2 = 1$, $\theta_1(x) := (x_1 - 1)^2$, $\theta_2(x) := (x_2 + 4)^2$ and a constraint $g_1^2(x) := x_1^2 + x_2^2 - 1 \leq 0$ for the second player and without constraints for the first player for simplicity. Then, we have $W = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$. For $\alpha = 2$, we get

$$V_2(x) \;=\; (x_1 - 1)^2 + (x_2 + 4)^2 - \min_{z_1 \in \mathbb{R}} \left[(z_1 - 1)^2 + (x_1 - z_1)^2\right] +$$

$$
\begin{aligned}
& - \min_{z_2 \in \left[-\sqrt{1-x_1^2}, \sqrt{1-x_1^2}\right]} \left[(z_2+4)^2 + (x_2-z_2)^2\right] \\
= \; & \frac{1}{2}(x_1-1)^2 + (x_2+4)^2 + \\
& - \begin{cases} \frac{1}{2}(x_2+4)^2, & \text{if } x_1^2 + \frac{1}{4}(x_2-4)^2 < 1, \\ \left(4-\sqrt{1-x_1^2}\right)^2 + \left(x_2+\sqrt{1-x_1^2}\right)^2, & \text{if } \frac{1}{2}x_2 - 2 \leq -\sqrt{1-x_1^2}, \\ \left(4+\sqrt{1-x_1^2}\right)^2 + \left(x_2-\sqrt{1-x_1^2}\right)^2, & \text{if } \frac{1}{2}x_2 - 2 \geq \sqrt{1-x_1^2}, \\ \infty, & \text{if } |x_1| > 1. \end{cases}
\end{aligned}
$$

Note that, for all $x \in W$, it holds that $\frac{1}{2}x_2 - 2 \leq -\sqrt{1-x_1^2}$. The graph of the optimal value function V_2 on the set W is illustrated in Figure 3a. Furthermore, the functions

$$f_2(x) = x_1^2 + x_2^2 + (x_1-1)^2 + (x_2+4)^2 + \delta_W(x)$$

and

$$
\begin{aligned}
h_2(x) \;=\; & x_1^2 + x_2^2 + \frac{1}{2}(x_1-1)^2 + \\
& + \begin{cases} \frac{1}{2}(x_2+4)^2, & \text{if } x_1^2 + \frac{1}{4}(x_2-4)^2 < 1, \\ \left(4-\sqrt{1-x_1^2}\right)^2 + \left(x_2+\sqrt{1-x_1^2}\right)^2, & \text{if } \frac{1}{2}x_2 - 2 \leq -\sqrt{1-x_1^2}, \\ \left(4+\sqrt{1-x_1^2}\right)^2 + \left(x_2-\sqrt{1-x_1^2}\right)^2, & \text{if } \frac{1}{2}x_2 - 2 \geq \sqrt{1-x_1^2}, \\ \infty, & \text{if } |x_1| > 1, \end{cases}
\end{aligned}
$$

have the following conjugates:

$$f_2^*(y) = \begin{cases} \frac{1}{8}\left((y_1+2)^2 + (y_2-8)^2\right) - 17, & \text{if } (y_1+2)^2 + (y_2-8)^2 < 16, \\ \sqrt{(y_1+2)^2 + (y_2-8)^2} - 19, & \text{else,} \end{cases}$$

and

$$h_2^*(y) = \begin{cases} \frac{1}{6}\left((y_1+1)^2 + (y_2-4)^2 - 51\right), & \text{if } (y_1-1)^2 + \frac{1}{4}(y_2-16)^2 < 9, \\ \frac{1}{2}\sqrt{4(y_1-1)^2 + (y_2-16)^2} - 18 + \frac{1}{8}y_2^2, & \text{else.} \end{cases}$$

