SMOOTHNESS PROPERTIES OF A REGULARIZED GAP FUNCTION FOR QUASI-VARIATIONAL INEQUALITIES[1]

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

Preprint 313 March 2013

[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 8, 2013

[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 continuity and differentiability properties for a reformulation of a finite-dimensional quasi-variational inequality (QVI) problem using a regularized gap function approach. For a special class of QVIs, this gap function is continuously differentiable everywhere, in general, however, it has nondifferentiability points. We therefore take a closer look at these nondifferentiability points and show, in particular, that under mild assumptions all locally minimal points of the reformulation are differentiability points of the regularized gap function. The results are specialized to generalized Nash equilibrium problems. Numerical results are also included and show that the regularized gap function provides a valuable approach for the solution of QVIs.

Key Words: Finite-dimensional quasi-variational inequalities, convex inequalities, regularized gap function, Hadamard directional differentiability, Gâteaux differentiability, Fréchet differentiability, Generalized Nash equilibrium problem, Generalized moving set.

1 Introduction

This paper considers the finite-dimensional quasi-variational inequality problem, QVI for short. To this end, let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a given vector-valued function and let $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued mapping such that $S(x)$ are closed and convex (possibly empty) sets for each given $x \in \mathbb{R}^n$. Then the QVI consists of finding a solution $x \in S(x)$ such that

$$F(x)^T(y - x) \geq 0 \qquad \forall y \in S(x). \tag{1}$$

If the set $S(x)$ is independent of x, i.e. $S(x) = S$ for all $x \in \mathbb{R}^n$ with some constant set $S \subseteq \mathbb{R}^n$, then the QVI reduces to the standard variational inequality (VI) problem, cf. the monograph [15] for an extensive discussion of VIs.

In the context of QVIs, the fixed point set of S,

$$X := \{x \in \mathbb{R}^n \mid x \in S(x)\} \tag{2}$$

plays a special role and is sometimes called the *feasible set* of the QVI from (1). In case of a VI, this set is equal to the constant set S and therefore justifies this terminology. In the present paper, also the (effective) domain of S,

$$M := \operatorname{dom} S = \{x \in \mathbb{R}^n \mid S(x) \neq \emptyset\},$$

will play a central role. Clearly, the relation

$$X \subseteq M \tag{3}$$

holds.

We assume that $S(x)$ has a representation of the form

$$S(x) = \{y \in \mathbb{R}^n \mid \; s_i(x, y) \leq 0 \;\; \forall i = 1, \ldots, m\}$$

with suitable functions $s_i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$. Then the feasible set X is given by

$$X = \{x \in \mathbb{R}^n \mid s_i(x, x) \leq 0 \;\; \forall i = 1, \ldots, m\}.$$

Throughout the paper, we make the following smoothness and convexity assumptions.

Assumption 1.1 *(a) The function F is continuous on $\mathbb{R}^n$.*

(b) The functions s_i, $i = 1, \ldots, m$, are continuous on $\mathbb{R}^n \times \mathbb{R}^n$.

(c) The functions $s_i(x, \cdot)$, $i = 1, \ldots, m$, are convex for each fixed $x \in \mathbb{R}^n$.

Note that, in particular, Assumptions 1.1 (b), (c) guarantee that $S(x)$ is indeed a closed and convex (possibly empty) set for any given $x \in \mathbb{R}^n$.

The QVI was formally introduced in a series of papers [5, 6, 7] by Bensoussan et al. It has soon become a powerful modelling tool for many different problems both in the finite and in the infinite-dimensional setting. An early summary may be found in the article by Mosco [29], the infinite-dimensional problem with several mechanical and engineering problems is discussed in the monograph [4] by Baiocchi and A. Capelo. For several other applications, we refer the reader to the list of references in the recent paper [13]. In the meantime, several applications coming from totally different origins can also be found in a test problem collection whose details are given in [14].

Unfortunately, the QVI turns out to be a difficult class of problems, and the numerical solution of QVIs is still a challenging task. To the best of our knowledge, the first method was proposed by Chan and Pang [8]. They consider a projection-type algorithm and prove a global convergence result under certain assumptions for the class of QVIs where the set-valued mapping S is given by $S(x) = c(x) + K$ for a suitable function $c : \mathbb{R}^n \to \mathbb{R}^n$ and a fixed closed and convex set $K \subseteq \mathbb{R}^n$. This particular class of problems is sometimes called a QVI with a 'moving set' $S(x)$ since the fixed set K moves along the mapping $c(x)$. There are a number of subsequent extensions of this approach, see, e.g., [30, 31, 33, 40, 41], which all use a projection-type or fixed-point iteration and essentially deal with the moving set case only in order to obtain suitable global convergence results. More recently, Pang and Fukushima [37] suggested a penalty-multiplier-type approach where they have to solve a sequence of (standard) VIs. They obtain a global convergence result for a class of problems not restricted to the moving set case, but their VI-subproblems are in general non-monotone and therefore difficult to solve. A very recent method by Facchinei et al. [13] applies a potential-reduction-type method to the corresponding KKT conditions and proves global convergence results for some classes of QVIs that go beyond the moving set case. Besides these (more or less) globally convergent approaches, there also exist some locally convergent Newton-type methods by Outrata et al., see, in particular, [34, 35, 36].

Apart from the previous classes of methods, there exist a number of different gap functions for QVIs, cf. [3, 9, 17, 19, 43] and the corresponding discussion in Section 2. In principal, these gap functions allow a reformulation of the QVI as an optimization problem and therefore the application of standard software. However, the disadvantage is that these gap functions are usually nonsmooth, so that the previous literature concentrates on error bound results or the local Lipschitz continuity and directional differentiability of these gap functions.

The main focus of this paper is an in-depth treatment of the (continuous) differentiability properties of one class of (regularized) gap functions for QVIs. In particular, we identify a class of QVIs with a generalized moving set where the gap function turns out to be continuously differentiable everywhere. We also show that, except for some pathological cases, the gap function is continuously differentiable at all minimal points.

The paper is organized in the following way: In Section 2, we recall the definition of a regularized gap function for the QVI from [9, 17, 43] and restate some of its basic properties. We then discuss three special classes of QVIs in Section 3, namely QVIs with a generalization of the moving set case for which the regularized gap function turns out to be continuously differentiable, further QVIs with set-valued mappings in product form, and finally, as an important application, the generalized Nash equilibrium problem. After this, we turn back to the general QVI, where the regularized gap function is typically nonsmooth.

Hence we investigate its continuity properties in Section 4 under suitable assumptions. We then discuss the differentiability properties of the gap function in Section 5. Our main result of Section 5 is that, apart from special cases, all locally minimal points of the reformulation are differentiability points of the gap function. Some numerical results are provided in Section 6, and we conclude with some final remarks in Section 7.

The notation used in this manuscript should be rather standard. We only point out that $\nabla F(x)$ denotes the transposed Jacobian of F at x, which is consistent with our notion of the gradient $\nabla f(x)$ of a real-valued function since this gradient is viewed as a column vector. Given a function f and a set $X \subseteq \mathbb{R}^n$, we say that f is continuous at $\bar{x} \in X$ *relative to* X if $f(x^k) \to f(\bar{x})$ for all sequences $\{x^k\} \subset X$ converging to $\bar{x}$.

2 Preliminaries on Gap Functions

There exist several gap functions for QVIs. All these gap functions were originally introduced for standard VIs and then extended to QVIs. We therefore first recall the definitions of the relevant gap functions for VIs in Section 2.1 and then present their counterparts for QVIs in Section 2.2, together with some elementary properties of one of these gap functions that plays a central role in our subsequent analysis. Note that there exist other gap functions both for VIs and QVIs which, however, do not play any role in our context, see, e.g., [32].

2.1 Gap Functions for Variational Inequalities

Recall that the (standard) variational inequality consists of finding a solution $x \in S$ such that

$$F(x)^T(y-x) \geq 0 \quad \forall y \in S \tag{4}$$

holds, where $S \subseteq \mathbb{R}^n$ is a nonempty, closed, and convex set, and $F : \mathbb{R}^n \to \mathbb{R}^n$ denotes a continuously differentiable function. The classical *gap function* for VI is defined by

$$g(x) := -\inf_{y \in S} F(x)^T(y-x)$$

and was introduced by Auslender [2], see also Hearn [23] and, e.g., the paper [28] for an algorithmic application. The gap function is nonnegative on S, and $g(\bar{x}) = 0$ for some $\bar{x} \in S$ holds if and only if $\bar{x}$ solves the VI. Hence the VI is equivalent to the constrained optimization problem

$$\min \; g(x) \quad \text{s.t.} \quad x \in S \tag{5}$$

with zero as the optimal value. However, unless S is compact, the objective function g is typically extended-valued, moreover, g is usually nondifferentiable.

In order to avoid these problems, Fukushima [16] and Auchmuty [1] independently developed the *regularized gap function*

$$g_\alpha(x) := -\min_{y \in S} \left[F(x)^T(y-x) + \frac{\alpha}{2}\|y-x\|^2 \right],$$

where $\alpha > 0$ denotes a given parameter. Similar to the gap function, one can show that also the regularized gap function is nonnegative on S, and $g_\alpha(\bar{x}) = 0$ for some $\bar{x} \in S$ holds if and only if $\bar{x}$ solves the VI. Moreover, g_α is finite-valued and continuously differentiable (by Danskin's Theorem) everywhere. Hence the VI is equivalent to a smooth optimization problem of the form (5) with g being replaced by g_α. This fact has been exploited, e.g., in the paper [45] which presents a simple globalization of the standard Josephy-Newton method based on the regularized gap function.

The main computational burden of the regularized gap function is the fact that the evaluation of $g_\alpha(x)$ is quite expensive for nonlinear (non-polyhedral) sets S since then one has to solve a convex optimization problem with a nonlinear feasible set, which is practically impossible. Motivated by this observation, Taji and Fukushima [44] introduced the following modification of the regularized gap function:

$$\tilde{g}_\alpha(x) := -\min_{y \in T(x)} \left[F(x)^T (y - x) + \frac{\alpha}{2} \|y - x\|^2 \right],$$

where $T(x)$ denotes the polyhedral approximation of S at x defined by

$$T(x) := \left\{ y \mid s_i(x) + \nabla s_i(x)^T (y - x) \leq 0 \;\; \forall i = 1, \ldots, m \right\}$$

and where we assume that the feasible set S has the representation $S = \{x \mid s_i(x) \leq 0 \;\; \forall i = 1, \ldots, m\}$ for some convex functions s_i. It was shown in [44] that, once again, the VI is equivalent to a constrained optimization problem like (5) with $\tilde{g}_\alpha$ replacing g, and with zero objective function value at the solution. However, in contrast to the regularized gap function g_α, the mapping $\tilde{g}_\alpha$ is, in general, not differentiable.

2.2 Gap Functions for Quasi-Variational Inequalities

Consider the QVI from (1). A direct extension of the classical gap function from VIs to QVIs seems to be due to Giannessi [19], who defines the mapping

$$g(x) := -\inf_{y \in S(x)} F(x)^T (y - x)$$

and shows that

- $g(x) \geq 0$ for all $x \in X$;
- $g(\bar{x}) = 0$ for some $\bar{x} \in X$ if and only if $\bar{x}$ solves the QVI,

where, we recall, X denotes the feasible set of a QVI from (2). Hence the QVI is equivalent to the constrained optimization problem

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

However, the objective function g is nondifferentiable, possibly extended-valued (both $g(x) = -\infty$ and $g(x) = +\infty$ may occur if $S(x) = \emptyset$ or g is unbounded from above). Further note that the set X might have a complicated structure.

An extension of the regularized gap function to QVIs is due to Taji [43] and was, in fact, introduced earlier by Dietrich [9] for a special class of QVIs in the infinite-dimensional setting, see also the very recent paper [3] by Aussel et al. This regularized gap function for QVIs is defined by

$$g_\alpha(x) := -\min_{y \in S(x)} \left[F(x)^T (y - x) + \frac{\alpha}{2} \|y - x\|^2 \right] \tag{6}$$

where $\alpha > 0$ denotes a given parameter. In view of Assumption 1.1, the function

$$\varphi_\alpha(x, y) := F(x)^T (y - x) + \frac{\alpha}{2} \|y - x\|^2 \tag{7}$$

is strongly convex in y for each fixed $x \in \mathbb{R}^n$. We therefore have the following remark.

Remark 2.1 For any $x \in M$ (the domain of S) the minimum in (6) is uniquely attained by the solution $y_\alpha(x)$ of the optimization problem

$$\min_y \ \varphi_\alpha(x, y) \quad \text{s.t.} \quad y \in S(x). \tag{8}$$

In particular, we have $g_\alpha(x) = -\varphi_\alpha(x, y_\alpha(x)) \in \mathbb{R}$. Note, however, that $g_\alpha(x) = -\infty$ holds for $x \notin M$, so that g_α is real-valued exactly on M. Consequently, due to (3), g_α is real-valued on X. $\Diamond$

The following result, whose proof may be found in [43], clarifies the relation between the regularized gap function g_α and the QVI (1) (recall once again that the set X in this result denotes the feasible set from (2)).

Proposition 2.2 *For all $x \in X$, we have $g_\alpha(x) \geq 0$. Moreover, $\bar{x}$ solves the QVI if and only if $g_\alpha(\bar{x}) = 0$ and $\bar{x} \in X$.*

Proposition 2.2 shows that the QVI is equivalent to finding an optimal point $\bar{x}$ of

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

with $g_\alpha(\bar{x}) = 0$. Unfortunately, and in contrast to standard VIs, simple examples show that the objective function of this problem is nondifferentiable in general, and for infeasible points $x \notin X$, it might also take the value $-\infty$ (compare Remark 2.1).

Based on this observation, it seems natural to replace g_α by the counterpart of the modified regularized gap function $\tilde{g}_\alpha$ from the previous subsection. In fact, this was done by Fukushima [17], but we skip the corresponding details here, mainly because it turns out that the regularized gap function has better differentiability properties. In fact, in an important special case to be discussed in the following section, the regularized gap function from (6) turns out to be smooth, whereas the modified regularized gap function from [17] would still be nonsmooth in general.

