ON ERROR BOUNDS AND NEWTON-TYPE METHODS
FOR GENERALIZED NASH EQUILIBRIUM PROBLEMS[*]

A. F. Izmailov[†] and M. V. Solodov[‡]

December 26, 2012 (revised April 14, 2013)

ABSTRACT

Error bounds (estimates for the distance to the solution set of a given problem) are key to analyzing convergence rates of computational methods for solving the problem in question, or sometimes even to justifying convergence itself. That said, for the generalized Nash equilibrium problems (GNEP), the theory of error bounds had not been developed in depth comparable to the fields of optimization and variational problems. In this paper, we provide a systematic approach which should be useful for verifying error bounds for both specific instances of GNEPs and for classes of GNEPs. These error bounds for GNEPs are based on more general results for constraints that involve complementarity relations (of independent interest) and cover those (few) GNEP error bounds that existed previously, and go beyond. In addition, they readily imply a Lipschitzian stability result for solutions of GNEPs, a subject where again very little had been known. As a specific application of error bounds, we discuss Newtonian methods for solving GNEPs. While we do not propose any significantly new methods in this respect, some new insights into applicability to GNEPs of various approaches and into their convergence properties are presented.

Dedicated to Professor Masao Fukushima, on the occasion of his 65th birthday.

[†] VMK Faculty, OR Department, Moscow State University, MSU, Uchebniy Korpus 2, Leninskiye Gory, 119991 Moscow, Russia.
Email: `izmaf@ccas.ru`
[‡] IMPA – Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ 22460-320, Brazil.
Email: `solodov@impa.br`

1   Introduction

For simplicity of presentation, we shall consider the generalized Nash equilibrium problem (GNEP) with two players and shared constraints only:

$$
\begin{array}{ll}
\text{minimize}_{x^1} \quad f_1(x^1,\, x^2) & \qquad \text{minimize}_{x^2} \quad f_2(x^1,\, x^2) \\
\text{subject to} \quad g(x^1,\, x^2) \leq 0, & \qquad \text{subject to} \quad g(x^1,\, x^2) \leq 0,
\end{array}
\tag{1.1}
$$

where the players' objective functions $f_1 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$ and $f_2 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$ and the constraint mapping $g : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^m$ are smooth. This simplified setting requires lighter notation, while being sufficient to state and explain our results. Extentions of all the statements to the case of more than two players, and/or to the case when individual (not shared) constraints are present, are straightforward. GNEP setting goes back to [35]; for a modern survey, see [14]. Some other relevant references on theory and numerical methods for GNEP are [20, 19, 11, 12, 13, 26, 32, 8, 21, 27].

As is well known, error bounds (i.e., estimates for the distance to the solution set of a given problem) are important for a great variety of reasons, among which are analysis and design of computational algorithms, stability/sensitivity, etc. For general references on error bounds, conditions required for them to hold, and their roles in optimization and variational problems, see, e.g., [33, 36] and [15, Ch. 6]. As commented in [12, Sections 3.3, 4.3] and [14, Section 5.6], error bounds for GNEP are not as well developed yet. This paper suggests a simple and natural way to deriving error bounds for GNEP under reasonable assumptions. Our approach is based on a more general result for constraint systems that involve complementarity relations, which is of independent interest. The obtained error bounds are further applied to the analysis of some Newton-type methods for GNEP.

One approach to (1.1) is to solve, simultaneously, the Karush–Kuhn–Tucker (KKT) optimality systems for the two optimization problems in (1.1); see [14]. I.e., the task is to solve the following system in the primal-dual space:

$$
\frac{\partial L_1}{\partial x^1}(x^1,\, x^2,\, \mu^1) = 0, \quad \frac{\partial L_2}{\partial x^2}(x^1,\, x^2,\, \mu^2) = 0,
$$
$$
\mu^1 \geq 0, \quad \langle \mu^1,\, g(x^1,\, x^2) \rangle = 0, \quad \mu^2 \geq 0, \quad \langle \mu^2,\, g(x^1,\, x^2) \rangle = 0, \quad g(x^1,\, x^2) \leq 0,
\tag{1.2}
$$

where $L_j : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \to \mathbb{R}$ is the Lagrangian of the corresponding optimization problem in (1.1), i.e.,

$$
L_j(x^1,\, x^2,\, \mu^j) = f_j(x^1,\, x^2) + \langle \mu^j,\, g(x^1,\, x^2) \rangle, \quad j = 1,\, 2.
$$

(Here, and throughout the rest of the paper, we use the notation $\frac{\partial}{\partial x^j}$ for the the partial derivative with respect to x^j. For instance, $\frac{\partial g}{\partial x^1}(x^1,\, x^2)$ is the Jacobian of the mapping $g(\cdot,\, x^2)$ at x^1.)

The following relations between the (primal) solutions of GNEP (1.1) and (primal-dual) solutions of system (1.2) are well-known [14, Theorem 4.6]. If $(\bar{x}^1,\, \bar{x}^2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ is an equilibrium, and the constraints of the first problem in (1.1) with $x^2 = \bar{x}^2$ satisfy an appropriate constraint qualification (CQ) [36] at $\bar{x}^1$, while the constraints of the second problem in (1.1) with $x^1 = \bar{x}^1$ satisfy a CQ at $\bar{x}^2$, then there exists $(\bar{\mu}^1,\, \bar{\mu}^2) \in \mathbb{R}^m \times \mathbb{R}^m$

such that $(\bar{x}^1, \bar{x}^2, \bar{\mu}^1, \bar{\mu}^2)$ solves (1.2). Conversely, if $f_1(\cdot, x^2)$ and the components of $g(\cdot, x^2)$ are convex for each $x^2 \in \mathbb{R}^{n_2}$, while $f_2(x^1, \cdot)$ and the components of $g(x^1, \cdot)$ are convex for each $x^1 \in \mathbb{R}^{n_1}$, then for every solution $(\bar{x}^1, \bar{x}^2, \bar{\mu}^1, \bar{\mu}^2)$ of (1.2) it holds that $(\bar{x}^1, \bar{x}^2)$ is an equilibrium. In particular, under the appropriate assumptions, by estimating the distance to the solution set of (1.2) one also estimates the distance to the solution set of GNEP (1.1).

If at a primal solution $\bar{x}$ of (1.2) we have coinciding Lagrange multipliers $\bar{\mu}^1 = \bar{\mu}^2$, this solution is called *variational equilibrium.* This is a simpler but important case, which indeed occurs in various practical models [14]; see also [28] for an example. Observe that imposing $\mu^1 = \mu^2$, (1.2) becomes an instance of the mixed complementarity problem [15] whose solutions can be expected to be isolated. In this situation, the error bounds and methods developed for this problem class (e.g., in [24, 17, 23, 2, 4, 5, 6]) are readily adaptable. We shall not deal with this case here, even though some details of such an adaptation should be of interest. In this paper, we consider the general case when the multipliers need not coincide.