Using Lemma 3.2 (e), we obtain that the GNEP has the unique solution $(\bar{x}_1, \bar{x}_2) = (1, 0)$. At this point, the function h_2, which is illustrated in Figure 3b, has 'infinite slope', and $\partial h_2(\bar{x}) = \emptyset$. Therefore, Theorem 3.13 is not applicable to determine a solution of the corresponding dual problem (22). Furthermore, the function $d_2^* = h_2^* - f_2^*$ is positive on $\mathbb{R}^2$, and it holds that $\lim_{y_1 \to \infty} d_2^*(y_1, 0) = 0$; see Figure 3c. Thus, the dual problem (22) has the infimum zero, but does not attain its infimum, hence it has no solution. $\Diamond$

4 Second-Order Properties of the Dual Gap Function

In this section, we show piecewise smoothness of the gradient mapping ∇d_α^* under certain conditions on the strategy-sets and the cost functions of the GNEP. This piecewise smoothness result follows from our parametric optimization result stated in Theorem 2.11. To this end, we use the following assumptions.

Assumption 4.1 *(a) The strategy-space mappings $X_\nu : \mathbb{R}^{n-n_\nu} \to \mathbb{R}^{n_\nu}$ are given by*

$$X_\nu(x^{-\nu}) := \left\{x^\nu \in \mathbb{R}^{n_\nu} \mid g_i^\nu(x^\nu, x^{-\nu}) \le 0 \quad \forall i = 1, \dots, m_\nu\right\} \quad (\nu = 1, \dots, N) \quad (23)$$

with convex and twice continuously differentiable functions $g_i^\nu : \mathbb{R}^n \to \mathbb{R}$ ($\nu = 1, \dots, N,\ i = 1, \dots, m_\nu$).

(b) The feasible set W of the GNEP (7) defined in (12) is nonempty.

(c) The cost functions θ_ν, $\nu = 1, \dots, N$, are convex and twice continuously differentiable.

Then, in particular, we get $W = \{x \in \mathbb{R}^n \mid g_i^\nu(x^\nu, x^{-\nu}) \le 0 \quad \forall i = 1, \dots, m_\nu,\ \nu = 1, \dots, N\}$ and $\operatorname{gph}\Omega = \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^n \mid g_i^\nu(z^\nu, x^{-\nu}) \le 0 \quad \forall i = 1, \dots, m_\nu,\ \nu = 1, \dots, N\}$. We start our analysis by showing that the gradient of the conjugate function f_α^* is piecewise smooth under the respective CRCQ assumptions.

Lemma 4.2 *Let Assumption 4.1 hold, and let $\bar{y} \in \mathbb{R}^n$ such that CRCQ holds at $\bar{x} := x_\alpha^{f^*}(\bar{y})$ with respect to the feasible set W. Then, there exists a neighborhood V of $\bar{y}$ such that $\nabla f_\alpha^* = x_\alpha^{f^*}$ is piecewise smooth on V.*

Proof. We define $\phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by $\phi(x, y) := \frac{\alpha}{2}\|x\|^2 + \sum_{\nu=1}^N \theta_\nu(x) - x^T y$. Then, ϕ is in particular strongly convex in x and C^2. Moreover, with $p := \sum_{\nu=1}^N m_\nu$ we define $c : \mathbb{R}^n \to \mathbb{R}^p$ by $c(x) := (g_i^\nu(x))_{(i=1,\dots,m_\nu,\, \nu=1,\dots,N)}$. Then, each function c_j ($j = 1, \dots, p$) is convex and C^2 by assumption, and the assertion follows from Theorem 2.11. □

We get a similar result for the gradient of the conjugate function h_α^*.

Lemma 4.3 *Let Assumption 4.1 hold, and let $\bar{y} \in \mathbb{R}^n$ be given such that CRCQ holds at $(\bar{x}, \bar{z}) := (x_\alpha^{h^*}(\bar{y}), z_\alpha^{h^*}(\bar{y}))$ with respect to $\operatorname{gph}\Omega$. Then, there exists a neighborhood V of $\bar{y}$ such that $\nabla h_\alpha^* = x_\alpha^{h^*}$ is piecewise smooth on V.*

Proof. The assertion follows from Theorem 2.11 similar to the proof of Lemma 4.2. □

The following theorem is the main result of this section and an immediate consequence of two foregoing lemmas.