To conclude this section, we introduce an example which not only illustrates Proposition 2.2, but will also serve to illustrate continuity and differentiability properties of g_α on X in Sections 4 and 5, respectively.

Example 2.3 Consider the QVI with $n = 2$, $F(x) = (1,1)^T$, and $S(x) = \{y \in \mathbb{R}^2 |\, s_i(x,y) \leq 0,\ i \in \{1,2,3\}\}$, where

$$s_1(x,y) = -2y_1 + x_2, \quad s_2(x,y) = x_1^2 + y_2^2 - 1, \quad s_3(x,y) = -x_1 - y_2.$$

Then Assumption 1.1 is satisfied, and we have $S(x) = S_1(x) \times S_2(x)$ with

$$\begin{aligned} S_1(x) &= \{y_1 \in \mathbb{R} |\, -2y_1 + x_2 \leq 0\} = \left[\frac{x_2}{2}, +\infty\right), \\ S_2(x) &= \{y_2 \in \mathbb{R} |\, x_1^2 + y_2^2 - 1 \leq 0,\ -x_1 - y_2 \leq 0\} \\ &= \left[\max\left\{-x_1, -\sqrt{1-x_1^2}\right\}, \sqrt{1-x_1^2}\right], \end{aligned}$$

so that $M = [-1/\sqrt{2}, 1] \times \mathbb{R}$ and

$$X = \{x \in \mathbb{R}^2 |\, -2x_1 + x_2 \leq 0,\ x_1^2 + x_2^2 - 1 \leq 0,\ -x_1 - x_2 \leq 0\},$$

see Fig. 1. For the regularized gap function with $\alpha > 0$ we obtain

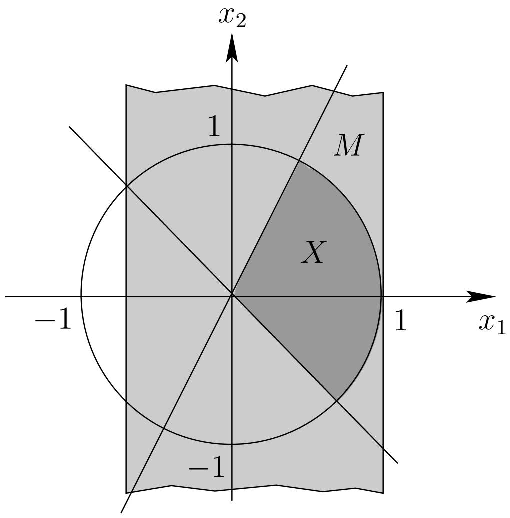

Figure 1: Illustration of the sets X and M in Example 2.3

$$\begin{aligned} g_\alpha(x) &= -\min_{y \in S(x)} \left[F(x)^T(y-x) + \frac{\alpha}{2}\|y-x\|^2\right] \\ &= x_1 + x_2 - \min_{y_1 \in S_1(x)} \left(y_1 + \frac{\alpha}{2}(y_1 - x_1)^2\right) - \min_{y_2 \in S_2(x)} \left(y_2 + \frac{\alpha}{2}(y_2 - x_2)^2\right), \end{aligned} \tag{9}$$

and for $x \in M$ the two components of $y_\alpha(x)$ are the unique optimal points corresponding to the two optimal values in (9). In fact, with

$$\varrho_1(x) := x_1 - \frac{x_2}{2}, \quad \varrho_2(x) := x_2 + \min\left\{x_1, \sqrt{1-x_1^2}\right\}, \quad \varrho_3(x) := x_2 - \sqrt{1-x_1^2},$$

we have

$$(y_\alpha(x))_1 = \begin{cases} x_1 - \varrho_1(x), & \text{if } \varrho_1(x) \le \frac{1}{\alpha}, \\ x_1 - \frac{1}{\alpha}, & \text{if } \frac{1}{\alpha} < \varrho_1(x), \end{cases}$$

$$(y_\alpha(x))_2 = \begin{cases} x_2 - \varrho_2(x), & \text{if } \varrho_2(x) \le \frac{1}{\alpha}, \\ x_2 - \frac{1}{\alpha}, & \text{if } \varrho_3(x) < \frac{1}{\alpha} < \varrho_2(x), \\ x_2 - \varrho_3(x), & \text{if } \frac{1}{\alpha} \le \varrho_3(x). \end{cases}$$

Using the corresponding indicator functions

$$\mathbb{1}_{\{1/\alpha < \varrho_1(x)\}}(x) = \begin{cases} 1, & \text{if } 1/\alpha < \varrho_1(x), \\ 0, & \text{else}, \end{cases}$$

etc., and (9), this results in

$$\begin{aligned} g_\alpha(x) &= \frac{1}{2\alpha}\Big(\mathbb{1}_{\{1/\alpha<\varrho_1(x)\}}(x) + \mathbb{1}_{\{\varrho_3(x)<1/\alpha<\varrho_2(x)\}}(x)\Big) + \Big(\varrho_1(x) - \frac{\alpha}{2}\varrho_1^2(x)\Big)\,\mathbb{1}_{\{\varrho_1(x)\le 1/\alpha\}}(x) \\ &\quad + \Big(\varrho_2(x) - \frac{\alpha}{2}\varrho_2^2(x)\Big)\,\mathbb{1}_{\{\varrho_2(x)\le 1/\alpha\}}(x) + \Big(\varrho_3(x) - \frac{\alpha}{2}\varrho_3^2(x)\Big)\,\mathbb{1}_{\{1/\alpha\le\varrho_3(x)\}}(x). \end{aligned}$$

Figure 2 shows the graph of the regularized gap function on the set X for $\alpha = 1$. One can

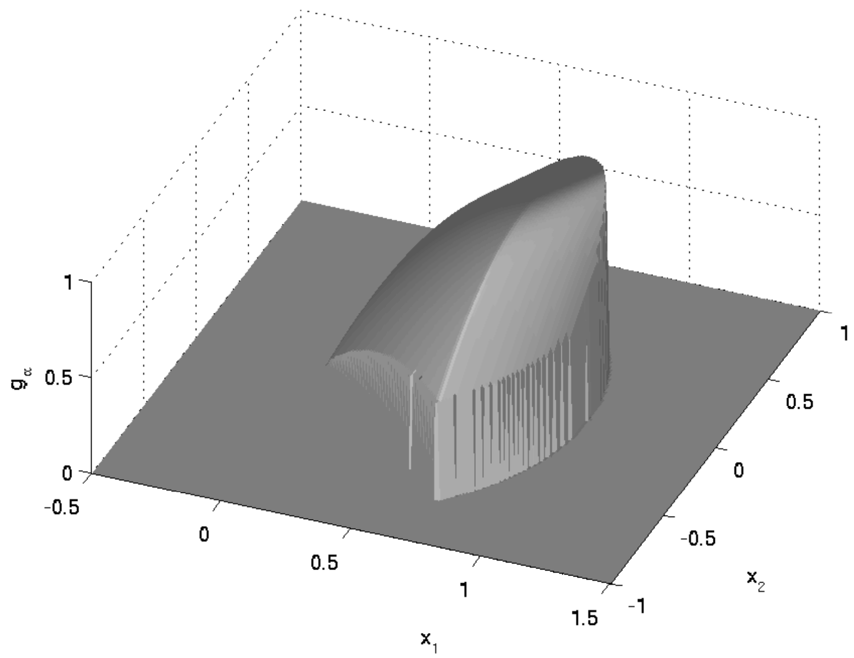

Figure 2: The regularized gap function for $\alpha = 1$ in Example 2.3

show that, indeed, $\bar{x} = 0$ is the unique globally minimal point of g_α on X with value zero so that, by Proposition 2.2, $\bar{x} = 0$ is the unique solution of the QVI. $\Diamond$

3 Special Classes of QVIs

Here we consider three special classes of QVIs. The first is a generalization of QVIs with 'moving sets', the second are QVIs with set-valued mappings in product form, and the third are generalized Nash equilibrium problems.

3.1 QVIs with Generalized Moving Sets

Many papers dealing with QVIs do not consider the general setting from (1), see, e.g., [8, 9, 30]. They only discuss the particular case where the set $S(x)$ has the form $S(x) = c(x) + K$ for some function $c : \mathbb{R}^n \to \mathbb{R}^n$ and a fixed nonempty, closed, and convex set $K \subseteq \mathbb{R}^n$. This class of QVIs is often called the 'moving set case' for reasons that should be clear from the left picture in Figure 3.

Here we consider a generalization of this case. To this end, let c be given as before and, for $p \leq n$, let $K \subseteq \mathbb{R}^p$ be a nonempty, closed, and convex set. In addition, assume that we have a matrix $Q(x) \in \mathbb{R}^{n\times p}$ of full (column) rank for all $x \in \mathbb{R}^n$. Then we consider the case where the set-valued mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ has the form

$$S(x) = c(x) + Q(x)K := \{c(x) + Q(x)z \mid z \in K\}. \tag{10}$$

Note that $S(x) \neq \emptyset$ holds in this case for any $x \in \mathbb{R}^n$, that is, we have $M = \mathbb{R}^n$. We call a QVI with the mapping S defined in this way the 'generalized moving set case'. In the special case $p = n$ and $Q(x) = I$ for all $x \in \mathbb{R}^n$ we re-obtain the 'moving set case'. Our generalization of this case for $p = n$ actually allows any $x-$dependent affine transformation $T(x, K) = c(x) + Q(x)K$ of K instead of just translation, that is, also scaling, rotation, reflection, and shearing, as shown in the right picture of Figure 3 for $p = n = 2$. For a further generalization of this approach see Remark 3.5 below.

With the exception of the recent paper [13], the QVIs with moving sets are essentially the only case that have been investigated in papers dealing with the numerical solution of QVIs, and for which a more or less complete convergence theory is available. For example, Dietrich [9] considers QVIs with moving sets only and notes that the regularized gap function is continuously differentiable in this case. It seems that this observation has been widely overlooked in the subsequent literature.

In this subsection, we want to generalize this observation by showing that the regularized gap function g_α from (6) is still smooth in the case where the set $S(x)$ is given by (10) with continuously differentiable functions c and Q. To this end, we first reformulate the minimization problem from (8) as

$$\begin{aligned}
&\min_y \ \varphi_\alpha(x, y) \quad \text{s.t.} \quad y \in S(x) \\
&\iff \min_y \ \varphi_\alpha(x, y) \quad \text{s.t.} \quad y \in c(x) + Q(x)K \\
&\iff \min_y \ \varphi_\alpha(x, y) \quad \text{s.t.} \quad \exists\, z \in K : \ y = c(x) + Q(x)z \\
&\iff \min_{y,z} \ \varphi_\alpha(x, y) \quad \text{s.t.} \quad y = c(x) + Q(x)z, \ z \in K
\end{aligned}$$

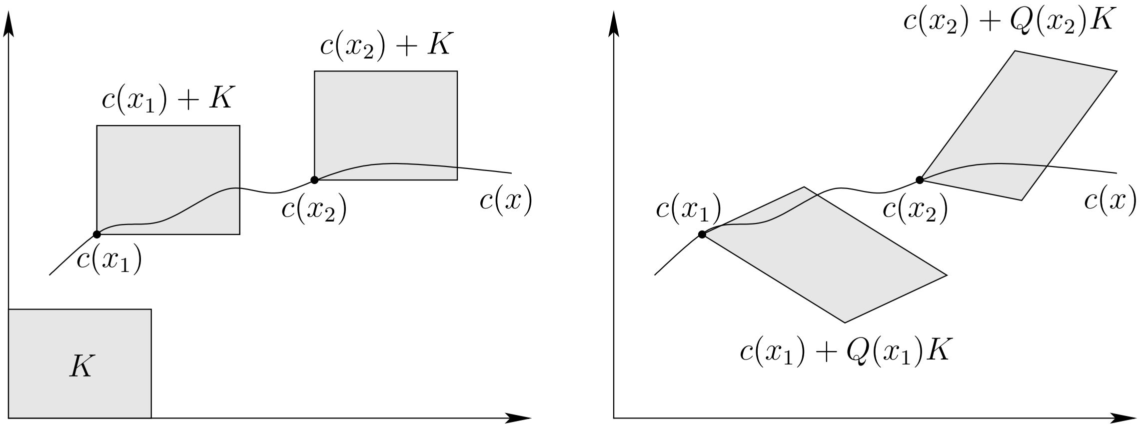

Figure 3: Examples for a 'moving set' (left) and a 'generalized moving set' (right)

$$\iff \quad \min_{z} \; \psi_\alpha(x,z) \quad \text{s.t.} \quad z \in K, \tag{11}$$

where

$$\begin{aligned} \psi_\alpha(x,z) \;&:=\; \varphi_\alpha\big(x, c(x) + Q(x)z\big) = F(x)^T(c(x) - x) + \frac{\alpha}{2}\|c(x) - x\|^2 \\ &\quad + \big(F(x) + \alpha(c(x) - x)\big)^T Q(x)z + \frac{\alpha}{2} z^T Q(x)^T Q(x) z \end{aligned}$$

is convex quadratic in z for each x. Note that the full rank of $Q(x)$ is actually not needed for the reformulation (11), but that under this assumption, for each fixed $x \in \mathbb{R}^n$, the function $\psi_\alpha(x,\cdot)$ is strongly convex with respect to z because $\nabla^2_{zz}\psi_\alpha(x,z) = \alpha Q(x)^T Q(x)$ is uniformly positive definite (in z). Therefore, problem (11) has a unique solution $z_\alpha(x)$ for all $x \in \mathbb{R}^n$, and we obtain

$$g_\alpha(x) = -\min_{y \in S(x)} \varphi_\alpha(x,y) = -\min_{z \in K} \psi_\alpha(x,z) = -\psi_\alpha(x, z_\alpha(x)).$$

The function $x \mapsto z_\alpha(x)$ turns out to be continuous on $\mathbb{R}^n$. Before we prove this assertion, we recall some definitions and results from set-valued analysis.