The rest of the paper is organized as follows. In Section 2 we provide an error bound result for general constraint systems that involve complementarity relations, showing in particular that piecewise error bounds imply the error bound for the full system in terms of its natural residual. In Section 2, we first put in evidence that the weakest general CQ that is currently known to imply an error bound for "generic" (without special structure) systems – the relaxed constant positive linear dependence condition [1] – cannot be expected to hold for the KKT systems associated to GNEP. Thus, a special analysis is needed. Using the result for systems with general complementarities, we then show that for the GNEP KKT systems, an error bound holds under some very reasonable assumptions (in particular, not involving strict complementarity). The obtained error bound further implies upper Lipschitz stability of solutions of GNEP KKT systems under smooth perturbations. In Section 4 we discuss some consequences of our error bound results for the convergence of Newtonian methods for GNEP. Among other things, we show applicability to GNEP of the algorithms for nonsmooth constrained equations with nonisolated solutions recently developed in [9, 10].

We now describe some of our notation and recall some CQs that will be of relevance for subsequent discussions. Given a vector $z \in \mathbb{R}^s$ and an index set $I \subset \{1, \ldots, s\}$, by $z_I \in \mathbb{R}^{|I|}$ we mean the vector comprised by the components of z indexed by I, where $|I|$ is the cardinality of I. Similarly, for a matrix M with s rows, M_I stands for the matrix given by the rows of M indexed by I.

Consider the constraint system

$$a(u) = 0, \quad b(u) \leq 0,$$

where $a : \mathbb{R}^p \to \mathbb{R}^q$ and $b : \mathbb{R}^p \to \mathbb{R}^s$ are sufficiently smooth, and let D be its solution set. For any $\bar{u} \in D$, define the set $I(\bar{u}) = \{i = 1, \ldots, s \mid b_i(\bar{u}) = 0\}$ of active inequality constraints. The *Mangasarian–Fromovitz CQ* (MFCQ) [30] holds at $\bar{u}$ if the system

$$(a'(\bar{u}))^{\mathrm{T}}\eta + (b'_{I(\bar{u})}(\bar{u}))^{\mathrm{T}}\tilde{\zeta} = 0, \quad \tilde{\zeta} \geq 0,$$

in the variables $(\eta, \tilde{\zeta}) \in \mathbb{R}^q \times \mathbb{R}^{|I(\bar{u})|}$ has only the zero solution (this is the dual form of MFCQ). (Here, and throughout the paper, prime denotes the Jacobian of the mapping in

2

question.) Clearly, MFCQ is implied by the *linear independence CQ* (LICQ), which consists of saying that the set $\{a'_j(\bar{u}) \mid j = 1, \ldots, q\} \cup \{b'_i(\bar{u}) \mid i \in I(\bar{u})\}$ is linearly independent.

Let $J \subset \{1, \ldots, q\}$ be such that $\{a'_j(\bar{u}) \mid j \in J\}$ is a basis (any basis) in the linear subspace spanned by $\{a'_j(\bar{u}) \mid j = 1, \ldots, q\}$. The *relaxed constant positive linear dependence constraint qualification* (RCPLD) [1] holds at $\bar{u} \in D$ if there exists a neighborhood U of $\bar{u}$ such that:

1. It holds that rank $a'(u)$ is constant for all $u \in U$.

2. For every $I \subset I(\bar{u})$, if there exist $\tilde{\eta} \in \mathbb{R}^{|J|}$ and $\tilde{\zeta} \in \mathbb{R}^{|I|}$, not all equal to zero and such that

$$(a'_J(\bar{u}))^{\mathrm{T}}\tilde{\eta} + (b'_I(\bar{u}))^{\mathrm{T}}\tilde{\zeta} = 0, \quad \tilde{\zeta} \geq 0, \tag{1.3}$$

   then

$$\{a'_j(u) \mid j \in J\} \cup \{b'_i(u) \mid i \in I\} \text{ is linearly dependent } \forall\, u \in U. \tag{1.4}$$

This condition is a relaxed version of the constant positive linear dependence CQ (CPLD) [34], which consists of saying that there exists a neighborhood U of $\bar{u}$ such that for any $J \subset \{1, \ldots, q\}$ and $I \subset I(\bar{u})$, if there exist $\tilde{\eta} \in \mathbb{R}^{|J|}$ and $\tilde{\zeta} \in \mathbb{R}^{|I|}$ not all equal to zero and such that (1.3) holds, then (1.4) holds as well. It is easy to see that MFCQ implies CPLD and, hence, RCPLD. Also, CPLD (and thus RCPLD) is implied by the well-known constant rank CQ (CRCQ) [22], and even by the relaxed constant rank CQ (RCRCQ) [31].

Note that RCPLD guarantees [1] that under twice differentiability of a and b near $\bar{u}$, the following error bound holds:

$$\mathrm{dist}(u, D) = O(\|a(u)\| + \|\max\{0, b(u)\}\|)$$

as $u \to \bar{u}$, where the max-operation is understood componentwise.

2 An error bound for constraint systems with complementarities

We start with a lemma concerned with error bounds for general constraint systems invloving complementarity relations:

$$F(u) = 0, \quad g(u) \leq 0, \quad G(u) \geq 0, \quad H(u) \geq 0, \quad \langle G(u), H(u)\rangle = 0, \tag{2.1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^l$, $g : \mathbb{R}^n \to \mathbb{R}^m$, $G : \mathbb{R}^n \to \mathbb{R}^s$ and $H : \mathbb{R}^n \to \mathbb{R}^s$. In particular, the feasible set here is as in mathematical programs with complementarity constraints (MPCC) [29]. Clearly, (1.2) is a special case of (2.1).

Lemma 2.1 below shows that piecewise error bounds for the branches of the set D given by (2.1) imply an error bound for D in terms of the natural residual of its constraints. This appears to be a rather natural observation, but we could not find it stated in MPCC literature. Below, it will be applied to derive error bounds for GNEP KKT systems (1.2). We note that it is also of an independent interest, as it should be useful for MPCC itself. As MPCC is not the subject of this paper, we mention only one example – Proposition 2.1 below – on exact penalization for MPCC.

Lemma 2.1 *Let $F : \mathbb{R}^n \to \mathbb{R}^l$, $g : \mathbb{R}^n \to \mathbb{R}^m$, $G : \mathbb{R}^n \to \mathbb{R}^s$ and $H : \mathbb{R}^n \to \mathbb{R}^s$ be continuous at $\bar{u} \in \mathbb{R}^n$ satisfying the complementarity constraint system (2.1), and let D stand for the solution set of this system. Set*

$$I_G = I_G(\bar{u}) = \{i = 1, \ldots, s \mid G_i(\bar{u}) = 0\}, \quad I_H = I_H(\bar{u}) = \{i = 1, \ldots, s \mid H_i(\bar{u}) = 0\}, \quad (2.2)$$

$$I_0 = I_G \cap I_H. \tag{2.3}$$

Assume that for each partition (I_1, I_2) of I_0 (i.e., $I_1 \cup I_2 = I_0$, $I_1 \cap I_2 = \emptyset$) the error bound

$$
\begin{aligned}
\mathrm{dist}(u, D) \;=\; & O\left(\|F(u)\| + \|\max\{0, g(u)\}\| + \|G_{I_G \setminus I_H}(u)\| + \|H_{I_H \setminus I_G}(u)\| \right. \\
& \left. + \|G_{I_1}(u)\| + \|H_{I_2}(u)\| + \|\min\{0, G_{I_2}(u)\}\| + \|\min\{0, H_{I_1}(u)\}\|\right)
\end{aligned}
$$
$$\tag{2.4}$$

holds as $u \to \bar{u}$.
    Then the error bound

$$\mathrm{dist}(u, D) = O(\|F(u)\| + \|\max\{0, g(u)\}\| + \|\min\{G(u), H(u)\}\|) \tag{2.5}$$

holds as $u \to \bar{u}$.