Theorem 4.4 *Let the assumptions of Lemmas 4.2 and 4.3 hold at $\bar{y} \in \mathbb{R}^n$. Then, the function ∇d_α^* is PC^1 near $\bar{y}$.*

Proof. The proof follows immediately from Lemma 4.2 and 4.3 together with the fact that $\nabla d_\alpha^* = \nabla h_\alpha^* - \nabla f_\alpha^*$. □

Corollary 4.5 *Let Assumption 4.1 hold, and let the functions g_i^ν ($\nu = 1, \dots, N,\ i = 1, \dots, m_\nu$) from (23) be affine-linear. Then, the function ∇d_α^* is PC^1 on $\mathbb{R}^n$.*

5 Numerical Results

Theorem 3.13 motivates to tackle a GNEP by solving the corresponding dual unconstrained minimization problem (22). The objective function of this minimization problem is, however, relatively expensive to calculate. On the other hand, our previous results show that each function evaluation automatically also provides the gradient. We therefore use the spectral gradient (SG) method from [31]. It has the advantage that only first order information is required and that, typically, no line search with extra function evaluations are needed. The method is defined by $y^{k+1} := y^k - t_k \nabla d_\alpha^*(y^k)$ with

$$t_0 := 1, \quad t_k := \frac{\|q^{k-1}\|^2}{(q^{k-1})^T r^{k-1}}, \quad q^{k-1} := y^k - y^{k-1}, \quad r^{k-1} := \nabla d_\alpha^*(y^k) - \nabla d_\alpha^*(y^{k-1})$$

if t_k satisfies a nonmonotone line search condition from [31]. We terminate the iteration if either $\|\nabla d_\alpha^*(y^k)\| \leq 10^{-6}$ or $d_\alpha^*(y^k) \leq 10^{-6}$ holds.

For the computation of the conjugate functions of f_α and h_α from Lemmas 3.10 and 3.11, respectively, we use the TOMLAB/SNOPT solver with settings `Prob.SOL.optPar(9)`$=10^{-8}$, `Prob.SOL.optPar(11)`$=10^{-8}$ and `Prob.SOL.optPar(12)`$=10^{-8}$, see the TOMLAB/SNOPT User's Guide on the web site `http://tomopt.com/tomlab/products/snopt/` for more information about the TOMLAB/SNOPT solver.

The test problems used here are: Examples 3.12 and 3.16 from Section 3, a class of test examples indicated by a capital T which are GNEP reformulations of a discrete approximation of a transportation problem defined as a generalized quasi-variational inequality problem in [34], as well as a subset of test problems from the report version [10] of the paper [11], indicated by a capital A. All these test examples satisfy Assumptions 3.3 (a) and (c), whereas the requirement (b) of this assumption is violated except for Examples 3.12, 3.16, A8, and A11. For our method to work also on the remaining examples, we used the strategy for the choice of the parameter α outlined after the statement of Lemma 3.9.

More precisely, in our implementation, whenever possible, we first choose, for each example, the parameter α as the smallest integer such that the minimal eigenvalues $\lambda_{min,1}$ and $\lambda_{min,2}$ of the Hessians $\nabla_x^2(-F_\alpha)$ and $\nabla_{(x,z)}^2(-H_\alpha)$ with the functions F_α and H_α defined in (20) and (21), respectively, are larger than 0.5. This choice of α guarantees that the functions f_α and h_α have all the desired properties. Such a suitable choice was easily possible for the two transportation problems T1 and T2 as well as for all test problems from [10] with quadratic cost functions, whereas the other test problems from that collection were excluded from our test set. Note that, without this particular choice of α, we usually get much worse results and often do not even converge to a solution.

The numerical results obtained with the SG method are summarized in Table 1. This table contains the following data: The name of the example, the number of players N, the number of variables n, the value of the chosen parameter α, the eigenvalues $\lambda_{min,1}$ and $\lambda_{min,2}$ of the corresponding Hessians $\nabla_x^2(-F_\alpha)$ and $\nabla_{(x,z)}^2(-H_\alpha)$, respectively, the starting point y^0, the number of iterations k, the cumulated number of dual gap function evaluations $\#d_\alpha^*$ until termination, the final value of the dual gap function $d_\alpha^*(y^k)$, and the final value of the gradient norm $\|\nabla d_\alpha^*(y^k)\|$.