Definition 3.1 *Let $X \subseteq \mathbb{R}^n$, $Y \subseteq \mathbb{R}^p$, and $\Phi : X \rightrightarrows Y$ be a set-valued mapping. Then Φ is called*

(a) lower semicontinuous *at $\bar{x} \in X$ relative to X if for all sequences $\{x^k\} \subseteq X$ with $x^k \to \bar{x}$ and all $\bar{y} \in \Phi(\bar{x})$ there exists a number $k_0 \in \mathbb{N}$ and a sequence $\{y^k\} \subseteq Y$ with $y^k \to \bar{y}$ and $y^k \in \Phi(x^k)$ for all $k \geq k_0$;*

(b) closed *at $\bar{x} \in X$ relative to X if for all sequences $\{x^k\} \subseteq X$ with $x^k \to \bar{x}$ and all sequences $y^k \to \bar{y}$ with $y^k \in \Phi(x^k)$ for all $k \in \mathbb{N}$ sufficiently large we have $\bar{y} \in \Phi(\bar{x})$;*

(c) continuous *at $\bar{x} \in X$ relative to X if it is lower semicontinuous and closed at $\bar{x} \in X$ relative to X;*

(d) lower semicontinuous*,* closed *or* continuous *on X relative to X if it is lower semicontinuous, closed or continuous at every $x \in X$ relative to X.*

The definition of a lower semicontinuous set-valued mapping is in the sense of Berge. Alternative names used in the literature are 'open mapping' (see [25]) and 'inner semicontinuous mapping' (see [39]). Note that, here and in the following, relative properties of functions and mappings are meant relative to $\mathbb{R}^n$ if not stated otherwise. The next result, which follows immediately from [25, Corollaries 8.1 and 9.1], is used to prove the continuity of z_α.

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

$$Y(x) := \big\{ z \in \Phi(x) \;\big|\; \sup_{y \in \Phi(x)} v(x,y) = v(x,z) \big\},$$

and assume that $Y(\bar{x})$ is a singleton. Then the set-valued mapping $x \mapsto Y(x)$ is continuous at $\bar{x}$ relative to X.

Proposition 3.3 *Let F be continuous on $\mathbb{R}^n$. Consider a QVI with $S(x)$ being defined by* (10) *with $p \leq n$, $K \subseteq \mathbb{R}^p$ being nonempty, closed, and convex, c and Q being continuous, and $Q(x)$ having full rank for each fixed $x \in \mathbb{R}^n$. Then the function $x \mapsto z_\alpha(x)$ is continuous on $\mathbb{R}^n$.*

Proof. First recall that $-\psi_\alpha(x,\cdot)$ is concave for each fixed $x \in \mathbb{R}^n$ and continuous on $\mathbb{R}^n \times \mathbb{R}^p$. Since K is a closed set, the constant set-valued mapping $x \mapsto K$ is continuous on $\mathbb{R}^n$. Moreover, K is convex. Furthermore, the set

$$Z_\alpha(x) := \big\{ \zeta \in K \;\big|\; \max_{z \in K} \big(-\psi_\alpha(x,z) \big) = -\psi_\alpha\big(x,\zeta\big) \big\}$$

is a singleton for all $x \in \mathbb{R}^n$ since the function $\psi_\alpha(x,\cdot)$ is strongly convex for each fixed $x \in \mathbb{R}^n$, and the set K is nonempty, closed, and convex. Therefore, Lemma 3.2 implies that the (singleton-valued) set-valued mapping $x \mapsto Z_\alpha(x) = \{z_\alpha(x)\}$ is continuous on $\mathbb{R}^n$. Hence, the function $x \mapsto z_\alpha(x)$ is continuous on $\mathbb{R}^n$. $\square$

Since we minimize the function $\psi_\alpha(x,\cdot)$ with respect to a fixed set K, we may apply Danskin's Theorem and Proposition 3.3 and immediately obtain the following result.

Proposition 3.4 *Let F be continuously differentiable on $\mathbb{R}^n$. Consider a QVI with $S(x)$ being defined by* (10) *with $p \le n$, $K \subseteq \mathbb{R}^p$ being nonempty, closed, and convex, c and Q being continuously differentiable, and $Q(x)$ having full rank for each fixed $x \in \mathbb{R}^n$. Then g_α is continuously differentiable with gradient*

$$\begin{aligned}\nabla g_\alpha(x) &= -\nabla_x \psi_\alpha(x,z)\big|_{z=z_\alpha(x)} \\ &= \Big[\nabla F(x)\big(x - c(x) - Q(x)z\big) + \\ &\quad +\big(I - \nabla c(x) - \nabla_x\left(Q(x)z\right)\big)\big(\alpha\big(c(x)+Q(x)z-x\big)+F(x)\big)\Big]_{z=z_\alpha(x)}, \quad (12)\end{aligned}$$

where $z_\alpha(x)$ denotes the unique solution of problem (11).

Remark 3.5 A careful analysis of the above proofs shows that the introduced 'generalized moving set case' with a nonempty, closed and convex set K can, in principle, be further generalized to the case $S(x) = T(x,K)$ with any continuously differentiable nonlinear mapping $T : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$ such that $\psi_\alpha(x,z) = \varphi_\alpha(x, T(x,z))$ is strongly convex in z for all fixed $x \in \mathbb{R}^n$. In applications, however, it might be cumbersome to check the strong convexity assumption on ψ_α. $\Diamond$

Example 3.6 Let $p = n = 2$, $K = \mathbb{R}^2_+$, and F be continuously differentiable on $\mathbb{R}^2$. On K we simultaneously impose the translation $c(x) := x$, the scaling $\lambda(x) > 0$ and the rotation by the angle $\omega(x)$ for $x \in \mathbb{R}^2$, and we assume that also the functions λ and ω are continuously differentiable on $\mathbb{R}^2$. Then we may set $Q(x) := \lambda(x)R(x)$ with the rotation matrix

$$R(x) := \begin{pmatrix} \cos(\omega(x)) & -\sin(\omega(x)) \\ \sin(\omega(x)) & \cos(\omega(x)) \end{pmatrix}$$

and $S(x) = x + Q(x)K$. Clearly, $Q(x)$ is nonsingular for all $x \in \mathbb{R}^2$, and we obtain

$$\psi_\alpha(x,z) = F(x)^T Q(x) z + \frac{\alpha\lambda^2(x)}{2} z^T z.$$

For given $x \in \mathbb{R}^2$ the unconstrained minimal point of $\psi_\alpha(x,\cdot)$ is

$$z^*_\alpha(x) = -\frac{1}{\alpha\lambda(x)} R(x)^T F(x).$$

Therefore, the minimal point of $\psi_\alpha(x,\cdot)$ on $K = \mathbb{R}^2_+$ is

$$z_\alpha(x) = \max\left\{0, -\frac{1}{\alpha\lambda(x)} R(x)^T F(x)\right\}$$

(with the maximum taken componentwise) for all $x \in \mathbb{R}^2$. The function z_α is obviously continuous on $\mathbb{R}^2$, and the gap function

$$g_\alpha(x) = -\psi_\alpha(x, z_\alpha(x)) = \frac{\alpha\lambda^2(x)}{2}\left\|z_\alpha(x)\right\|^2 = \frac{1}{2\alpha}\left\|\max\{0, -R(x)^T F(x)\}\right\|^2$$

is also known to be continuously differentiable on $\mathbb{R}^2$. By Proposition 3.3 and Proposition 3.4, respectively, we get the same results for this particular example. Note that g_α does not depend on the scaling function λ.

Due to $0 \in K$ we have $x \in S(x) = x + Q(x)K$ for all x, so that $X = \mathbb{R}^2$ and, by Proposition 2.2, the solutions of the QVI are exactly the unconstrained minimal points of g_α with value zero, that is, the $x \in \mathbb{R}^2$ with

$$\max\{0, -R(x)^T F(x)\} = 0.$$

Thus, the solutions of the QVI are formed by the set $\{x \in \mathbb{R}^2 | \; R(x)^T F(x) \geq 0\}$. For a plot of the regularized gap function with the special choices $F(x) := x$, $\omega(x) := x_1 + x_2$ and $\alpha = 1$ see Figure 4. $\Diamond$

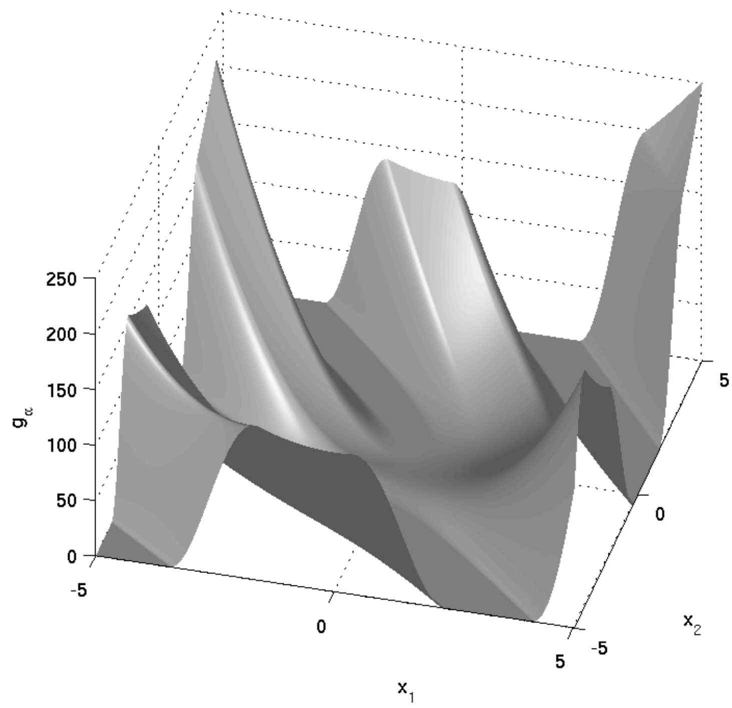

Figure 4: The regularized gap function with $\alpha = 1$ in Example 3.6

In Section 5, we will investigate the smoothness properties of the regularized gap function g_α in the general case.

3.2 QVIs with Set-valued Mappings in Product Form

Motivated by Example 2.3 (and Section 3.3 below), let us consider QVIs with a set-valued mapping S in product form, that is, for some $N \in \mathbb{N}$ and $n_\nu \in \mathbb{N}$, $\nu = 1, \ldots, N$, with

$n_1 + n_2 + \ldots + n_N = n$ there exist set-valued mappings $S_\nu : \mathbb{R}^n \rightrightarrows \mathbb{R}^{n_\nu}$, $\nu = 1, \ldots, N$ such that

$$S(x) = S_1(x) \times S_2(x) \times \ldots \times S_N(x)$$

holds for all $x \in \mathbb{R}^n$. After partitioning the variables $x = (x^1, \ldots, x^N)$ and $y = (y^1, \ldots, y^N)$ as well as the function $F(x) = (F^1(x), \ldots, F^N(x))$ accordingly, we may use the separability with respect to y of the function φ_α from (7) to obtain

$$\begin{aligned} g_\alpha(x) &= -\min_{y \in S(x)} \left[F(x)^T (y - x) + \frac{\alpha}{2} \|y - x\|^2 \right] \\ &= -\sum_{\nu=1}^N \min_{y^\nu \in S_\nu(x)} \left[F^\nu(x)^T (y^\nu - x^\nu) + \frac{\alpha}{2} \|y^\nu - x^\nu\|^2 \right] = \sum_{\nu=1}^N g_\alpha^\nu(x) \end{aligned} \tag{13}$$

with

$$g_\alpha^\nu(x) := -\min_{y^\nu \in S_\nu(x)} \left[F^\nu(x)^T (y^\nu - x^\nu) + \frac{\alpha}{2} \|y^\nu - x^\nu\|^2 \right], \quad \nu = 1, \ldots, N. \tag{14}$$

Lemma 3.7 *For all $x \in X$ and $\nu \in \{1, \ldots, N\}$, we have $g_\alpha^\nu(x) \geq 0$.*

Proof. For any $\nu \in \{1, \ldots, N\}$ choose some $x \in X$. Then we have

$$(x^1, x^2, \ldots, x^N) \in S_1(x) \times S_2(x) \times \ldots \times S_N(x)$$

and, in particular, $x^\nu \in S_\nu(x)$. Consequently, $g_\alpha^\nu(x)$ is minorized by the value of $-\left[F^\nu(x)^T (y^\nu - x^\nu) + \frac{\alpha}{2} \|y^\nu - x^\nu\|^2 \right]$ at $y^\nu := x^\nu$, which shows the assertion. □

The combination of Proposition 2.2, (13), and Lemma 3.7 immediately yields the following separation result.

Theorem 3.8 *A point $\bar{x}$ solves a QVI with set-valued mapping in product form if and only if $\bar{x}$ is the globally minimal point of g_α^ν on X with value zero for all $\nu = 1, \ldots, N$.*

Next, we combine the ideas of generalized moving sets from Section 3.1 with set-valued mappings in product form. In fact, the product form and the resulting separability allow each set $S_\nu(x)$, $\nu = 1, \ldots, N$, to be written as an *independent* generalized moving set, that is,

$$S_\nu(x) = \{ c^\nu(x) + Q^\nu(x) z \mid z \in K^\nu \} \tag{15}$$

where, for $p_\nu \leq n_\nu$, the set $K^\nu \subseteq \mathbb{R}^{p_\nu}$ is nonempty, closed, and convex, the functions $c^\nu : \mathbb{R}^n \to \mathbb{R}^{n_\nu}$ and $Q^\nu : \mathbb{R}^n \to \mathbb{R}^{n_\nu \times p_\nu}$ are continuous, and $Q^\nu(x)$ has full rank for all $x \in \mathbb{R}^n$. The proof of the assertion in Proposition 3.4 then translates word by word to a proof of the assertion that, under additional differentiability assumptions on F, c^ν and Q^ν, the function g_α^ν from (14) is continuously differentiable for each $\nu = 1, \ldots, N$ with known gradient.