Proof.   Observe first that if $u \in \mathbb{R}^n$ is close enough to $\bar{u}$, then (by the continuity of the functions involved) it holds that

$$G_i(u) = \min\{G_i(u), H_i(u)\} \quad \forall i \in I_G \setminus I_H, \quad H_i(u) = \min\{G_i(u), H_i(u)\} \quad \forall i \in I_H \setminus I_G. \tag{2.6}$$

Moreover, for any $u \in \mathbb{R}^n$ it evidently holds that

$$|\min\{0, G_i(u)\}| \le |\min\{G_i(u), H_i(u)\}|, \quad |\min\{0, H_i(u)\}| \le |\min\{G_i(u), H_i(u)\}| \tag{2.7}$$
$$\forall i = 1, \ldots, s.$$

Considering any $u \in \mathbb{R}^n$, we now define the corresponding partition (I_1, I_2) of I_0 as follows: we assign $i \in I_0$ to I_1 if $G_i(u) \le H_i(u)$, and to I_2 otherwise. Then

$$G_i(u) = \min\{G_i(u), H_i(u)\} \quad \forall i \in I_1, \quad H_i(u) = \min\{G_i(u), H_i(u)\} \quad \forall i \in I_2. \tag{2.8}$$

The claimed error bound (2.5) now follows from the error bounds (2.4) by (2.6)–(2.8), and by the fact that the number of different partitions of I_0 is finite.   ∎

The following comments are useful for clarifying the nature of this development, and its possible other applications. Observe that in the right-hand side of (2.4) we have the residual of the constraints

$$F(u) = 0, \;\; g(u) \le 0, \;\; G_{(I_G \setminus I_H) \cup I_1}(u) = 0, \;\; H_{I_1}(u) \ge 0, \;\; G_{I_2}(u) \ge 0, \;\; H_{(I_H \setminus I_G) \cup I_2}(u) = 0, \tag{2.9}$$

which define the branch $D_{(I_1, I_2)}$ of the solution set D of system (2.1). Specifically, locally (near $\bar{u}$) the set D coincides with the union of such branches over all partitions of I_0. It then follows that any CQ [36] which guarantees the error bound

$$
\begin{aligned}
\mathrm{dist}\left(u,\, D_{(I_1,\, I_2)}\right) \;\; = \;\; & O\left(\|F(u)\| + \|\max\{0,\, g(u)\}\| + \|G_{I_G \setminus I_H}(u)\| + \|H_{I_H \setminus I_G}(u)\| \right. \\
& \left. + \|G_{I_1}(u)\| + \|H_{I_2}(u)\| + \|\min\{0,\, G_{I_2}(u)\}\| + \|\min\{0,\, H_{I_1}(u)\}\|\right)
\end{aligned}
\tag{2.10}
$$

for each of the branches (i.e., "piecewise CQ") implies the needed error bound (2.5). For instance, according to this argument, the so-called MPCC-linear independence constraint qualification implies (2.5), since it implies LICQ for each branch. However, it is important to emphasize that various much weaker conditions do the job as well. For example, it is sufficient to assume the piecewise MFCQ, or piecewise (R)CRCQ, or piecewise (R)CPLD; i.e, MFCQ or (R)CRCQ or (R)CPLD for the constraints defining each branch of (2.1) at a feasible $\bar{u} \in \mathbb{R}^n$. Out of all the CQs mentioned, RCPLD of [1] is the weakest, and thus could in principle be assumed for each branch to yield the sharpest condition for the error bound. However, it is still appealing to leave the possibility of *different* CQs holding for different branches. The reason is that although in the *current* state-of-the-art on the subject RCPLD is weaker than the other relevant conditions (MFCQ, RCRCQ, ...) this might change in the future if new CQs are developed. To support this line of thought, it is enough to mention that RCPLD is a new condition which did not exist at the time when [36] was written, for example; and within CQs that give error bounds and are covered in that reference some are independent (neither one implies the other). In principle, it is possible that such a situation may emerge again in the future, and then assuming different CQs for different branches would be useful.

The possibility that complementarity constraints may satisfy some piecewise CQ (weaker than the strongest MPCC-LICQ) is perhaps "geometrically" clear, but specific examples for illustration can also be constructed explicitly as follows. Consider the constraint system

$$
F(u) = 0, \quad h(u) = 0, \quad g(u) \le 0
\tag{2.11}
$$

with some $F : \mathbb{R}^n \to \mathbb{R}^l$, $h : \mathbb{R}^n \to \mathbb{R}^s$ and $g : \mathbb{R}^n \to \mathbb{R}^m$. Suppose that this system satisfies some specific CQ (any of the mentioned above, for example) at its feasible point $\bar{u}$. Define $G : \mathbb{R}^n \to \mathbb{R}^s$ and $H : \mathbb{R}^n \to \mathbb{R}^s$ as follows: $G \equiv (1, \dots, 1)$, $H = h$. Then for the complementarity constraints in (2.1) the point $\bar{u}$ is feasible, and it gives rise to the single valid branch given by the constraint system in (2.11), with this branch satisfying the needed CQ.

As an immediate application of Lemma 2.1 to MPCC, we mention the following exact penalization result. It can be readily justified following the classical exact penalization principle; see, e.g., [29] or other related statements such as [3, Proposition 3.111]. However, just like the error bound of Lemma 2.1 itself, we could not find the statement of Proposition 2.1 on exact penalization in MPCC literature (i.e., this statement under these assumptions; but there are a number of somewhat related results, of course).

Proposition 2.1 *In addition to the assumptions of Lemma 2.1, let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz-continuous at $\bar{u}$, and assume that $\bar{u}$ is a (strict) local solution of the problem*

$$
\begin{aligned}
\text{minimize} \quad & f(u) \\
\text{subject to} \quad & u \in D,
\end{aligned}
$$

where D is the set defined by (2.1).