Example	N	n	α	$\lambda_{min,1}$	$\lambda_{min,2}$	y^0	k	$\#d_\alpha$	$d^*_\alpha(y^k)$	$\|\nabla d^*_\alpha(y^k)\|$
Ex. 3.12	2	2	2	2.00	0.76	$(0,\dots,0)$	0	1	0.0000e+00	0.0000e+00
						$(1,\dots,1)$	5	6	1.0856e-07	1.9022e-04
						$(10,\dots,10)$	3	4	2.5000e-01	7.9040e-07
Ex. 3.16	2	2	1	3.00	1.38	$(0,\dots,0)$	83	141	1.7000e+01	9.9880e-07
						$(1,\dots,1)$	64	106	1.7000e+01	9.9216e-07
						$(10,\dots,10)$	64	108	1.7000e+01	9.3826e-07
Ex. A3	3	7	63	13.86	0.83	$(0,\dots,0)$	970	1377	7.4902e-08	1.0444e-05
						$(1,\dots,1)$	1973	2809	4.0346e-07	5.5712e-05
						$(10,\dots,10)$	1499	2133	4.7735e-07	3.3865e-04
Ex. A5	3	7	14	4.76	0.61	$(0,\dots,0)$	20	22	6.7906e-07	1.7271e-04
						$(1,\dots,1)$	17	18	7.0005e-07	1.1031e-04
						$(10,\dots,10)$	32	35	3.3339e-07	1.7712e-04
Ex. A7	4	20	117	4.85	0.63	$(0,\dots,0)$	15	16	5.0958e-08	2.1327e-05
						$(1,\dots,1)$	15	16	5.2761e-08	2.1706e-05
						$(10,\dots,10)$	15	16	6.0704e-08	2.3002e-05
Ex. A8	3	3	2	2.00	0.76	$(0,\dots,0)$	11	12	1.4745e-07	1.3367e-04
						$(1,\dots,1)$	5	6	4.7894e-07	2.8679e-04
						$(10,\dots,10)$	9	10	1.2500e-01	3.3832e-07
Ex. A11	3	3	1	3.00	1.38	$(0,\dots,0)$	4	5	4.9960e-15	3.7652e-08
						$(1,\dots,1)$	0	1	0.0000e+00	5.6501e-10
						$(10,\dots,10)$	2	3	1.4334e-11	2.0243e-06
Ex. A12	2	2	2	2.00	1.00	$(0,\dots,0)$	2	3	1.8645e-11	2.4924e-06
						$(1,\dots,1)$	2	3	2.4218e-10	6.9597e-06
						$(10,\dots,10)$	2	3	3.9645e-10	8.9048e-06
Ex. A13	3	3	2	2.02	0.78	$(0,\dots,0)$	13	15	9.9189e-07	2.0379e-04
						$(1,\dots,1)$	14	16	2.2737e-13	4.6215e-07
						$(10,\dots,10)$	20	25	-2.2737e-13	5.7849e-07
Ex. A15	3	6	3	3.02	0.60	$(0,\dots,0)$	1094	1433	9.9879e-07	1.0102e-05
						$(1,\dots,1)$	1099	1470	9.1871e-07	1.1981e-04
						$(10,\dots,10)$	1443	1922	3.3720e-07	3.7221e-04
Ex. A17	2	3	2	1.63	0.79	$(0,\dots,0)$	7	8	8.9804e-09	3.8128e-05
						$(1,\dots,1)$	8	9	0.0000e+00	2.7595e-08
						$(10,\dots,10)$	8	9	2.2402e-07	1.9045e-04
Ex. A18	2	12	2	2.00	0.76	$(0,\dots,0)$	23	25	9.3132e-10	2.6248e-06
						$(1,\dots,1)$	36	37	2.9153e-07	2.3941e-04
						$(10,\dots,10)$	27	32	1.4095e-08	4.8729e-05
Ex. T1	2	2400	2	2.76	0.57	$(0,\dots,0)$	18	19	3.2072e-07	3.0320e-04
						$(1,\dots,1)$	20	21	3.0623e-07	1.9514e-04
						$(10,\dots,10)$	21	22	3.3050e-07	2.2579e-04
Ex. T2	2	4800	2	2.76	0.57	$(0,\dots,0)$	19	20	1.7288e-08	5.3392e-05
						$(1,\dots,1)$	20	21	2.3871e-07	2.5285e-04
						$(10,\dots,10)$	21	22	6.2399e-07	2.9929e-04