To prepare the statement of this result note that, for $\nu = 1, \dots, N$, we may rewrite the function g_α^ν from (14) as

$$g_\alpha^\nu(x) = -\min_{y^\nu \in S_\nu(x)} \varphi_\alpha^\nu(x, y^\nu)$$

with

$$\varphi_\alpha^\nu(x, y^\nu) := F^\nu(x)^T(y^\nu - x^\nu) + \frac{\alpha}{2}\|y^\nu - x^\nu\|^2$$

for all $x \in X$. In analogy to (11), upon defining

$$\psi_\alpha^\nu(x, z^\nu) := \varphi_\alpha^\nu(x, c^\nu(x) + Q^\nu(x)z^\nu)$$

one can show that also

$$g_\alpha^\nu(x) = -\min_{z^\nu \in K^\nu} \psi_\alpha^\nu(x, z^\nu)$$

as well as

$$\nabla g_\alpha^\nu(x) \quad = \quad -\nabla_x \psi_\alpha^\nu(x, z^\nu)\big|_{z^\nu = z_\alpha^\nu(x)} \tag{16}$$

hold, where $z_\alpha^\nu(x)$ denotes the unique solution of the problem

$$\min_{z^\nu} \psi_\alpha^\nu(x, z^\nu) \quad \text{s.t.} \quad z^\nu \in K^\nu. \tag{17}$$

Consequently, (13) yields the following result.

Theorem 3.9 *Consider a QVI with set-valued mapping in product form and generalized moving sets of the form* (15) *where, for $p_\nu \leq n_\nu$, the set $K^\nu \subseteq \mathbb{R}^{p_\nu}$ is nonempty, closed, and convex, the functions F, c^ν and Q^ν are continuously differentiable, and $Q^\nu(x)$ has full rank for all $x \in \mathbb{R}^n$, $\nu = 1, \dots, N$. Then g_α is continuously differentiable with $\nabla g_\alpha(x) = \sum_{\nu=1}^N \nabla g_\alpha^\nu(x)$ and $\nabla g_\alpha^\nu(x)$ given by* (16).

Note that, under the above assumptions, $S(x)$ can be written as a generalized moving set in the form $S(x) = c(x) + Q(x)K$ with the nonempty, closed, and convex set $K = K^1 \times \ldots \times K^N$ in product form as well as

$$c(x) = \begin{pmatrix} c^1(x) \\ \vdots \\ c^N(x) \end{pmatrix} \quad \text{and} \quad Q(x) = \begin{pmatrix} Q^1(x) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & Q^N(x) \end{pmatrix},$$

where $Q(x)$ has full rank for all $x \in \mathbb{R}^n$.

3.3 Application to Generalized Nash Equilibrium Problems

A GNEP consists of a finite number of players $\nu = 1, \dots, N$ for some number $N \in \mathbb{N}$. Each player controls a set of variables $x^\nu \in \mathbb{R}^{n_\nu}$ for some positive number $n_\nu \in \mathbb{N}$. The vector $x = (x^1, x^2, \dots, x^N) \in \mathbb{R}^n$ with $n := n_1 + n_2 + \ldots + n_N$ denotes the set of all variables stacked together. This vector is sometimes also written as $(x^\nu, x^{-\nu})$, where $x^{-\nu}$ subsumes all subvectors x^μ except for $\mu = \nu$. This notation is particularly useful in order to stress the importance of the ν-th block vector x^ν within x. Each player ν has its own objective function $\theta_\nu : \mathbb{R}^n \to \mathbb{R}$, possibly depending on the entire vector x, as well as a strategy set $X_\nu(x^{-\nu}) \subseteq \mathbb{R}^{n_\nu}$, possibly depending in the variables $x^{-\nu}$ of all the other players. A (generalized) *Nash equilibrium* or simply a *solution* of the GNEP is a vector $\bar{x} = (\bar{x}^1, \dots, \bar{x}^N) \in X_1(\bar{x}^{-1}) \times \dots \times X_1(\bar{x}^{-N})$ such that, for each player ν, the subvector $\bar{x}^\nu$ solves the minimization problem

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

cf. [12] for a survey on GNEPs.

A GNEP is called *player convex* if for each ν and each $x^{-\nu}$ the function $\theta_\nu(x^\nu, x^{-\nu})$ is convex in the variable x^ν, and the strategy set $X_\nu(x^{-\nu})$ is closed and convex for all $\nu \in \{1, \dots, N\}$ and $x \in \mathbb{R}^n$. Throughout this subsection, we therefore make the following smoothness and convexity assumptions.

Assumption 3.10 *(a) The cost functions $\theta_\nu : \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable for all $\nu \in \{1, \dots, N\}$.*

(b) $\theta_\nu(\cdot, x^{-\nu})$ is convex in the variable x^ν for all fixed $x^{-\nu}$ and all $\nu \in \{1, \dots, N\}$.

(c) The sets $X_\nu(x^{-\nu})$ are closed and convex for all $\nu \in \{1, \dots, N\}$ and $x \in \mathbb{R}^n$.

Under Assumption 3.10, it is well-known, see, e.g., [12, 21], that a GNEP is equivalent to a QVI in the sense that $\bar{x}$ is a solution of the GNEP if and only if $\bar{x}$ solves the corresponding QVI with F being defined by

$$F^{GNEP}(x) := F(x) := \begin{pmatrix} \nabla_{x^1}\theta_1(x) \\ \vdots \\ \nabla_{x^N}\theta_N(x) \end{pmatrix}$$

and $S(x)$ having the product structure (cf. Section 3.2)

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

The regularized gap function of this particular QVI therefore reads

$$\begin{aligned} g_\alpha(x) &= -\min_{y \in S(x)} \left[F^{GNEP}(x)^T(y - x) + \frac{\alpha}{2}\|y - x\|^2 \right] \\ &= -\min_{y \in S(x)} \left[\sum_{\nu=1}^N \left(\nabla_{x^\nu}\theta_\nu(x^\nu, x^{-\nu})^T(y^\nu - x^\nu) + \frac{\alpha}{2}\|y^\nu - x^\nu\|^2 \right) \right] \\ &= -\sum_{\nu=1}^N \min_{y^\nu \in X_\nu(x^{-\nu})} \left[\nabla_{x^\nu}\theta_\nu(x^\nu, x^{-\nu})^T(y^\nu - x^\nu) + \frac{\alpha}{2}\|y^\nu - x^\nu\|^2 \right]. \end{aligned}$$

Taking into account the convexity of θ_ν as a function of x^ν, it follows that

$$\begin{aligned} g_\alpha(x) &\geq -\sum_{\nu=1}^{N} \min_{y^\nu \in X_\nu(x^{-\nu})} \left[\theta_\nu(y^\nu, x^{-\nu}) - \theta_\nu(x^\nu, x^{-\nu}) + \frac{\alpha}{2}\|y^\nu - x^\nu\|^2\right] \\ &= -\min_{y \in S(x)} \Phi_\alpha(x, y) \;=:\; V_\alpha(x), \end{aligned}$$

where

$$\Phi_\alpha(x, y) := \sum_{\nu=1}^{N} \left(\theta_\nu(y^\nu, x^{-\nu}) - \theta_\nu(x^\nu, x^{-\nu}) + \frac{\alpha}{2}\|y^\nu - x^\nu\|^2\right). \tag{18}$$

The functions $-\Phi_\alpha$ and V_α defined in this way are the *regularized Nikaido-Isoda function* and the corresponding *regularized value function*, respectively, known both from theoretical and numerical considerations in the GNEP context, see, e.g., [11, 24]. We summarize the previous discussion in the following result.

Lemma 3.11 *Let Assumption 3.10 hold. Consider a QVI arising from a player convex GNEP, and let g_α and V_α be the corresponding regularized gap function and regularized value function, respectively. Then $g_\alpha(x) \geq V_\alpha(x)$ holds for all $x \in \mathbb{R}^n$.*

The previous result implies, for example, that any error bound result for V_α also gives an error bound result for the regularized gap function g_α, whereas the converse might not be true.

Next, we also study the differentiability properties of the regularized value function V_α of player convex GNEPs in the generalized moving set case $S(x) = c(x) + Q(x)K$ defined by (10). In fact, due to the inherent product structure of $S(x)$ in the GNEP case, we have

$$S(x) = S_1(x) \times \ldots \times S_N(x)$$

with $S_\nu(x) = X_\nu(x^{-\nu})$, $\nu = 1, \ldots, N$, so that we may use independent generalized moving sets for each player as defined in (15):

$$X_\nu(x^{-\nu}) = \{c^\nu(x^{-\nu}) + Q^\nu(x^{-\nu})z \mid z \in K^\nu\} \tag{19}$$

where, for $p_\nu \leq n_\nu$, the set $K^\nu \subseteq \mathbb{R}^{p_\nu}$ is nonempty, closed, and convex, the functions $c^\nu : \mathbb{R}^{n-n_\nu} \to \mathbb{R}^{n_\nu}$ and $Q^\nu : \mathbb{R}^{n-n_\nu} \to \mathbb{R}^{n_\nu \times p_\nu}$ are continuous, and $Q^\nu(x^{-\nu})$ has full rank for all $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$.

Note that, under the additional assumption of continuous differentiability of the functions F^{GNEP} (that is, twice continuous differentiability of the functions θ_ν), c^ν and Q^ν, $\nu = 1, \ldots, N$, the regularized *gap* function g_α is continuously differentiable with known gradient by Theorem 3.9. The corresponding analysis for the regularized *Nikaido-Isoda* function V_α is similar to the one given in Section 3.2. A first difference is that in the description

$$V_\alpha(x) = -\min_{y \in S(x)} \Phi_\alpha(x, y)$$

the function Φ_α from (18) is not separable with respect to all components of y, while the function φ_α from (7) in the description

$$g_\alpha(x) = -\min_{y \in S(x)} \varphi_\alpha(x, y)$$

is. However, Φ_α obviously is separable with respect to the vectors $y^1, \dots, y^N$ which suffices to mimic the proof of continuous differentiability of g_α in Theorem 3.9 to show continuous differentiability of V_α. In fact, the separability allows us to write $V_\alpha(x) = \sum_{\nu=1}^{N} V_\alpha^\nu(x)$ with

$$V_\alpha^\nu(x) := -\min_{y^\nu \in X_\nu(x^{-\nu})} \Phi_\alpha^\nu(x, y^\nu)$$

and

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

or, equivalently,

$$V_\alpha^\nu(x) = -\min_{z^\nu \in K^\nu} \Psi_\alpha^\nu(x, z^\nu)$$

with

$$\Psi_\alpha^\nu(x, z^\nu) := \Phi_\alpha^\nu(x, c^\nu(x^{-\nu}) + Q^\nu(x^{-\nu})z^\nu)$$

for $\nu = 1, \dots, N$. As a second difference to the analysis of the gap function, the strong convexity of Ψ_α^ν in z^ν is slightly less apparent. In fact, the convexity of Φ_α^ν in y^ν implies the convexity of Ψ_α^ν in z^ν. Moreover, by the full rank of $Q^\nu(x)$, the matrix

$$\nabla_{z^\nu z^\nu} \Psi_\alpha^\nu(x, z^\nu) = Q^\nu(x^{-\nu})^T \left(\nabla_{y^\nu y^\nu} \Phi_\alpha^\nu(x, y^\nu)|_{y^\nu = c^\nu(x^{-\nu}) + Q^\nu(x^{-\nu})z^\nu} \right) Q^\nu(x^{-\nu})$$

with

$$\nabla_{y^\nu y^\nu} \Phi_\alpha^\nu(x, y^\nu) = \nabla_{y^\nu y^\nu} \theta_\nu(y^\nu, x^{-\nu}) + \alpha I$$

is uniformly positive definite (in z^ν), so that Ψ_α^ν even is strongly convex in z^ν. Therefore, for each $\nu = 1, \dots, N$ the problem

$$\min_{z^\nu} \Psi_\alpha^\nu(x, z^\nu) \quad \text{s.t.} \quad z^\nu \in K^\nu$$

has a unique solution $z_\alpha^\nu(x)$, and along the lines of Section 3.2 we obtain that V_α^ν is continuously differentiable with

$$\nabla V_\alpha^\nu(x) = -\nabla_x \Psi_\alpha^\nu(x, z^\nu)|_{z^\nu = z_\alpha^\nu(x)} \tag{20}$$

where

$$\begin{aligned}
\nabla_{x^\nu} \Psi_\alpha^\nu(x, z^\nu) &= \Big(-\nabla_{x^\nu} \theta(x^\nu, x^{-\nu}) - \alpha(y^\nu - x^\nu) \Big)|_{y^\nu = c^\nu(x^{-\nu}) + Q^\nu(x^{-\nu})z^\nu}, \\
\nabla_{x^{-\nu}} \Psi_\alpha^\nu(x, z^\nu) &= \Big(\nabla_{x^{-\nu}} \theta_\nu(y^\nu, x^{-\nu}) - \nabla_{x^{-\nu}} \theta_\nu(x^\nu, x^{-\nu}) \\
&\quad + \left(\nabla_{x^{-\nu}} c^\nu(x^{-\nu}) + \nabla_{x^{-\nu}} (Q^\nu(x^{-\nu}) z^\nu) \right) \\
&\quad \cdot \left(\nabla_{x^\nu} \theta_\nu(y^\nu, x^{-\nu}) + \alpha(y^\nu - x^\nu) \right) \Big)|_{y^\nu = c^\nu(x^{-\nu}) + Q^\nu(x^{-\nu})z^\nu}.
\end{aligned}$$

The following theorem summarizes the previous discussion.