Then for any $c > 0$ large enough, $\bar{u}$ is a (strict) local solution of the problem

$$
\begin{aligned}
\text{minimize} \quad & f(u) + c(\|F(u)\| + \|\max\{0, g(u)\}\| + \|\min\{G(u), H(u)\}\|) \\
\text{subject to} \quad & u \in \mathbb{R}^n.
\end{aligned}
$$

3   Error bounds and upper Lipschitz stability for GNEP

We now turn our attention to the GNEP joint KKT systems (1.2). For a given solution $(\bar{x}^1, \bar{x}^2, \bar{\mu}^1, \bar{\mu}^2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m$ of (1.2), define the index sets

$$
A = A(\bar{x}^1, \bar{x}^2) = \{i = 1, \ldots, m \mid g_i(\bar{x}^1, \bar{x}^2) = 0\}, \quad N = N(\bar{x}^1, \bar{x}^2) = \{1, \ldots, m\} \setminus A,
$$

$$
A_+^j = A_+^j(\bar{x}^1, \bar{x}^2, \bar{\mu}^j) = \{i \in A \mid \bar{\mu}^j > 0\}, \quad A_0^j = A_0^j(\bar{x}^1, \bar{x}^2, \bar{\mu}^j) = A \setminus A_+^j, \quad j = 1, 2,
$$

$$
A_+ = A_+(\bar{x}^1, \bar{x}^2, \bar{\mu}^1, \bar{\mu}^2) = A_+^1 \cup A_+^2, \quad A_0 = A_0(\bar{x}^1, \bar{x}^2, \bar{\mu}^1, \bar{\mu}^2) = A_0^1 \cap A_0^2.
$$

The first issue to settle is whether some general CQs that imply an error bound can be expected to hold for the GNEP KKT systems (1.2). As discussed above, the weakest CQ giving the error bound that is currently known is RCPLD. Note also that there is no a priori reason that may allow one to claim that RCPLD is not a useful condition for (1.2). However, we next exhibit that RCPLD for (1.2) is an atypical property (see also the model Example 3.1 below, where RCPLD does not hold).

The Jacobian of the equalities in (1.2) has the form

$$
\begin{pmatrix}
\dfrac{\partial^2 L_1}{\partial x^1 \partial x^1}(x^1, x^2, \mu^1) & \dfrac{\partial^2 L_1}{\partial x^1 \partial x^2}(x^1, x^2, \mu^1) & \left(\dfrac{\partial g}{\partial x^1}(x^1, x^2)\right)^{\mathrm{T}} & 0 \\[2ex]
\dfrac{\partial^2 L_2}{\partial x^1 \partial x^2}(x^1, x^2, \mu^2) & \dfrac{\partial^2 L_2}{\partial x^2 \partial x^2}(x^1, x^2, \mu^2) & 0 & \left(\dfrac{\partial g}{\partial x^2}(x^1, x^2)\right)^{\mathrm{T}} \\[2ex]
\left(\dfrac{\partial g}{\partial x^1}(x^1, x^2)\right)^{\mathrm{T}} \mu^1 & \left(\dfrac{\partial g}{\partial x^2}(x^1, x^2)\right)^{\mathrm{T}} \mu^1 & g(x^1, x^2) & 0 \\[2ex]
\left(\dfrac{\partial g}{\partial x^1}(x^1, x^2)\right)^{\mathrm{T}} \mu^1 & \left(\dfrac{\partial g}{\partial x^2}(x^1, x^2)\right)^{\mathrm{T}} \mu^1 & 0 & g(x^1, x^2)
\end{pmatrix},
$$

$$(3.1)$$

where all the vectors in the last two rows of this $(n_1 + n_2 + 1 + 1) \times (n_1 + n_2 + m + m)$ matrix are considered as row vectors. If any of these last two rows at the solution belongs to the subspace spanned by its first $n_1 + n_2$ rows, then there is no reason to expect that this matrix has constant rank near the solution. Suppose neither of the last two rows of (3.1) at

the solution belongs to the subspace spanned by the first $n_1 + n_2$ rows, and assume that, e.g., the next-to-the-last row

$$\left(\left(\frac{\partial g_{A_+^1}}{\partial x^1}(\bar{x}^1, \bar{x}^2) \right)^{\mathrm{T}} \bar{\mu}_{A_+^1}^1 \quad \left(\frac{\partial g_{A_+^1}}{\partial x^2}(\bar{x}^1, \bar{x}^2) \right)^{\mathrm{T}} \bar{\mu}_{A_+^1}^1 \quad g(\bar{x}^1, \bar{x}^2) \quad 0 \right) \tag{3.2}$$

in (3.1) computed at the solution enters some linearly independent system of rows spanning the subspace spanned by all rows, and that either $A_+^1 \neq \emptyset$ (the case when $\bar{\mu}^1 = \bar{\mu}^2 = 0$ is clearly very special), or $A \neq \emptyset$ and $N \neq \emptyset$. The gradients of active at the solution inequalities

$$g_{A_+^1}(x^1, x^2) \leq 0, \quad -\mu_N^1 \leq 0$$

in (1.2) form the rows of the matrix

$$\left(\begin{array}{cccc} \dfrac{\partial g_{A_+^1}}{\partial x^1}(x^1, x^2) & \dfrac{\partial g_{A_+^1}}{\partial x^2}(x^1, x^2) & 0 & 0 \\ 0 & 0 & -I_N & 0 \end{array} \right), \tag{3.3}$$

where, and from now on, I denotes the identity matrix of an appropriate size. Recalling that $\bar{\mu}_{A_+^1}^1 > 0$, $g_A(\bar{x}^1, \bar{x}^2) = 0$, $g_N(\bar{x}^1, \bar{x}^2) < 0$, we conclude that the row in (3.2) can be expressed as a linear combination of the rows in (3.3) computed at the solution, and with nonnegative coefficients not all equal to zero. At the same time, if g_A is not identically equal to zero near $(\bar{x}^1, \bar{x}^2)$ (which again would be a very special case), there is no reason why the next-to-the-last row in (3.1) might belong to the subspace spanned by the rows in (3.3), except "by accident". Hence, RCPLD should not be expected hold for GNEP KKT systems, in general. The model Example 3.1 below illustrates this fact on a specific problem.

Therefore, since we cannot invoke RCPLD, specific error bound analysis tailored to GNEP is required, which is our next subject. Recall that the *strict complementarity condition* for GNEP KKT systems (1.2) means that $A_0^1 = \emptyset$ and $A_0^2 = \emptyset$; see [12, Assumption 1]. In Theorem 3.1 below, we *do not* make this assumption. Our result extends [12, Theorem 9], where strict complementarity is imposed. Also, in Theorem 3.1 we assume, for simplicity, linear independence of certain elements. However, according to the discussion following Lemma 2.1, it can be easily seen that weaker or different types of assumptions are possible. In particular, making a suitable change in the assumptions, we can easily recover [12, Theorem 8] which deals with quadratic objective functions of the players and linear constraints; see Remark 3.1 below. Moreover, there are still other possible assumptions; see again Remark 3.1.