Table 1: Numerical results with the spectral gradient method

The calculations with the SG method were quite successful for most instances, except for Example 3.16 with all starting points and Examples 3.12 and A8 with the third starting point, where the function value is not small enough and the iteration is terminated since the norm of the gradient gets small, hence we are close to a non-optimal stationary point. Note that the failure in Example 3.16 was to be expected based on the considerations in this example! Moreover, there are two cases, namely Example 3.12 with first starting point and Example A11 with second starting point, where the starting point already provides a solution of the dual unconstrained minimization problem (22). Furthermore, the bad convergence speed in Examples A3 and A15 leads to large number of iterations although the calculations of solutions for all starting points were successful. In all other test examples, we observed far better convergence properties. The iteration is terminated after a relatively small number of iterations, in particular, taking into account that the SG method is just a first-order gradient method.

6 Final Remarks

This paper shows that a class of generalized Nash equilibrium problems can be reformulated, using some results from variational and convex analysis, as an unconstrained and smooth optimization problem. There are a couple of questions still open for future research. First, when is a stationary point of our unconstrained optimization problem already a global minimum and, therefore, a solution of the GNEP? Second, can we develop a second-order method with fast local convergence by exploiting the fact that the gradient is still piecewise smooth? Third, what happens for the case where the function evaluations are done only inexactly? The last point is quite interesting from a practical perspective since the evaluation of our unconstrained objective function requires the solution of suitable optimization problems which are strongly convex, but which might be difficult to compute exactly at least in the non-quadratic case.

References

[1] A. Auslender: *Optimisation: Méthodes Numériques.* Masson, Paris, 1976.

[2] D. Aussel, R. Correa, M. Marechal: *Gap functions for quasivariational inequalities and generalized Nash equilibrium problems.* J. Optim. Theory Appl. 151, 2011, pp. 474–488.

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

[4] S. Dempe, D. Pallaschke: *Quasidifferentiability of optimal solutions in parametric nonlinear optimization.* Optimization 40, 1997, pp. 1–24.

[5] H. Dietrich: *A smooth dual gap function solution to a class of quasivariational inequalities.* J. Math. Anal. Appl. 235, 1999, pp. 380–393.

[6] H. Dietrich: *Optimal control problems for certain quasivariational inequalities.* Optimization 49, 2001, pp. 67–93.

[7] A. Dreves, C. Kanzow: *Nonsmooth optimization reformulations characterizing all solutions of jointly convex generalized Nash equilibrium problems.* Comput. Optim. Appl. 50, 2011, pp. 23–48.

[8] A. Dreves, C. Kanzow, O. Stein: *Nonsmooth optimization reformulations of player convex generalized Nash equilibrium problems.* J. Global Optim. 53, 2012, pp. 587–614.

[9] F. Facchinei, C. Kanzow: *Generalized Nash equilibrium problems.* Ann. Oper. Res. 175, 2010, pp. 177–211.

[10] 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.

[11] F. Facchinei, C. Kanzow: *Penalty methods for the solution of generalized Nash equilibrium problems.* SIAM J. Optim. 20, 2010, pp. 2228–2253.

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

[13] M. Fukushima: *A class of gap functions for quasi-variational inequality problems.* J. Ind. Manag. Optim. 3, 2007, pp. 165–171.

[14] F. Giannessi: *Separation of sets and gap functions for quasi-variational inequalities.* In F. Giannessi, A. Maugheri (eds.): *Variational Inequality and Network Equilibrium Problems.* Plenum Press, New York, 1995, pp. 101–121.