Theorem 3.12 *Consider a GNEP with strategy spaces of generalized moving set form* (19) *where, for $p_\nu \leq n_\nu$, the set $K^\nu \subseteq \mathbb{R}^{p_\nu}$ is nonempty, closed, and convex, the functions θ_ν are twice continuously differentiable, the functions c^ν and Q^ν are continuously differentiable, and $Q^\nu(x^{-\nu})$ has full rank for all $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$, $\nu = 1, \ldots, N$. Then V_α is continuously differentiable with $\nabla V_\alpha(x) = \sum_{\nu=1}^N \nabla V_\alpha^\nu(x)$ and $\nabla V_\alpha^\nu(x)$ given by* (20).

4 Continuity Properties and the Domain of g_α for General QVIs

The first part of this section shows that the solution y_α of the problem (8) is continuous at $\bar{x} \in M = \operatorname{dom} S$ if $S(\bar{x})$ satisfies the Slater condition, i.e., if there exists some $\bar{y} \in \mathbb{R}^n$ satisfying $s_i(\bar{x}, \bar{y}) < 0$ for all $i = 1, \ldots, m$. We therefore define the 'degenerate point set'

$$D_1 := \{x \in M \mid \text{the set } S(x) \text{ violates the Slater condition}\}.$$

Note that continuity of y_α, in particular, implies the continuity of the regularized gap function g_α at $\bar{x}$. The corresponding analysis is similar to the one given in [10] and [11] for certain objective functions arising in the context of jointly and player convex GNEPs, respectively.

After two generalizations of our main result, the second part of this section then studies a topological property of the set $M \setminus D_1$.

Theorem 4.1 *Let Assumption 1.1 hold and let the set-valued mapping S be lower semicontinuous at $\bar{x} \in M$. Then the functions y_α and g_α are continuous at $\bar{x}$.*

Proof. Recall that $\varphi_\alpha(x, \cdot)$ is convex for each fixed $x \in \mathbb{R}^n$ and continuous on $\mathbb{R}^n \times \mathbb{R}^n$. Therefore, $-\varphi_\alpha(x, \cdot)$ is concave for each fixed $x \in \mathbb{R}^n$ and continuous on $\mathbb{R}^n \times \mathbb{R}^n$.

The set-valued mapping S is closed since its graph

$$\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mid s_i(x, y) \leq 0 \quad \forall i = 1, \ldots, m\}$$

is a closed set in view of continuity of s_i, $i = 1, \ldots, m$, see [25, Theorem 2]. Due to Assumption 1.1, $S(x)$ is convex for all $x \in \mathbb{R}^n$. Moreover, the set

$$Y_\alpha(x) = \left\{ z \in S(x) \;\middle|\; \max_{y \in S(x)} \big(-\varphi_\alpha(x, y) \big) = -\varphi_\alpha(x, z) \right\}$$

is a singleton with the unique element $y_\alpha(x)$ for all $x \in M$, see Remark 2.1. Therefore, Lemma 3.2 implies that the (singleton-valued) set-valued mapping $x \mapsto \{y_\alpha(x)\}$ is continuous at $\bar{x}$. Hence, the function $x \mapsto y_\alpha(x)$ is continuous at $\bar{x}$. Moreover, $g_\alpha(x) = -\varphi_\alpha(x, y_\alpha(x))$ is continuous at $\bar{x}$ as a composition of continuous functions. $\square$

As an immediate consequence of Theorem 4.1, we obtain the following result.

Corollary 4.2 *Let Assumption 1.1 hold. Then y_α and g_α are continuous on $M \setminus D_1$.*

Proof. Let $\bar{x} \in M \setminus D_1$. Due to Assumption 1.1 and the Slater condition for $S(\bar{x})$, the set-valued mapping S is lower semicontinuous at $\bar{x}$ (see [25, Theorem 12]). Therefore, Theorem 4.1 implies that the functions y_α and g_α are continuous at $\bar{x}$. $\square$

Let us illustrate the previous result in the context of Example 2.3.

Example 4.3 In the situation of Example 2.3, for $x \in M$ the set $S(x) = S_1(x) \times S_2(x)$ satisfies the Slater condition if and only if $S_1(x)$ as well as $S_2(x)$ possess a Slater point. Clearly, $S_1(x)$ satisfies the Slater condition for all $x \in M$. On the other hand, $S_2(x)$ violates the Slater condition exactly for all x with $x_1 = -1/\sqrt{2}$ and for all x with $x_1 = 1$. Hence, we obtain $D_1 = \Big(\{-1/\sqrt{2}\} \cup \{1\}\Big) \times \mathbb{R}$ and, by Corollary 4.2, the functions y_α and g_α are continuous on $M \setminus D_1 = (-1/\sqrt{2}, 1) \times \mathbb{R}$. $\diamondsuit$

Direct inspection of the functions y_α and g_α from Example 2.3 shows that they are actually continuous at least on all of X ($\subseteq M$) (relative to X). This motivates to relax the assumption of Corollary 4.2 in the spirit of [11, Theorem 3.5] for generalized Nash equilibrium problems. To this end, let us define the set

$$D_1' := \{x \in M \mid \text{the set } S(x) \text{ violates the Slater condition and is not a singleton}\}.$$

The following result shows that y_α and hence also g_α are continuous on the set $X \setminus D_1'$ (relative to X), i.e., they are continuous at every point $x \in X$ (relative to X) where $S(x)$ either satisfies the Slater condition or reduces to a single point. Note that the latter degenerate case occurs quite frequently, e.g., in the context of GNEPs.

Theorem 4.4 *Let Assumption 1.1 hold. Then y_α and g_α are continuous on $X \setminus D_1'$ (relative to X).*

Proof. Let $\bar{x} \in X \setminus D_1'$. In view of (3) and Corollary 4.2 we only have to consider the case that $S(\bar{x})$ is a singleton. Due to $\bar{x} \in X$, we actually have $S(\bar{x}) = \{\bar{x}\}$. Choose any sequence $\{x^k\} \subseteq X$ with $x^k \to \bar{x}$. Then for each $k \in \mathbb{N}$ we have $x^k \in S(x^k)$, so that S turns out to be lower semi-continuous at $\bar{x}$ (relative to X). Theorem 4.1 now yields the assertion. $\square$

Unfortunately, in Example 2.3 we obtain $X \cap D_1 = X \cap D_1' = \{(1,0)\}$ as $S((1,0)) = [0,+\infty) \times \{0\}$ violates the Slater condition while *not* being a singleton. Hence, Theorem 4.4 may not be evoked to show continuity of y_α and g_α on all of X (relative to X). However, the product form of the set-valued mapping S in Example 2.3 justifies to modify the assumptions of Theorem 4.4. Let us consider the general case of a set-valued mapping in product form (cf. Section 3.2)

$$S(x) = S_1(x) \times S_2(x) \times \ldots \times S_N(x)$$

and define

$$D_1'' := \{x \in M \mid \text{For some } \nu \in \{1, \ldots, N\} \text{ the set } S_\nu(x) \text{ violates the Slater condition and is not a singleton}\}.$$

Recall that this product structure of $S(x)$ arises quite naturally in the GNEP context (s. Sec. 3.3 below).

Theorem 4.5 *Let Assumption 1.1 hold, and let S be given in product form. Then the functions y_α and g_α are continuous on $X \setminus D_1''$ (relative to X).*

Proof. Let $\bar{x} \in X \setminus D_1''$. Then for each $\nu \in \{1, \ldots, N\}$ the set $S_\nu(\bar{x})$ either satisfies the Slater condition or coincides with the singleton $\{\bar{x}^\nu\}$. Choose any sequence $\{x^k\} \subseteq X$ with $x^k \to \bar{x}$ and any $\bar{y} \in S(\bar{x})$, that is, we have $x^{\nu,k} \to \bar{x}^\nu$ and $\bar{y}^\nu \in S_\nu(\bar{x})$, $\nu = 1, \ldots, N$. For those $\nu \in \{1, \ldots, N\}$ with $S_\nu(\bar{x})$ satisfying the Slater condition, the set-valued mapping S_ν is lower semi-continuous at $\bar{x}$, so that for sufficiently large k a sequence $y^{\nu,k} \in S_\nu(x^k)$ with $y^{\nu,k} \to \bar{y}^\nu$ exists. On the other hand, for $\nu \in \{1, \ldots, N\}$ with $S_\nu(\bar{x}) = \{\bar{x}^\nu\}$, as in the proof of Theorem 4.4, we may choose $y^{\nu,k} := x^{\nu,k} \in S_\nu(x^k)$ and obtain $y^{\nu,k} = x^{\nu,k} \to \bar{x}^\nu = \bar{y}^\nu$. This shows the lower semi-continuity of S at $\bar{x}$ (relative to X), and Theorem 4.1 yields the assertion. □

Note that in Example 2.3 we have $X \setminus D_1'' = X$, so that Theorem 4.5 finally yields the continuity of y_α and g_α on all of X (relative to X).

Let us return to the set D_1 which is also important in our analysis of the differentiability properties of g_α. In fact, in Section 5, we shall study differentiability of g_α at points from the topological interior of the domain of g_α where, in view of Remark 2.1,

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

coincides with M. Therefore, their topological interiors also coincide:

$$\operatorname{int} \operatorname{dom} g_\alpha = \operatorname{int} M. \tag{21}$$

The following result relates the set $M \setminus D_1$ to the interior of the domain of g_α.

Lemma 4.6 *Let Assumption 1.1 hold. Then the set $M \setminus D_1$ is open and satisfies*

$$M \setminus D_1 \subseteq \operatorname{int} \operatorname{dom} g_\alpha. \tag{22}$$

Proof. Let $\bar{x} \in M \setminus D_1$. Then there exists some $\bar{y} \in \mathbb{R}^n$ satisfying $s_i(\bar{x}, \bar{y}) < 0$ for all $i = 1, \ldots, m$. Due to continuity of the functions s_i, $i = 1, \ldots, m$, we can choose a neighborhood U of $\bar{x}$ such that for all $x \in U$ also $s_i(x, \bar{y}) < 0$ is satisfied for all $i = 1, \ldots, m$. Therefore, for all $x \in U$ the set $S(x)$ satisfies the Slater condition, that is, we have $x \in M \setminus D_1$. This shows that $M \setminus D_1$ is open. In particular, U is contained in $\operatorname{dom} S = M$. This implies $\bar{x} \in \operatorname{int} M$ and, due to (21), shows the second assertion. □

Remark 4.7 Lemma 4.6 guarantees that the set $\operatorname{dom} g_\alpha \setminus D_1$ is an open subset of $\operatorname{int} \operatorname{dom} g_\alpha$, so that we will be able to study Fréchet differentiability of g_α on $\operatorname{dom} g_\alpha \setminus D_1$ in Section 5. We point out that, under stronger convexity and regularity assumptions, along the lines of [22, Theorem 3.9] one can also show the reverse inclusion in Lemma 4.6, that is, the topological boundary of $\operatorname{dom} g_\alpha$ coincides with D_1. For an illustration of this result see Example 4.3. $\Diamond$

5 Differentiability Properties for General QVIs

Assumption 1.1 together with the following Assumption 5.1 are the blanket assumptions for this section.

Assumption 5.1 *The functions F and s_i, $i = 1, \ldots, m$, are continuously differentiable.*

We want to study differentiability properties of g_α. To this end, we have to make sure that we consider differentiability only at points from the interior of the domain of g_α, since otherwise it makes no sense to talk about (Fréchet) differentiability. In view of Lemma 4.6, it is reasonable to investigate the differentiability of the function g_α on the set $M \setminus D_1$. To this end, consider once again the convex optimization problem from (8). In view of Remark 2.1, this problem has a unique optimal point $y_\alpha(x)$ for all $x \in M$, in particular, for all $x \in M \setminus D_1$. Let

$$L_\alpha(x, y, \lambda) := \varphi_\alpha(x, y) + \sum_{i=1}^{m} \lambda^i s_i(x, y)$$

denote the Lagrange function of the optimization problem (8), and let

$$KKT_\alpha(x) := \left\{ \lambda \in \mathbb{R}^m \;\middle|\; F(x) + \alpha\big(y_\alpha(x) - x\big) + \sum_{i=1}^{m} \lambda^i \nabla_y s_i\big(x, y_\alpha(x)\big) = 0, \right.$$
$$\left. \lambda^i \geq 0,\ \lambda^i s_i\big(x, y_\alpha(x)\big) = 0 \ \ \forall i = 1, \ldots, m \right\}$$

be the set of Karush-Kuhn-Tucker multipliers for $y_\alpha(x) \in S(x)$. Note that the convex polyhedron $KKT_\alpha(x)$ is a convex polytope if and only if $S(x)$ satisfies the Slater condition [18], that is, for $x \in M \setminus D_1$. Furthermore,

$$\mathcal{I}_\alpha(x) := \big\{ i \mid s_i\big(x, y_\alpha(x)\big) = 0 \big\}$$

will denote the set of active indices of $y_\alpha(x) \in S(x)$.