Define the mapping $F : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$,

$$F(u) = \left(\begin{array}{c} \dfrac{\partial L_1}{\partial x^1}(x^1, x^2, \mu^1) \\ \dfrac{\partial L_2}{\partial x^2}(x^1, x^2, \mu^2) \end{array} \right), \tag{3.4}$$

where $u = (x^1, x^2, \mu^1, \mu^2)$. Then the natural residual of GNEP KKT systems (1.2) is given by $\Phi : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m \to (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}) \times \mathbb{R}^m \times \mathbb{R}^m$,

$$\Phi(u) = \left(\begin{array}{c} F(u) \\ \min\{\mu^1, -g(x^1, x^2)\} \\ \min\{\mu^2, -g(x^1, x^2)\} \end{array} \right). \tag{3.5}$$

7

Theorem 3.1 *Let f_1, f_2 and g be twice differentiable near $(\bar{x}^1, \bar{x}^2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, with their second derivatives being continuous at $(\bar{x}^1, \bar{x}^2)$. Let $\bar{u} = (\bar{x}^1, \bar{x}^2, \bar{\mu}^1, \bar{\mu}^2)$ be a solution of system (1.2), and assume that the matrix*

$$
\begin{pmatrix}
\dfrac{\partial^2 L_1}{\partial x^1 \partial x^1}(\bar{x}^1, \bar{x}^2, \bar{\mu}^1) & \dfrac{\partial^2 L_1}{\partial x^1 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^1) & \left(\dfrac{\partial g_{A_+^1}}{\partial x^1}(\bar{x}^1, \bar{x}^2)\right)^{\mathrm{T}} & 0 \\[3ex]
\dfrac{\partial^2 L_2}{\partial x^1 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^2) & \dfrac{\partial^2 L_2}{\partial x^2 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^2) & 0 & \left(\dfrac{\partial g_{A_+^2}}{\partial x^2}(\bar{x}^1, \bar{x}^2)\right)^{\mathrm{T}} \\[3ex]
\dfrac{\partial g_A}{\partial x^1}(\bar{x}^1, \bar{x}^2) & \dfrac{\partial g_A}{\partial x^2}(\bar{x}^1, \bar{x}^2) & 0 & 0
\end{pmatrix}
\tag{3.6}
$$

has the full row rank.

Then for the solution set $\bar{U}$ of system (1.2), the error bound

$$
\mathrm{dist}(u, \bar{U}) = O(\|\Phi(u)\|) \tag{3.7}
$$

holds as $u = (x^1, x^2, \mu^1, \mu^2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m$ tends to $\bar{u}$, where Φ is given by (3.5).

Proof. Define the mappings $G : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m$ and $H : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m$,

$$
G(u) = \begin{pmatrix} \mu^1 \\ \mu^2 \end{pmatrix}, \quad H(u) = \begin{pmatrix} -g(x^1, x^2) \\ -g(x^1, x^2) \end{pmatrix}.
$$

Then (1.2) is equivalent to the complementarity constraint system

$$
F(u) = 0, \quad G(u) \geq 0, \quad H(u) \geq 0, \quad \langle G(u), H(u) \rangle = 0,
$$

with F given by (3.4).

Defining the sets I_G, I_H and I_0 according to (2.2), (2.3), we obtain that each partition (I_1, I_2) of I_0 corresponds to the pair of partitions (I_1^1, I_2^1) of A_0^1, and (I_1^2, I_2^2) of A_0^2, and the corresponding branch of the form (2.9) is given by

$$
\begin{aligned}
&\frac{\partial L_1}{\partial x^1}(x^1, x^2, \mu^1) = 0, \quad \frac{\partial L_2}{\partial x^2}(x^1, x^2, \mu^2) = 0, \quad \mu_N^1 = 0, \quad \mu_N^2 = 0, \\
&\mu_{I_1^1}^1 = 0, \quad \mu_{I_1^2}^2 = 0, \quad g_{I_1^1}(x^1, x^2) \leq 0, \quad g_{I_1^2}(x^1, x^2) \leq 0, \\
&\mu_{I_2^1}^1 \geq 0, \quad \mu_{I_2^2}^2 \geq 0, \quad g_{I_2^1}(x^1, x^2) = 0, \quad g_{I_2^2}(x^1, x^2) = 0, \\
&\qquad\qquad\qquad g_{A_+}(x^1, x^2) = 0.
\end{aligned}
\tag{3.8}
$$

Simplifying these relations by removing the duplicated ones, as well as those inequalities that are implied by the corresponding equalities, we obtain the system

$$
\begin{aligned}
&\frac{\partial L_1}{\partial x^1}(x^1, x^2, \mu^1) = 0, \quad \frac{\partial L_2}{\partial x^2}(x^1, x^2, \mu^2) = 0, \quad \mu_N^1 = 0, \quad \mu_N^2 = 0, \\
&\mu_{I_1^1}^1 = 0, \quad \mu_{I_1^2}^2 = 0, \quad g_{A_0 \setminus (I_2^1 \cup I_2^2)}(x^1, x^2) \leq 0, \\
&\mu_{I_2^1}^1 \geq 0, \quad \mu_{I_2^2}^2 \geq 0, \quad g_{(I_2^1 \cup I_1^2) \cap A_0}(x^1, x^2) = 0, \\
&\qquad\qquad\qquad g_{A_+}(x^1, x^2) = 0.
\end{aligned}
\tag{3.9}
$$

8

If the matrix in (3.6) has the full row rank, then LICQ holds for this constraint system at $(\bar{x}^1, \bar{x}^2, \bar{\mu}^1, \bar{\mu}^2)$, implying the error bound for this system. By the construction of (3.9), this error bound evidently implies the error bound for the constraint system (3.8) (since the latter system simply involves some extra constraints giving rise to extra terms in the right-hand side of the error bound, while the set characterized by the two systems is the same). This error bound for (3.8) is precisely (2.10) which, in turn, implies (2.4). Therefore, by Lemma 2.1, the error bound (2.5) is valid. According to the definition of G and H, the latter is exactly (3.7) with Φ defined in (3.5). ∎

Remark 3.1 Observing that the full row rank of the matrix in (3.6) is just LICQ for all the branches (i.e., for all systems of the form (3.9)), the discussion following Lemma 2.1 puts in evidence that the nondegeneracy assumption of Theorem 3.1 can be weakened. Specifically, it can be replaced by the weaker piecewise MFCQ, piecewise (R)CRCQ, or even piecewise (R)CPLD. This immediately gives further generalizations of [12, Theorem 9] (in addition to removing strict complementarity, already mentioned).

In particular, if f_1 and f_2 are quadratic functions, and g is an affine mapping, then the constraints in (3.9) (in fact, already in (3.8)) are linear, and hence, automatically satisfy CRCQ and thus the error bound holds. This modification of Theorem 3.1 recovers [12, Theorem 8].

It is also instructive to consider the special case of $A_0 = \emptyset$. Note that this is still weaker than strict complementarity, as $A_0 = A_0^1 \cap A_0^2 = \emptyset$ is possible even if $A_0^1 \neq \emptyset$ and $A_0^2 \neq \emptyset$. For this case, we have the following.