[15] G. Gürkan, J.-S. Pang: *Approximations of Nash equilibria.* Math. Program. 117, 2009, pp. 223–253.

[16] P.T. Harker: *Generalized Nash games and quasi-variational inequalities.* Eur. J. Oper. Res. 54, 1991, pp. 81–94.

[17] N. Harms, T. Hoheisel, C. Kanzow: *On a smooth dual gap function for a class of quasi-variational inequalities.* J. Optim. Theory Appl., to appear, DOI: 10.1007/s10957-014-0536-4.

[18] N. Harms, C. Kanzow, O. Stein: *On differentiability properties of player convex generalized Nash equilibrium problems.* Optimization, to appear, DOI: 10.1080/02331934.2012.752822.

[19] N. Harms, C. Kanzow, O. Stein: *Smoothness properties of a regularized gap function for quasi-variational inequalities.* Optim. Method. Softw. 29, 2014, pp. 720–750.

[20] A. von Heusinger, C. Kanzow: *Optimization reformulations of the generalized Nash equilibrium problem using Nikaido-Isoda-type functions.* Comput. Optim. Appl. 43, 2009, pp. 353–377.

[21] A. von Heusinger, C. Kanzow, M. Fukushima: *Newton's method for computing a normalized equilibrium in the generalized Nash game through fixed point formulation.* Math. Program. 132, 2012, pp. 99-123.

[22] J.-B. Hiriart-Urruty, C. Lemaréchal: *Fundamentals of Convex Analysis.* Springer, Berlin, 2001.

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

[24] T. Horst, N.V. Thoai: *DC programming: Overview.* J. Optim. Theory Appl. 103, 1999, pp. 1–43.

[25] R. Janin: *Directional derivative of the marginal function in nonlinear programming.* Math. Program. Stud. 21, 1984, pp. 110–126.

[26] J.B. Krawczyk, S. Uryasev: *Relaxation algorithms to find Nash equilibria with economic applications.* Environ. Model. Assess. 5, 2000, pp. 63–73.

[27] K. Kubota, M. Fukushima: *Gap function approach to the generalized Nash equilibrium problem.* J. Optim. Theory Appl. 144, 2010, pp. 511–531.

[28] H. Nikaido, K. Isoda: *Note on noncooperative convex games.* Pac. J. Math. 5, 1955, pp. 807–815.

[29] J.S. Pang, D. Ralph: *Piecewise smoothness, local invertibility, and parametric analysis of normal maps.* Math. Oper. Res. 21, 1996, pp. 401–426.

[30] D. Ralph, S. Dempe: *Directional derivatives of the solution of a parametric nonlinear program.* Math. Program. 70, 1995, pp. 159–172.

[31] M. Raydan: *The Barzilai and Borwein gradient method for the large unconstrained minimization problem.* SIAM J. Optim. 7, 1997, pp. 26–33.

[32] R.T. Rockafellar, R.J.-B. Wets: *Variational Analysis.* Springer, Berlin, 1998.

[33] S. Scholtes: *Introduction to Piecewise Differentiable Equations.* Springer, New York, 2012.

[34] L. Scrimali: *Mixed behavior network equilibria and quasi-variational inequalities* J. Ind. Manag. Optim. 5, 2009, pp. 363–379.

[35] I. Singer: *A Fenchel-Rockafellar type duality theorem for maximization.* B. Aust. Math. Soc. 20, 1979, pp. 193–198.

[36] K. Taji: *On gap functions for quasi-variational inequalities.* Abstr. Appl. Anal., 2008, Article ID 531361.

[37] J.F. Toland: *Duality in non-convex optimization.* J. Math. Anal. Appl. 66, 1978, pp. 399–415.

[38] J.F. Toland: *A duality principle for non-convex optimisation and the calculus of variations.* Arch. Ration. Mech. An. 71, 1979, pp. 41–61.

[39] S. Uryasev, R.Y. Rubinstein: *On relaxation algorithms in computation of noncooperative equilibria.* IEEE T. Automat. Contr. 39, 1994, pp. 1263–1267.