Before stating the next result, we recall that a real-valued function 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 5.2 *Let Assumptions 1.1 and 5.1 hold and let $x \in M \setminus D_1$. Then the regularized gap function g_α is Hadamard directionally differentiable at x with*

$$g'_\alpha(x; d) = \min_{\lambda \in KKT_\alpha(x)} \left[\Big(F(x) - \big(\nabla F(x) - \alpha I\big)\big(y_\alpha(x) - x\big) - \sum_{i=1}^m \lambda^i \nabla_x s_i\big(x, y_\alpha(x)\big)\Big)^T d \right] \tag{23}$$

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

Proof. Since $x \in M \setminus D_1$, the set $S(x)$ satisfies the Slater condition. A standard result from parametric optimization (see, e.g., [20, 26, 38]) then states that the optimal value function of (8), that is, $-g_\alpha$, is Hadamard directionally differentiable at x with

$$(-g_\alpha)'(x; d) = \max_{\lambda \in KKT_\alpha(x)} \big(\nabla_x L_\alpha(x, y, \lambda)|_{y = y_\alpha(x)} \big)^T d$$

for all $d \in \mathbb{R}^n$. After a short calculation, this shows the assertion. □

Remark 5.3 Note that, in the assertion of Theorem 5.2 and in the following, for any $x \in M \setminus D_1$ and any $\lambda \in KKT_\alpha(x)$ one may replace the term $\sum_{i=1}^m \lambda^i \nabla_x s_i\big(x, y_\alpha(x)\big)$ by $\sum_{i \in \mathcal{I}_\alpha(x)} \lambda^i \nabla_x s_i\big(x, y_\alpha(x)\big)$. ◊

The formula (23) for the directional derivative of g_α at some $x \in M \setminus D_1$ simplifies if not only the optimal point set $\{y_\alpha(x)\}$ of (8) is a singleton, but also the Karush-Kuhn-Tucker set $KKT_\alpha(x)$. This motivates to define a next 'degenerate point set'

$$D_2 := \{x \in M \mid \text{the set } KKT_\alpha(x) \text{ is not a singleton}\}.$$

As mentioned before, the convex polyhedron $KKT_\alpha(x)$ is a convex polytope if and only $x \in M \setminus D_1$. Hence, for $x \in D_1$ the set $KKT_\alpha(x)$ is either empty or unbounded, but certainly not a singleton. This shows the relation

$$D_1 \subseteq D_2. \tag{24}$$

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. Theorem 5.2 and (24) lead to the following result.

Corollary 5.4 *Let Assumptions 1.1 and 5.1 hold, and let $x \in M \setminus D_2$ with $KKT_\alpha(x) = \{\lambda_\alpha(x)\}$. Then the regularized gap function g_α is Gâteaux differentiable at x with*

$$g'_\alpha(x;d) = \Big(F(x) - \big(\nabla F(x) - \alpha I\big)\big(y_\alpha(x) - x\big) - \sum_{i=1}^{m} \lambda_\alpha^i(x) \nabla_x s_i\big(x, y_\alpha(x)\big)\Big)^T d \tag{25}$$

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

For algebraic characterizations of the sets D_1 and D_2 recall that the Mangasarian Fromovitz constraint qualification, MFCQ for short, holds at $y_\alpha(x) \in S(x)$ if there exists a $d \in \mathbb{R}^n$ satisfying $\nabla_y s_i\big(x, y_\alpha(x)\big)^T d < 0$ for all $i \in \mathcal{I}_\alpha(x)$. Note that, because of the convexity of the functions $s_i(x, \cdot)$, $i = 1, \ldots, m$, for each fixed x, MFCQ holds at $y_\alpha(x)$ if and only if the Slater condition for $S(x)$ is satisfied. Hence, we have the characterization

$$D_1 = \{x \in M \mid \text{MFCQ is violated at } y_\alpha(x) \text{ in } S(x)\}.$$

Furthermore, it is known from [27] that the strict Mangasarian Fromovitz constraint qualification, SMFCQ for short, at $y_\alpha(x) \in S(x)$ characterizes a unique KKT multiplier $\lambda_\alpha(x)$ at the optimal point $y_\alpha(x)$; here, SMFCQ holds at $y_\alpha(x)$ in $S(x)$ with the multiplier $\lambda_\alpha \in KKT_\alpha(x)$ if the gradients

$$\nabla_y s_i\big(x, y_\alpha(x)\big), \quad i \in \mathcal{I}_\alpha^+(x) = \big\{i \in \mathcal{I}_\alpha(x) \mid \lambda_\alpha^i > 0\big\},$$

are linearly independent, and there exists a $d \in \mathbb{R}^n$ satisfying

$$\begin{aligned} \nabla_y s_i\big(x, y_\alpha(x)\big)^T d < 0 \quad & \forall i \in \mathcal{I}_\alpha^0(x) = \big\{i \in \mathcal{I}_\alpha(x) \mid \lambda_\alpha^i = 0\big\}, \\ \nabla_y s_i\big(x, y_\alpha(x)\big)^T d = 0 \quad & \forall i \in \mathcal{I}_\alpha^+(x). \end{aligned}$$

Therefore we arrive at

$$D_2 = \{x \in M \mid \text{SMFCQ is violated at } y_\alpha(x) \text{ in } S(x)\},$$

which, since SMFCQ implies MFCQ at $y_\alpha(x)$, yields an alternative proof of (24).

Finally, the linear independence constraint qualification, LICQ for short, is said to hold at $y_\alpha(x) \in S(x)$ if the vectors $\nabla_y s_i\big(x, y_\alpha(x)\big)$ $(i \in \mathcal{I}_\alpha(x))$ are linearly independent. As LICQ implies SMFCQ at $y_\alpha(x) \in S(x)$, the set

$$D_3 = \{x \in M \mid \text{LICQ is violated at } y_\alpha(x) \text{ in } S(x)\}$$

satisfies

$$D_1 \subseteq D_2 \subseteq D_3. \tag{26}$$

For the proof of the next result recall that, if a function $f : U \to \mathbb{R}$ with open domain U is Gâteaux differentiable on U, and the partial derivatives of f are continuous at $\bar{x} \in U$, then f is continuously differentiable at $\bar{x}$.

Theorem 5.5 *Let Assumptions 1.1 and 5.1 hold, and let $\bar{x} \in M \setminus D_3$ with $KKT_\alpha(\bar{x}) = \{\lambda_\alpha(\bar{x})\}$. Then the regularized gap function g_α is continuously differentiable in a neighborhood of $\bar{x}$ with*

$$\nabla g_\alpha(\bar{x}) = F(\bar{x}) - \big(\nabla F(\bar{x}) - \alpha I\big)\big(y_\alpha(\bar{x}) - \bar{x}\big) - \sum_{i=1}^{m} \lambda_\alpha^i(\bar{x}) \nabla_x s_i\big(\bar{x}, y_\alpha(\bar{x})\big).$$

Proof. First, due to (26) and Lemma 4.6, $\bar{x}$ is an interior point of $\operatorname{dom} g_\alpha$, and there is some neighborhood U of $\bar{x}$ such that for all $x \in U$ the optimal point $y_\alpha(x) \in S(x)$ satisfies the Slater condition. By Corollary 4.2, the function y_α is actually continuous on U. Consequently, since LICQ is stable under perturbations, U may be chosen such that LICQ holds at $y_\alpha(x) \in S(x)$ for each $x \in U$. This implies that KKT_α is single-valued on U, say $KKT_\alpha(x) = \{\lambda_\alpha(x)\}$ for $x \in U$. Corollary 5.4 thus guarantees that g_α is Gâteaux differentiable on U with (25). By [26, Lemma 2] the set-valued mapping KKT_α is locally bounded and closed on U. As it is also singleton-valued in our case, the function λ_α is continuous on U, so that the partial derivatives of g_α are continuous at $\bar{x}$. This shows continuous differentiability of g_α at $\bar{x}$ with the asserted gradient. Since the partial derivatives of g_α actually are continuous on all of U, also continuous differentiability of g_α on U follows. □

Remark 5.6 The main reason to use D_3 instead of the smaller set D_2 in the assumption of Theorem 5.5 is the lack of stability of SMFCQ (cf. also Example 5.7 below). On the other hand, a different sufficient condition for continuous differentiability of g_α can be obtained in cases when SMFCQ is stable. In particular, if the set $\mathcal{I}_\alpha^0(x) = \{i \in \mathcal{I}_\alpha(x) \mid \lambda_\alpha^i = 0\}$ remains constant under small perturbations of x (e.g., due to $\mathcal{I}_\alpha^0(x) = \emptyset$, i.e., strict complementary slackness), then continuity arguments show that SMFCQ is stable at $y_\alpha(x)$ under sufficiently small perturbations of x. After this observation, along the lines of the proof of Theorem 5.5 one can show continuous differentiability of g_α on a neighborhood of $\bar{x}$. ◊

Example 5.7 Let us illustrate our results for the QVI from Example 2.3 and check differentiability properties of the regularized gap function g_α on $X \setminus D_1$. Note that Assumptions 1.1 and 5.1 hold for this example. By Theorem 5.5, g_α is continuously differentiable at each $x \in X \setminus D_3$ with known gradient. In the following, we will determine the sets $X \cap (D_3 \setminus D_1)$ and $X \cap (D_2 \setminus D_1)$ as well as the corresponding directional derivatives of g_α.

By definition of D_3 one has

$$X \cap (D_3 \setminus D_1) = \{x \in X \setminus D_1 \mid \text{LICQ is violated at } y_\alpha(x) \text{ in } S(x)\}$$

so that we have to check for violation of LICQ. The involved gradients are

$$\nabla_y s_1(x, y_\alpha(x)) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \quad \nabla_y s_2(x, y_\alpha(x)) = \begin{pmatrix} 0 \\ 2(y_\alpha(x))_2 \end{pmatrix}, \quad \nabla_y s_3(x, y_\alpha(x)) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Some tedious calculations show that the activities are characterized as follows, where we use the functions ϱ_i from Example 2.3:

$$\begin{aligned}
\{x \in X \setminus D_1 |\ 1 \in \mathcal{I}_\alpha(x)\} &= \{x \in X \setminus D_1 |\ \varrho_1(x) \le 1/\alpha\},\\
\{x \in X \setminus D_1 |\ 2 \in \mathcal{I}_\alpha(x)\} &= \{x \in X \setminus D_1 |\ \varrho_2(x) \le 1/\alpha,\ x_1 \ge 1/\sqrt{2}\}\\
&\quad \cup \{x \in X \setminus D_1 |\ 1/\alpha \le \varrho_3(x)\},\\
\{x \in X \setminus D_1 |\ 3 \in \mathcal{I}_\alpha(x)\} &= \{x \in X \setminus D_1 |\ \varrho_2(x) \le 1/\alpha,\ x_1 \le 1/\sqrt{2}\}.
\end{aligned}$$

In particular, if $2 \in \mathcal{I}_\alpha(x)$, for all $x \in X \setminus D_1$ with $\varrho_2(x) \le 1/\alpha,\ x_1 \ge 1/\sqrt{2}$ we find

$$\nabla_y s_2(x, y_\alpha(x)) = \begin{pmatrix} 0 \\ -2\sqrt{1-x_1^2} \end{pmatrix} \neq 0,$$

and for all $x \in X \setminus D_1$ with $1/\alpha \le \varrho_3(x)$

$$\nabla_y s_2(x, y_\alpha(x)) = \begin{pmatrix} 0 \\ 2\sqrt{1-x_1^2} \end{pmatrix} \neq 0,$$

so that

$$X \cap (D_3 \setminus D_1) = \{x \in X \setminus D_1 |\ \{2,3\} \subseteq \mathcal{I}_\alpha(x)\}.$$

As $\varrho_3(x) < \varrho_2(x)$ holds for all $x \in X \setminus D_1$, this implies

$$\begin{aligned}
X \cap (D_3 \setminus D_1) &= \left\{ x \in X \setminus D_1 \middle|\ \varrho_2(x) \le \frac{1}{\alpha},\ x_1 = \frac{1}{\sqrt{2}} \right\}\\
&= \left\{ x \in X \setminus D_1 \middle|\ x_2 + \frac{1}{\sqrt{2}} \le \frac{1}{\alpha},\ x_1 = \frac{1}{\sqrt{2}} \right\}\\
&= \left\{ \frac{1}{\sqrt{2}} \right\} \times \left[-\frac{1}{\sqrt{2}}, \min\left\{ \frac{1}{\sqrt{2}}, \frac{1}{\alpha} - \frac{1}{\sqrt{2}} \right\} \right].
\end{aligned}$$

For sufficiently small $\alpha > 0$, that is, for $\alpha \le 1/\sqrt{2}$, this results in

$$X \cap (D_3 \setminus D_1) = \{x \in X \setminus D_1 |\ x_1 = 1/\sqrt{2}\}$$

and, as will become apparent below, the latter corresponds to a 'concave kink in the graph of g_α on X along the line segment connecting the boundary points $(1/\sqrt{2}, -1/\sqrt{2})$ and $(1/\sqrt{2}, 1/\sqrt{2})$ of X'.