Proposition 3.1 *Let f_1, f_2 and g be twice differentiable near $(\bar{x}^1, \bar{x}^2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, with their second derivatives being continuous at $(\bar{x}^1, \bar{x}^2)$. Let $\bar{u} = (\bar{x}^1, \bar{x}^2, \bar{\mu}^1, \bar{\mu}^2)$ be a solution of system (1.2) such that $A_0 = \emptyset$, the matrix*

$$
\begin{pmatrix}
\dfrac{\partial^2 L_1}{\partial x^1 \partial x^1}(\bar{x}^1, \bar{x}^2, \bar{\mu}^1) & \dfrac{\partial^2 L_1}{\partial x^1 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^1) & \left(\dfrac{\partial g_A}{\partial x^1}(\bar{x}^1, \bar{x}^2)\right)^{\mathrm{T}} & 0 \\[3mm]
\dfrac{\partial^2 L_2}{\partial x^1 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^2) & \dfrac{\partial^2 L_2}{\partial x^2 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^2) & 0 & \left(\dfrac{\partial g_A}{\partial x^2}(\bar{x}^1, \bar{x}^2)\right)^{\mathrm{T}} \\[3mm]
\dfrac{\partial g_A}{\partial x^1}(\bar{x}^1, \bar{x}^2) & \dfrac{\partial g_A}{\partial x^2}(\bar{x}^1, \bar{x}^2) & 0 & 0
\end{pmatrix}
\tag{3.10}
$$

has the full row rank, and there exists $(\bar{\xi}^1, \bar{\xi}^2, \bar{\zeta}^1, \bar{\zeta}^2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m$ such that

$$
\frac{\partial^2 L_1}{\partial x^1 \partial x^1}(\bar{x}^1, \bar{x}^2, \bar{\mu}^1)\bar{\xi}^1 + \frac{\partial^2 L_1}{\partial x^1 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^1)\bar{\xi}^2 + \left(\frac{\partial g}{\partial x^1}(\bar{x}^1, \bar{x}^2)\right)^{\mathrm{T}}\bar{\zeta}^1 = 0,
$$

$$
\frac{\partial^2 L_2}{\partial x^1 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^2)\bar{\xi}^1 + \frac{\partial^2 L_2}{\partial x^2 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^2)\bar{\xi}^2 + \left(\frac{\partial g}{\partial x^2}(\bar{x}^1, \bar{x}^2)\right)^{\mathrm{T}}\bar{\zeta}^1 = 0,
$$

$$
\bar{\zeta}_N^1 = 0, \quad \bar{\zeta}_N^1 = 0, \quad \frac{\partial g_A}{\partial x^1}(\bar{x}^1, \bar{x}^2)\bar{\xi}^1 + \frac{\partial g_A}{\partial x^2}(\bar{x}^1, \bar{x}^2)\bar{\xi}^2 = 0,
$$

and

$$
\bar{\zeta}_{A_0^1}^1 > 0, \quad \bar{\zeta}_{A_0^2}^2 > 0.
$$

9

Then the error bound (3.7) holds.

Proof. If $A_0 = \emptyset$ then $A_+ = A$, and systems (3.9) reduce to

$$\frac{\partial L_1}{\partial x^1}(x^1,\, x^2,\, \mu^1) = 0, \quad \frac{\partial L_2}{\partial x^2}(x^1,\, x^2,\, \mu^2) = 0, \quad \mu_N^1 = 0, \quad \mu_N^2 = 0,$$
$$\mu_{I_1^1}^1 = 0, \quad \mu_{I_1^2}^2 = 0, \quad \mu_{I_2^1}^1 \ge 0, \quad \mu_{I_2^2}^2 \ge 0, \quad g_A(x^1,\, x^2) = 0. \tag{3.11}$$

Observe that all branches given by these systems are contained in the maximal branch given by

$$\frac{\partial L_1}{\partial x^1}(x^1,\, x^2,\, \mu^1) = 0, \quad \frac{\partial L_2}{\partial x^2}(x^1,\, x^2,\, \mu^2) = 0, \quad \mu_N^1 = 0, \quad \mu_N^2 = 0,$$
$$\mu_{A_0^1}^1 \ge 0, \quad \mu_{A_0^2}^2 \ge 0, \quad g_A(x^1,\, x^2) = 0, \tag{3.12}$$

corresponding to $I_1^1 = I_1^2 = \emptyset$, $I_2^1 = I_2^2 = \{1,\, \ldots,\, m\}$.

Under the stated assumptions, MFCQ holds for (3.12). Then the error bound (2.10) holds for the specified branch $(I_1,\, I_2)$, implying the error bound (2.4) for all branches given by (3.11) (since (3.11) differs from (3.12) merely by some equalities replacing the corresponding inequalities, which can only increase the estimating term in the right-hand-side of (2.4)). Therefore, by Lemma 2.1, the error bound (3.7) is again valid in this case. ∎

Further, it can be readily seen that the error bound (3.7) implies the upper Lipschitz stability of the solutions of system (1.2) subject to canonical perturbations of the optimization problems in (1.1), and hence, subject to any smooth perturbations. Note, however, that the converse implication cannot be established as it is done for mixed complementarity problems in [18, Theorem 2] (thus covering KKT systems associated to variational problems). We conjecture that this converse implication is generally not valid for GNEP KKT systems.

Corollary 3.1 *Under the assumptions of Theorem 3.1, the estimate*

$$\mathrm{dist}(u(\sigma),\, \bar{U}) = O(\|\sigma\|)$$

holds as $\sigma = (a^1,\, a^2,\, b) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m$ tends to zero, for any solution $u(\sigma) = (x^1(\sigma),\, x^2(\sigma),\, \mu^1(\sigma),\, \mu^2(\sigma))$ of the perturbed system

$$\frac{\partial L_1}{\partial x^1}(x^1,\, x^2,\, \mu^1) = a^1, \quad \frac{\partial L_2}{\partial x^2}(x^1,\, x^2,\, \mu^2) = a^2,$$
$$\mu^1 \ge 0, \quad \langle \mu^1,\, g(x^1,\, x^2) - b \rangle = 0, \quad \mu^2 \ge 0, \quad \langle \mu^2,\, g(x^1,\, x^2) - b \rangle = 0, \quad g(x^1,\, x^2) \le b, \tag{3.13}$$

close enough to $\bar{u}$.

Proof. By Theorem 3.1, $\mathrm{dist}(u(\sigma),\, \bar{U}) = O(\|\Phi(u(\sigma))\|)$ as $u(\sigma)$ tends to $\bar{u}$, where Φ is given by (3.5).

Assuming now that $u(\sigma)$ solves (3.13) for a given σ, first note that for the part F of Φ, defined in (3.4), (3.13) implies that $F(u(\sigma)) = (a^1, a^2)$. It thus remains to show that the norm of the complementarity residuals

$$r^j(\sigma) = \min\{\mu^j(\sigma), -g(x^1(\sigma), x^2(\sigma))\}, \quad j = 1, 2,$$

is bounded above by $\|b\|$. For any $j = 1, 2$ and any $i = 1, \ldots, m$ consider the two possible cases: $\mu_i^j(\sigma) = 0$ or $\mu_i^j(\sigma) > 0$.

If $\mu_i^j(\sigma) = 0$ and $g_i(x^1(\sigma), x^2(\sigma)) \leq 0$ then $r_i^j(\sigma) = 0 \leq |b_i|$, and if $g_i(x^1(\sigma), x^2(\sigma)) > 0$ then $|r_i^j(\sigma)| = g_i(x^1(\sigma), x^2(\sigma)) \leq |b_i|$ by (3.13).

If $\mu_i^j(\sigma) > 0$ then $g_i(x^1(\sigma), x^2(\sigma)) = b_i$. Then if $r_i^j(\sigma) = -g_i(x^1(\sigma), x^2(\sigma))$, evidently $|r_i^j(\sigma)| = |b_i|$. And in the remaining case, $0 < r_i^j(\sigma) = \mu_i^j(\sigma) \leq -g_i(x^1(\sigma), x^2(\sigma)) = -b_i$, again implying the needed property. ∎

Remark 3.2 Again, according to the discussion following Lemma 2.1, the assumptions of Corollary 3.1 can be weakened: LICQ for all branch problems can be replaced by piecewise MFCQ, piecewise (R)CRCQ, or even piecewise (R)CPLD.

The assumptions of Theorem 3.1 (and thus also of Corollary 3.1) are rather natural, in the sense that they hold in various examples used in GNEP literature. The following is [14, Example 1.1], which is considered a model example containing the most important structural features of GNEPs.

Example 3.1 Consider the GNEP

$$
\begin{array}{ll}
\text{minimize}_{x^1} \quad (x^1 - 1)^2 & \quad \text{minimize}_{x^2} \quad (x^2 - 1/2)^2 \\
\text{subject to} \quad x^1 + x^2 \leq 1, & \quad \text{subject to} \quad x^1 + x^2 \leq 1.
\end{array}
$$

The primal-dual solution set of this problem has the form

$$\bar{U} = \{u = (x^1, x^2, \mu^1, \mu^2) \mid x^1 = t, \, x^2 = 1 - t, \, \mu^1 = 2(1 - t), \, \mu^2 = 2(t - 1/2), \, t \in [1/2, 1]\}.$$

Consider first the case when $t \in (1/2, 1)$. Then $A_0^1 = A_0^2 = \emptyset$, and the matrix in (3.6) has the form

$$
\begin{pmatrix}
2 & 0 & 1 & 0 \\
0 & 2 & 0 & 1 \\
1 & 1 & 0 & 0
\end{pmatrix}.
$$

This matrix has full row rank, and hence, Theorem 3.1 and Corollary 3.1 apply.

Now consider the case when $t = 1$, that is, consider the equilibrium $(\bar{x}^1, \bar{x}^2) = (1, 0)$ with the associated multipliers $(\bar{\mu}^1, \bar{\mu}^2) = (0, 1)$. Then $A_0^2 = \emptyset$, and the matrix in (3.6) has the form

$$
\begin{pmatrix}
2 & 0 & 0 \\
0 & 2 & 1 \\
1 & 1 & 0
\end{pmatrix}.
$$

This square matrix is nondegenerate, and hence, Theorem 3.1 and Corollary 3.1 apply again.

The case when $t = 1/2$ can be considered similarly to $t = 1$, and with the same conclusion.

Observe that RCPLD does not hold in this example: the Jacobian of equality constraints in (1.2) does not have constant rank near any solution of this system.

4 Newton-type methods

In this section, we first comment briefly on the consequences of our error bound results for the convergence of the smooth Levenberg–Marquardt and Gauss-Newton methods when applied to the GNEP KKT systems (possibly modified). This part requires the strict complementarity condition (to ensure local smoothness of the mapping). Our main contribution is for the general case without strict complementarity, where we show that the recently proposed in [9] (see also [10]) algorithms for constrained nonsmooth equations with nonisolated solutions can be adapted to GNEP and have local quadratic convergence rate under our assumptions (of Theorem 3.1).

First note that assuming the strict complementarity condition $A_0^1 = \emptyset$ and $A_0^2 = \emptyset$, the mapping Φ defined in (3.5) is smooth near the solution $\bar{u} = (\bar{x}^1, \bar{x}^2, \bar{\mu}^1, \bar{\mu}^2)$. Nevertheless, even in this case, the standard Newton method is not applicable to the system

$$\Phi(u) = 0,$$

because of its inherent degeneracy (not only at the solution in question, but in fact even in its neighbourhood). Indeed, for any $i \in A = A_+^1 = A_+^2$ it holds that

$$\min\{\mu_i^1, -g_i(x^1, x^2)\} = \min\{\mu_i^2, -g_i(x^1, x^2)\} = -g_i(x^1, x^2)$$

for all $u = (x^1, x^2, \mu^1, \mu^2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m$ close enough to $\bar{u}$. Therefore, near $\bar{u}$, the corresponding two components of Φ are identical, and hence, the Jacobian of Φ is degenerate everywhere around $\bar{u}$.

However, the local convergence results in [37, 18, 16] for variants of the Levenberg–Marquardt method can be applicable; see [12, Theorem 5]. Starting at a point u^k, the method computes the next iterate u^{k+1} as the solution of the equation

$$(\Phi'(u^k))^{\mathrm{T}}\Phi(u^k) + ((\Phi'(u^k))^{\mathrm{T}}\Phi'(u^k) + \alpha(u^k)I)(u - u^k) = 0,$$

where $\alpha(u^k) > 0$ is the regularization parameter. Various variants of the method differ essentially in the choices of regularization parameters. As pointed out in [12, Theorem 5], the assumptions required for convergence of such methods are the strict complementarity and the error bound. Hence, combining the corresponding results in [37, 18, 16] with Theorem 3.1, we immediately conclude that these methods possess local quadratic convergence to some point in $\bar{U}$ if the strict complementarity condition holds, and the matrix in (3.10) has the full row rank. Observe that in Example 3.1, these assumptions are satisfied at any solution corresponding to $t \in (1/2, 1)$. Moreover, as highlighted in Remark 3.1, the full row rank of (3.6) can in fact be replaced by much weaker conditions that still guarantee the error bound; for example, piecewise RCPLD.

Furthermore, under the assumptions of strict complementarity and the full row rank of (3.6), the following and different Newtonian approach is possible. To the best of our knowledge, it had not been considered previously. For any starting point $u^0 = ((x^1)^0, (x^2)^0, (\mu^1)^0, (\mu^2)^0)$ close enough to $\bar{u}$, and for any $i = 1, \ldots, m$ satisfying

$$(\mu_i^1)^0 \geq -g_i((x^1)^0, (x^2)^0), \quad (\mu_i^2)^0 \geq -g_i((x^1)^0, (x^2)^0), \tag{4.1}$$

we remove from Φ any one of the corresponding components to construct the reduced mapping $\tilde{\Phi} : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}) \times \mathbb{R}^{|A|} \times \mathbb{R}^{|N|} \times \mathbb{R}^{|N|}$,

$$\tilde{\Phi}(u) = \begin{pmatrix} F(u) \\ -g_A(x^1, x^2) \\ \mu_N^1 \\ \mu_N^2 \end{pmatrix}$$

(observe that under the strict complementarity condition, the set of i satisfying (4.1) locally coincides with A). We thus obtain the *underdetermined* system

$$\tilde{\Phi}(u) = 0,$$

to which we apply the Gauss-Newton method: having a point u^k, the next iterate u^{k+1} is the solution of the equation

$$(\tilde{\Phi}'(u^k))^{\mathrm{T}} \tilde{\Phi}(u^k) + (\tilde{\Phi}'(u^k))^{\mathrm{T}} \tilde{\Phi}'(u^k)(u - u^k) = 0.$$

If the matrix in (3.10) has the full row rank, then $\tilde{\Phi}'(\bar{u})$ has the full row rank as well. Therefore, if u^0 is close enough to $\bar{u}$ then the Gauss-Newton method thus constructed is well-defined and converges superlinearly to some point in $\bar{U}$ [25, Theorem 2.4.2].