The example exhibits a more interesting feature, however, for $\alpha > 1/\sqrt{2}$ when

$$X \cap (D_3 \setminus D_1) = \left\{ \frac{1}{\sqrt{2}} \right\} \times \left[-\frac{1}{\sqrt{2}}, \frac{1}{\alpha} - \frac{1}{\sqrt{2}} \right].$$

In the following we will see that this corresponds to a 'concave kink in the graph of g_α on X along the line segment connecting the boundary point $(1/\sqrt{2}, -1/\sqrt{2})$ and the *interior* point $(1/\sqrt{2}, -1/\sqrt{2}+1/\alpha)$ of X'. For $\alpha = 1\,(> 1/\sqrt{2})$, this kink is visualized in Figure 2. For simplicity, in the remainder of this example, let us focus on the case $\alpha = 1$ with

$$X \cap (D_3 \setminus D_1) = \left\{ x(t) := \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} + t \right) \;\middle|\; t \in [0,1] \right\}.$$

To identify the set $X \cap (D_2 \setminus D_1)$, next we compute the sets $KKT_1(x(t))$ for $t \in [0,1]$. It is not hard to see that $1 \in \mathcal{I}_1(x(t))$ if and only if $t \geq 3/\sqrt{2} - 2$. Some more computations show that

$$KKT_1(x(t)) = \left\{ (1-s) \begin{pmatrix} 0 \\ \frac{1-t}{\sqrt{2}} \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 0 \\ 1-t \end{pmatrix} \;\middle|\; s \in [0,1] \right\}$$

for all $t \in [0, 3/\sqrt{2} - 2)$, and

$$KKT_1(x(t)) = \left\{ (1-s) \begin{pmatrix} \frac{1}{2}\left(1 - \frac{3}{2\sqrt{2}} + \frac{t}{2}\right) \\ \frac{1-t}{\sqrt{2}} \\ 0 \end{pmatrix} + s \begin{pmatrix} \frac{1}{2}\left(1 - \frac{3}{2\sqrt{2}} + \frac{t}{2}\right) \\ 0 \\ 1-t \end{pmatrix} \;\middle|\; s \in [0,1] \right\}$$

for all $t \in [3/\sqrt{2} - 2, 1]$. Hence, $KKT_1(x(t))$ contains more than one multiplier for all $t \in [0,1)$, whereas $KKT_1(x(1))$ is a *singleton*. In other words, for $t = 1$, that is, at 'the interior end point of the kink' $x(1) = (1/\sqrt{2}, 1 - 1/\sqrt{2})$, SMFCQ holds at $y_1(x(1))$ in $S(x(1))$ while LICQ is violated. We arrive at

$$X \cap (D_2 \setminus D_1) = \left\{ x(t) := \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} + t \right) \;\middle|\; t \in [0,1) \right\}.$$

In particular, by Corollary 5.4, g_1 is Gâteaux differentiable at $x(1)$, but SMFCQ is *unstable* at $y_1(x(1))$ in $S(x(1))$, as it is violated at $y_1(x(t))$ in $S(x(t))$ with $t < 1$. In the following we shall see that, indeed, g_1 is not Gâteaux differentiable at $x(t)$ with $t < 1$. To this end, we compute the Hadamard directional derivatives of g_1 at $x(t)$ with the formula from Theorem 5.2. The appearing derivatives are

$$\nabla F(x) = 0,\ \nabla_x s_1(x, y_1(x)) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ \nabla_x s_2(x, y_1(x)) = \begin{pmatrix} 2x_1 \\ 0 \end{pmatrix},\ \nabla_x s_3(x, y_1(x)) = \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$

and for $d \in \mathbb{R}^n$, we obtain

$$g_1'(x(t); d) = (1-t) \cdot \begin{cases} (d_1 + d_2), & \text{if } d_1 \leq 0 \\ (-d_1 + d_2), & \text{if } d_1 > 0 \end{cases}$$

for all $t \in [0, 3/\sqrt{2} - 2)$ as well as

$$g_1'(x(t); d) = \left(1 - \frac{3}{2\sqrt{2}} + \frac{t}{2}\right) d_1 - \frac{1}{2}\left(1 - \frac{3}{2\sqrt{2}} + \frac{t}{2}\right) d_2 + (1-t) \cdot \begin{cases} (d_1 + d_2), & \text{if } d_1 \leq 0 \\ (-d_1 + d_2), & \text{if } d_1 > 0 \end{cases}$$

for all $t \in [3/\sqrt{2} - 2, 1)$. This shows that g_1 is not Gâteaux differentiable at $x(t)$ with $t < 1$, but that a 'concave kink' occurs in the graph of g_1 along $X \cap (D_2 \setminus D_1)$. Note that at $x(1)$ we have

$$g_1'(x(1); d) = \frac{3}{2}\left(1 - \frac{1}{\sqrt{2}}\right)\left(d_1 - \frac{d_2}{2}\right)$$

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

We point out that the main argument in the proof of Theorem 5.5 needs Gâteaux differentiability of g_1 not only at the point under consideration, but also on a whole neighborhood. In the present example, Gâteaux differentiability of g_1 at $x(1)$ does not extend to a whole neighborhood. $\Diamond$

The observed differentiability properties in Example 5.7 particularly guarantee that any local minimizer $\bar{x}$ of g_α on X either lies in D_1, or g_α is at least Gâteaux differentiable at $\bar{x}$, where usually even continuous differentiability occurs at $\bar{x}$. In the sequel we will show that, under mild assumptions, this also holds in the general case.

To this end, we will use the linearization cone to $X = \{x \in \mathbb{R}^n | \, s_i(x,x) \leq 0, \, i = 1, \ldots, m\}$ at a point x, which is easily seen to be given by

$$\mathcal{L}_X(x) := \left\{ d \in \mathbb{R}^n \;\middle|\; \big(\nabla_x s_i(x,x) + \nabla_y s_i(x,x)\big)^T d \leq 0, \; \forall i \in I_0(x) \right\}$$

with the active index set $I_0(x) := \{i \in \{1, \ldots, m\} \mid s_i(x,x) = 0\}$. Similar to [22], we define the 'degenerate point set' D_4 as a set of points in D_2 with

$$\operatorname{span}\Big\{\nabla_x s_i\big(x, y_\alpha(x)\big), \; i \in \mathcal{I}_\alpha(x)\Big\} \cap \operatorname{span}\Big\{\nabla_x s_i(x,x) + \nabla_y s_i(x,x), \; i \in I_0(x)\Big\} \neq \{0\}, \tag{27}$$

so

$$D_4 := \big\{ x \in D_2 \;\big|\; (27) \text{ holds for } y_\alpha(x) \in S(x) \big\}.$$

For the next result, we need the following assumption which is not to be confused with LICQ at $y_\alpha(x) \in S(x)$, as here the gradients are taken with respect to x.

Assumption 5.8 *The vectors* $\nabla_x s_i\big(x,y\big)|_{y=y_\alpha(x)}$ $(i \in \mathcal{I}_\alpha(x))$ *are linearly independent for all* $x \in D_2 \setminus (D_1 \cup D_4)$.

Proposition 5.9 *Let Assumptions 1.1, 5.1, and 5.8 hold, and let* $\bar{x} \in D_2 \setminus (D_1 \cup D_4)$. *Then there exists a vector* $d \in \mathbb{R}^n$ *solving the system*

$$g_\alpha'(\bar{x}; d) < 0, \quad \big(\nabla_x s_i(\bar{x},\bar{x}) + \nabla_y s_i(\bar{x},\bar{x})\big)^T d \leq 0, \; i \in I_0(\bar{x}). \tag{28}$$

Proof. Assume that (28) does not possess a solution $d \in \mathbb{R}^n$. By Theorem 5.2 this implies the inconsistency of

$$\Big(F(\bar{x}) - \big(\nabla F(\bar{x}) - \alpha I\big)\big(y_\alpha(\bar{x}) - \bar{x}\big) - \sum_{i \in \mathcal{I}_\alpha(\bar{x})} \lambda^i \nabla_x s_i\big(\bar{x}, y_\alpha(\bar{x})\big)\Big)^T d < 0,$$

$$\big(\nabla_x s_i(\bar{x},\bar{x}) + \nabla_y s_i(\bar{x},\bar{x})\big)^T d \le 0,\ \ i \in I_0(\bar{x}),$$

for any $\lambda \in KKT_\alpha(\bar{x})$. By the Lemma of Farkas, this system is inconsistent if and only if there exist scalars $\gamma_i(\lambda) \ge 0$, $i \in I_0(\bar{x})$, with

$$\begin{aligned} &F(\bar{x}) - \big(\nabla F(\bar{x}) - \alpha I\big)\big(y_\alpha(\bar{x}) - \bar{x}\big) - \sum_{i \in \mathcal{I}_\alpha(\bar{x})} \lambda^i \nabla_x s_i\big(\bar{x}, y_\alpha(\bar{x})\big) \\ &+ \sum_{i \in I_0(\bar{x})} \gamma_i(\lambda)\big(\nabla_x s_i(\bar{x},\bar{x}) + \nabla_y s_i(\bar{x},\bar{x})\big) = 0. \end{aligned} \tag{29}$$

Because of $\bar{x} \in D_2 \setminus D_1$, there exist two different multipliers $\hat{\lambda} \neq \tilde{\lambda}$ with $\hat{\lambda},\ \tilde{\lambda} \in KKT_\alpha(\bar{x})$. Then equation (29) holds for $\lambda = \hat{\lambda}$ as well as for $\lambda = \tilde{\lambda}$. Subtracting and rearranging these two equations leads to

$$\sum_{i \in I_0(\bar{x})} \Big(\gamma_i(\hat{\lambda}) - \gamma_i(\tilde{\lambda})\Big)\big(\nabla_x s_i(\bar{x},\bar{x}) + \nabla_y s_i(\bar{x},\bar{x})\big) = \sum_{i \in \mathcal{I}_\alpha(\bar{x})} (\hat{\lambda}^i - \tilde{\lambda}^i)\nabla_x s_i\big(\bar{x}, y_\alpha(\bar{x})\big),$$

where the left hand side is some element of

$$\text{span}\Big\{\nabla_x s_i(\bar{x},\bar{x}) + \nabla_y s_i(\bar{x},\bar{x}),\ i \in I_0(\bar{x})\Big\},$$

and the right hand side is some element of

$$\text{span}\Big\{\nabla_x s_i\big(\bar{x}, y_\alpha(\bar{x})\big),\ i \in \mathcal{I}_\alpha(\bar{x})\Big\}.$$

The right hand side cannot be trivial in view of $\hat{\lambda} \neq \tilde{\lambda}$ and Assumption 5.8. Hence, (27) holds, which is a contradiction to $\bar{x} \in D_2 \setminus D_4$. Therefore, our assumption is wrong, and there exists a vector $d \in \mathbb{R}^n$ solving the system (28). □

Before we present the main result of this section, we recall that the tangent (or contingent or Bouligand) cone to X at point x is defined by

$$\mathcal{T}_X(x) := \Big\{d \in \mathbb{R}^n \ \Big|\ \exists t_k \searrow 0,\ d^k \to d:\ x + t_k d^k \in X \text{ for all } k \in \mathbb{N}\Big\}.$$

It is well-known that the relation $\mathcal{T}_X(x) \subseteq \mathcal{L}_X(x)$ always holds (see, e.g., [42]), and the Abadie constraint qualification (ACQ) is said to hold at $x \in X$ if $\mathcal{T}_X(x) = \mathcal{L}_X(x)$.

Assumption 5.10 *The ACQ holds for all $x \in D_2 \setminus (D_1 \cup D_4)$.*

Theorem 5.11 *Let Assumptions 1.1, 5.1, 5.8 and 5.10 hold. Then any local minimizer $\bar{x}$ of g_α on X either lies in $D_1 \cup D_4$, or g_α is at least Gâteaux differentiable at $\bar{x}$. If, in the latter case, LICQ holds at $y_\alpha(\bar{x}) \in S(\bar{x})$, then g_α is continuously differentiable at $\bar{x}$.*

Proof. Let $\bar{x}$ be a local minimizer of g_α on X. We distinguish the cases $\bar{x} \in D_2$ and $\bar{x} \in X \setminus D_2$.

First, let $\bar{x} \in D_2$. Then either $\bar{x} \in D_1 \cup D_4$ or, by Proposition 5.9, there exists a vector $d \in \mathbb{R}^n$ solving the system (28). We shall show that the latter leads to a contradiction. In fact, because of $\big(\nabla_x s_i(\bar{x}, \bar{x}) + \nabla_y s_i(\bar{x}, \bar{x})\big)^T d \leq 0$ for all $i \in I_0(\bar{x})$, this d is an element of the linearization cone $\mathcal{L}_X(\bar{x})$. Due to Assumption 5.10, d also belongs to the tangent cone $\mathcal{T}_X(\bar{x})$. Hence, there exist sequences $t_k \searrow 0$ and $d^k \to d$ with $\bar{x} + t_k d^k \in X$ for all $k \in \mathbb{N}$. As $\bar{x}$ is a local minimizer of g_α on X, we have $g_\alpha(\bar{x} + t_k d^k) \geq g_\alpha(\bar{x})$ and

$$\frac{g_\alpha(\bar{x} + t_k d^k) - g_\alpha(\bar{x})}{t_k} \geq 0 \tag{30}$$

for all sufficiently large $k \in \mathbb{N}$. By Theorem 5.2, the function g_α is Hadamard directionally differentiable at $\bar{x}$. Hence, the limit of the left-hand side in (30) exists and is equal to $g'_\alpha(\bar{x}, d)$ (note that just directionally differentiability in the ordinary sense is not sufficient for this implication). Consequently, it holds $g'_\alpha(\bar{x}, d) \geq 0$. This is a contradiction to (28).

In the second case, let $\bar{x} \in X \setminus D_2$. In view of Corollary 5.4 and (3), g_α is Gâteaux differentiable at $\bar{x}$. This completes the proof of the first part of the assertion.

The second part immediately follows from Theorem 5.5. □

Corollary 5.12 *Let Assumptions 1.1, 5.1, 5.8 hold, and assume that all constraint functions s_i are linear. Then any local minimizer $\bar{x}$ of g_α on X either lies in $D_1 \cup D_4$, or the function g_α is at least Gâteaux differentiable at $\bar{x}$. If, in the latter case, LICQ holds at $y_\alpha(\bar{x}) \in S(\bar{x})$, then g_α is continuously differentiable at $\bar{x}$.*

Proof. Due to linearity of all constraint functions s_i, the ACQ holds everywhere in X (see, e.g., [42]). Then Theorem 5.11 yields the statements. □

6 Numerical Results

This section presents numerical results for the solution of QVIs based on the optimization reformulation

$$\min_x \; g_\alpha(x) \quad \text{s.t.} \quad x \in X \tag{31}$$

from Proposition 2.2, where g_α denotes the regularized gap function and X is the feasible set of the QVI, cf. (2). In order to apply suitable standard software to this problem, we have to distinguish two cases: First, we have a QVI with a generalized moving set in which case (31) represents a smooth (continuously differentiable) optimization problem. Second, if the constraints are not given by a generalized moving set, g_α is not necessarily everywhere continuously differentiable, although our analysis shows that, also in this case, except for some pathological situations, we can expect differentiability at all locally minimal points.

Since, for the nondifferentiable case, numerical results are presented in the previous paper [22] for the special case of generalized Nash equilibrium problems, we decided to concentrate on QVIs defined by generalized moving sets in this section. More precisely, we consider both QVIs with (standard) moving sets and QVIs with generalized moving sets as defined in Section 3.1.

To this end, we recall that the generalized gap function g_α is well defined for all $x \in \mathbb{R}^n$ in the moving and generalized moving set cases whenever $K \neq \emptyset$. This observation is important since this allows to apply software that might generate non-feasible iterates. In particular, this enables us to use the TOMLAB/SNOPT 7.2-9 solver as the working horse for problem (31), especially since this method does not use any second-order derivatives. However, we compare the results also with the TOMLAB/KNITRO 8.0.0 solver applied to (31) although, formally, this solver uses second-order information and, therefore, is not a feasible method in our case since the regularized gap function g_α may not be twice continuously differentiable everywhere. For more information about TOMLAB/SNOPT and TOMLAB/KNITRO, we refer to the TOMLAB/SNOPT and TOMLAB /KNITRO User Guides on the web sites `http://tomopt.com/tomlab/products/snopt/` and `http://tomopt.com/tomlab/products/knitro/`, respectively.

For both solvers, we provide the starting point x^0 as well as the function and gradient values (including the derivative of g_α from (12)) for each test problem. Moreover, for KNITRO, we use the active set Sequential Linear-Quadratic Programming (SLQP) optimizer by setting `Prob.KNITRO.options.ALG=3`. Apart from this, all standard options are taken for both methods. Our implementation uses the regularization parameter $\alpha = 1$ for all test problems.

We use two groups of test examples: The first group consists of all the QVIs with (standard) moving sets from the recent test problem collection [14] (called MovSet*). For the second group, we modify these test problems to QVIs with generalized moving sets (called GenMovSet*) defined by the diagonal matrix $Q(x) = \text{diag}\left(\frac{1}{x_1^2+1}, \dots, \frac{1}{x_n^2+1}\right)$. The corresponding numerical results for the first group are presented in Table 1, whereas Table 2 contains the numerical results for the second group.

For each test example, Tables 1 and 2 contain the following data: The name of the example, the number of variables n, the number of constraints s_i, $i = 1, \dots, m$, the starting point x^0 (all components of this starting point are equal to the number given here), and for both solvers the number of iterations k needed until convergence and the final value of the generalized gap function g_α in column g_α^{opt} (whenever a solution was found). Here, the starting points in Table 1 are those taken from the paper [14] and implemented in the corresponding M-file `startingPoints.m`. The same starting points are used for the generalized moving set examples. The results for examples MovSet4* and GenMovSet4* with the starting point equal to the zero vector (as suggested in [14]) are not contained in Tables 1 and 2 since the zero vector turned out to be a solution of these test problems and are immediately identified as such from both solvers.

Tables 1 and 2 show that all test examples can be solved within a very reasonable number of iterations except for examples MovSet2B and GenMovSet2B with the second

Ex.	n	m	x^0	SNOPT Solver		KNITRO Solver	
				k	g_α^{opt}	k	g_α^{opt}
MovSet1A	5	1	0	9	8.119032e-09	6	1.771996e-09
			10	14	8.168694e-09	8	3.695276e-11
MovSet1B	5	1	0	57	-1.455251e-09	7	5.913887e-10
			10	89	-4.106141e-08	16	5.888718e-10
MovSet2A	5	1	0	9	3.127895e-13	5	4.689504e-10
			10	18	-1.963065e-11	9	4.697078e-10
MovSet2B	5	1	0	35	3.129177e-09	9	-1.499496e-05
			10	–	failure	–	failure
MovSet3A1	1000	1	0	55	1.542633e-06	6	-1.572717e-09
			10	54	1.542746e-06	11	1.503841e-09
MovSet3B1	1000	1	0	57	5.208333e-08	7	4.823794e-10
			10	56	5.211922e-08	12	4.416943e-10
MovSet3A2	2000	1	0	64	4.339869e-11	7	1.318250e-11
			10	63	3.043553e-11	11	1.420134e-11
MovSet3B2	2000	1	0	63	1.095111e-07	7	1.616324e-11
			10	63	1.095374e-07	13	9.701867e-11
MovSet4A1	400	801	10	3	4.216834e-12	3	5.494870e-13
MovSet4B1	400	801	10	3	3.046541e-12	3	-1.763913e-13
MovSet4A2	800	1601	10	4	2.139371e-12	3	7.364564e-13
MovSet4B2	800	1601	10	4	-2.618998e-13	3	8.076459e-13

Table 1: Table with numerical results for QVIs with moving sets from paper [14]

Ex.	n	m	x^0	SNOPT Solver		KNITRO Solver	
				k	g_α^{opt}	k	g_α^{opt}
GenMovSet1A	5	1	0	10	-8.048828e-13	6	2.996280e-08
			10	18	4.050013e-12	13	2.996280e-08
GenMovSet1B	5	1	0	21	-1.286942e-02	11	7.618806e-06
			10	18	-1.853720e-04	15	7.806321e-06
GenMovSet2A	5	1	0	8	1.976154e-11	6	7.765171e-09
			10	18	-3.330922e-10	10	7.763598e-09
GenMovSet2B	5	1	0	28	1.352352e-09	14	1.985551e-06
			10	–	failure	–	failure
GenMovSet3A1	1000	1	0	29	5.991367e-10	8	9.817330e-12
			10	42	6.008491e-10	17	3.991185e-10
GenMovSet3B1	1000	1	0	31	3.184530e-11	8	2.014215e-10
			10	43	3.388897e-11	17	1.906384e-10
GenMovSet3A2	2000	1	0	34	1.226018e-09	9	4.932545e-10
			10	51	1.221392e-09	16	-3.373292e-08
GenMovSet3B2	2000	1	0	36	7.742417e-11	8	-5.451358e-11
			10	59	6.534881e-11	18	-6.147936e-10
GenMovSet4A1	400	801	10	12	5.694374e-03	10	1.327288e-08
GenMovSet4B1	400	801	10	12	4.728919e-03	10	1.370384e-08
GenMovSet4A2	800	1601	10	13	6.742428e-14	10	2.652667e-08
GenMovSet4B2	800	1601	10	12	1.069513e-02	10	2.810001e-08

Table 2: Table with numerical results for QVIs with generalized moving sets

k	x^k	$g_\alpha(x^k)$	g_α counts
0	(10, 10)	1.274434e+01	1
1	(9.84901583, 9.84901583)	3.854426e-01	3
2	(9.82327425, 9.82327425)	1.297152e-02	5
3	(9.81753271, 9.81753271)	1.194829e-06	6
4	(9.81747717, 9.81747717)	6.141179e-12	7
5	(9.81747704, 9.81747704)	-7.787916e-20	8

Table 3: Table with numerical results for Example 3.6

starting point. These tables also indicate that the number of iterations needed by KNITRO is sometimes significantly smaller than the corresponding numbers for SNOPT. A possible explanation might be the fact that KNITRO uses second-order information. We also believe that this fact is responsible for the higher accuracy that is sometimes obtained by the KNITRO solver. In fact, SNOPT terminates for three of the four test examples called GenMovSet4* with the function value of g_α being around $10^{-2} - 10^{-3}$, whereas KNITRO is able to get much closer to zero. Nevertheless, the termination by SNOPT was successful in the sense that the standard stopping criteria of this solver were reached.

Note also that, in some cases, upon termination we have a negative function value g_α^{opt} in the corresponding columns of Tables 1 and 2. These negative values arise for two reasons: First, if the final iterate x^k is slightly outside the feasible region, then g_α might be negative. Second, negative values may arise due to inexact function evaluations (recall that the evaluation of g_α at a point x requires the solution of an optimization problem which, fortunately, automatically also gives the gradient $\nabla g_\alpha(x)$).

Finally, in Table 3, we come back to our Example 3.6 and present the corresponding iteration history, with all calculations being done by SNOPT. More precisely, for each iteration k, Table 3 provides the iteration vector x^k, the value of g_α at x^k as well as the cumulated number of evaluations of the mapping g_α. Table 3 illustrates that the calculation of a solution for the starting point $x^0 = (10, 10)$ finishes successfully and has a fast local convergence rate. We also tried a number of different starting points, and were always able to find a solution up to the required accuracy. Note, however, that Example 3.6 has infinitely many solutions, hence the method finds different solutions when using different starting points.

7 Final Remarks

This paper studied smoothness properties of a regularized gap function for QVIs as well as connections between QVIs and GNEPs. While, under general convexity assumptions and except for pathological cases, continuous differentiability of the regularized gap function was shown at all locally minimal points of the optimization reformulation of the QVI, the concept of generalized moving sets even allowed to show continuous differentiability of the regularized gap function on its whole domain. Our numerical results cover the latter case, as we treated the first case for GNEPs already in [22].

We believe that, under stronger convexity assumptions, also the directional differentiability behaviour of the regularized gap function on the degenerate point set D_1 may be understood which would lead to an improvement of Theorem 5.11. On the other hand, under weaker convexity assumptions as, for example, quasi-convexity of the functions s_i, $i = 1, \ldots, m$, most of the results shown in this article may still be valid. We leave these questions for future research.

References

[1] G. Auchmuty: *Variational principles for variational inequalities.* Numerical Functional Analysis and Optimization 10, 1989, pp. 863–874.

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

[3] D. Aussel, R. Correa, and M. Marechal: *Gap functions for quasivariational inequalities and generalized Nash equilibrium problems.* Journal of Optimization Theory and Applications 151, 2011, pp. 474–488.

[4] C. Baiocchi and A. Capelo: *Variational and Quasivariational Inequalities: Applications to Free Boundary Problems.* Wiley, New York, 1984.

[5] A. Bensoussan, M. Goursat, and J.-L. Lions: *Contrôle impulsionnel et inéquations quasi-variationnelles stationnaires.* C. R. Acad. Sci. Paris Sér. A 276, 1973, pp. 1279–1284.

[6] A. Bensoussan and J.-L. Lions: *Nouvelle formulation de problèmes de contrôle impulsionnel et applications.* C. R. Acad. Sci. Paris Sér. A 276, 1973, pp. 1189–1192.

[7] A. Bensoussan and J.-L. Lions: *Nouvelles méthodes en contrôle impulsionnel.* Applied Mathematics and Optimization 1, 1975, pp. 289–312.

[8] D. Chan and J.-S. Pang: *The generalized quasi-variational inequality problem.* Mathematics of Operations Research 7, 1982, pp. 211–222.

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

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

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

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

[13] F. Facchinei, C. Kanzow, and S. Sagratella: *Solving quasi-variational inequalities via their KKT conditions.* Mathematical Programming, DOI:10.1007/s10107-013-0637-0.

[14] F. Facchinei, C. Kanzow, and S. Sagratella: *QVILIB: A library of quasi-variational inequality test problems.* Pacific Journal of Optimization, to appear.

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

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

[17] M. Fukushima: *A class of gap functions for quasi-variational inequality problems.* Journal of Industrial and Management Optimization 3, 2007, pp. 165–171.

[18] J. Gauvin, *A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming.* Mathematical Programming 12, 1977, pp. 136–138.

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

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

[21] P.T. Harker: *Generalized Nash games and quasi-variational inequalities.* European Journal of Operational Research 54, 1991, pp. 81–94.

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

[23] D.W. Hearn: *The gap function of a convex program.* Operations Research Letters 1, 1982, pp. 67–71.

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

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

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

[27] J. Kyparisis: *On uniqueness of Kuhn-Tucker multipliers in nonlinear programming.* Mathematical Programming 32, 1985, pp. 242–246.

[28] P. Marcotte and J.P. Dussault: *A sequential linear programming algorithm for solving monotone variational inequalities.* SIAM Journal on Control and Optimization 27, 1989, pp. 1260–1278.

[29] U. Mosco: *Implicit variational problems and quasi variational inequalities.* In: J. Gossez, E. Lami Dozo, J. Mawhin, and L. Waelbroeck (eds.): *Nonlinear Operators and the Calculus of Variations.* Lecture Notes in Mathematics Vol. 543, Springer, 1976, pp. 83–156.

[30] Y. Nesterov and L. Scrimali: *Solving strongly monotone variational and quasi-variational inequalities.* CORE Discussion Paper 2006/107, Catholic University of Louvain, Center for Operations Research and Econometrics, 2006.

[31] M.A. Noor: *An iterative scheme for a class of quasi variational inequalities.* Journal of Mathematical Analysis and Applications 110, 1985, pp. 463–468.

[32] M.A. Noor: *On merit functions for quasivariational inequalities.* Journal of Mathematical Inequalities 1, 2007, pp. 259–268.

[33] M.A. Noor: *On general quasi-variational inequalities.* J. King Saud University 24, 2012, pp. 81–88.

[34] J.V. Outrata and M. Kocvara: *On a class of quasi-variational inequalities.* Optimization Methods and Software 5, 1995, pp. 275–295.

[35] J.V. Outrata, M. Kocvara, and J. Zowe: *Nonsmooth approach to optimization problems with equilibrium constraints.* Kluwer Academic Publishers, Dordrecht and Boston, 1998.

[36] J.V. Outrata and J. Zowe: *A Newton method for a class of quasi-variational inequalities.* Computational Optimization and Applications 4, 1995, pp. 5–21.

[37] J.-S. Pang and M. Fukushima: *Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games.* Computational Management Science 2, 2005, pp. 21–56 (Erratum: ibid 6, 2009, pp. 373–375).

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

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

[40] I.P. Ryazantseva: *First-order methods for certain quasi-variational inequalities in Hilbert space.* Computational Mathematics and Mathematical Physics 47, 2007, pp. 183–190.

[41] A.H. Siddiqi and Q.H. Ansari: *Strongly nonlinear quasivariational inequalities.* Journal of Mathematical Analysis and Applications 149, 1990, pp. 444–450.

[42] O. Stein: *On constraint qualifications in non-smooth optimization.* Journal of Optimization Theory and Applications 121, 2004, pp. 647–671.

[43] K. Taji: *On gap functions for quasi-variational inequalities.* Abstract and Applied Analysis, 2008, Article ID 531361.

[44] K. Taji and M. Fukushima: *A new merit function and a successive quadratic programming algorithm for variational inequality problems.* SIAM Journal on Optimization 6, 1996, pp. 704–713.

[45] K. Taji, M. Fukushima, and T. Ibaraki: *A globally convergent Newton method for solving strongly monotone variational inequalities.* Mathematical Programming 58, 1993, pp. 369–383.