We now turn our attention to the case when strict complementarity can be violated, like for those solutions in Example 3.1 that correspond to $t = 1/2$ and to $t = 1$. Recall that in this case Φ is not differentiable at the solution, and the solution cannot be expected to be isolated. An attractive approach to deal with this (difficult) situation is the recently proposed linear-programming-Newton method of [9]; see also [10] for related constrained nonsmooth Levenberg–Marquardt method, as well as inexact versions. Consider the constrained equation

$$\Phi(u) = 0, \quad u \in U, \tag{4.2}$$

where Φ is a given mapping (not related to GNEP, for now) and U is some set (typically, closed and convex, or even polyhedral). The method of [9] for solving (4.2) is as follows. For the current iterate $u^k \in U$, select $J(u^k)$ as an element in the suitable generalized derivative of Φ at u^k (the Jacobian $\Phi'(u^k)$, if Φ is differentiable at u^k). Compute the next iterate u^{k+1} as (part of) the solution (u^{k+1}, t_{k+1}) with some $t_{k+1} \in \mathbb{R}$ of the subproblem

$$\begin{aligned} \text{minimize}_{(t, u)} \quad & t \\ \text{subject to} \quad & \|\Phi(u^k) + J(u^k)(u - u^k)\| \leq t\|\Phi(u^k)\|^2, \\ & \|u - u^k\| \leq t\|\Phi(u^k)\|, \\ & u \in U. \end{aligned} \tag{4.3}$$

Note that if U is polyhedral and the ∞-norm is employed, then this is a linear programming (LP) problem. It should be noted that writing the optimality conditions for (4.3), it can be observed that it is equivalent to an iteration of the constrained Levenberg–Marquardt method for (4.2), with a special choice of the regularization parameter.

Local quadratic convergence of the method described by (4.3) is established in [9] under four assumptions, which therefore have to be verified in our context. (We note that the assumptions for the extensions in [10] are the same.) This however is not all that straightfoward;

moreover, simply applying the algorithm to the mapping Φ given by (3.5) is not practical unless the constraints in (1.1) are affine. We proceed to discuss those issues in detail.

Assumption 4 in [9] formalizes the needed quality of approximation for Φ around u^k provided by the linearization using the substitute $J(u^k)$ of the derivative at u^k. There is no need to state it here formally; we shall just give the reasons why it holds in our case, referring to the coresponding results in [9]. Assuming that f_1, f_2 and g are twice differentiable near $(\bar{x}^1,\, \bar{x}^2)$ with their second derivatives being locally Lipschitz-continuous at this point, the mapping Φ in (3.5) is piecewise smooth near $\bar{u}$ and appropriate substitutes for the derivative of this mapping are the Jacobians of active smooth pieces. According to [9, Theorem 4, Corollary 5], with this choice, Assumption 4 holds for the constrained equation (4.2) if the set U is defined either by

$$U = \{u = (x^1,\, x^2,\, \mu^1,\, \mu^2) \mid \mu^1 \geq 0,\, \mu^2 \geq 0,\, g(x^1,\, x^2) \leq 0\}$$

or by

$$U = \{u = (x^1,\, x^2,\, \mu^1,\, \mu^2) \mid \mu^1 - g(x^1,\, x^2) \geq 0,\, \mu^2 - g(x^1,\, x^2) \geq 0\}.$$

At this point, one might wonder why not solve (4.2) with U being the whole space, without constraints, since the solution set of the unconstrained equation already gives all GNEP KKT solutions. The reason is that the required Assumption 4 need not hold if the complementarity relations are reformulated via the unconstrained min-operation as in the definition of Φ, without the constraints as above added; see [9, Section 3.3 and Example 3].

Next note that unless g is affine, (4.3) with U defined in either of the ways stated above is a nonlinearly-contrained problem, in general too expensive to solve at each iteration. Thus, instead of solving the constrained equation (4.2), we propose the following construction (in fact analogous to the one for variational KKT systems in [9]). We introduce an auxiliary variable $v \in \mathbb{R}^m$ and define the mapping $\Psi : (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m) \times \mathbb{R}^m \to (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}) \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m$,

$$\Psi(u,\, v) = \begin{pmatrix} F(u) \\ g(x^1,\, x^2) + v \\ \min\{\mu^1,\, v\} \\ \min\{\mu^2,\, v\} \end{pmatrix}, \tag{4.4}$$

where F is given by (3.4), and consider the constrained equation

$$\Psi(u,\, v) = 0, \quad (u,\, v) \in \Omega, \tag{4.5}$$

where the set Ω is given either by

$$\Omega = \{(u,\, v) = ((x^1,\, x^2,\, \mu^1,\, \mu^2),\, v) \mid \mu^1 \geq 0,\, \mu^2 \geq 0,\, v \geq 0\}$$

or by

$$\Omega = \{(u,\, v) = ((x^1,\, x^2,\, \mu^1,\, \mu^2),\, v) \mid \mu^1 - v \geq 0,\, \mu^2 - v \geq 0\}.$$

It is evident that $(\bar{u},\, -g(\bar{x}^1,\, \bar{x}^2))$ is a solution of (4.5).

To solve the GNEP KKT systems (1.2), the algorithm is now as follows. For the current iterate $(u^k,\, v^k) \in \Omega$, select $J(u^k,\, v^k)$ as the Jacobian of any of the smooth pieces of Ψ active

at (u^k, v^k). Compute the next iterate (u^{k+1}, v^{k+1}) as (part of) the solution $(u^{k+1}, v^{k+1}, t_{k+1})$ with some $t_{k+1} \in \mathbb{R}$ of the subproblem

$$
\begin{array}{ll}
\text{minimize}_{(t, u, v)} & t \\
\text{subject to} & \|\Psi(u^k, v^k) + J(u^k, v^k)(u - u^k, v - v^k)\| \leq t\|\Psi(u^k, v^k)\|^2, \\
& \|(u - u^k, v - v^k)\| \leq t\|\Psi(u^k, v^k)\|, \\
& (u, v) \in \Omega.
\end{array}
\tag{4.6}
$$

Note that using the ∞-norm, this is *always* an LP, regardless of the problem data in (1.1). We now have to come back to the four assumptions of [9].

By the same reasoning as above for Φ, Assumption 4 in [9] holds around the solution $(\bar{u}, \bar{v})$, with $\bar{v} = -g(\bar{x}^1, \bar{x}^2)$, for Ψ and the specified choice of J.

Assumption 1 in [9] states that a constant times the distance to the solution set bounds *from above* (unlike from below, like in the error bound property) the residual of the equation mapping. This holds automatically here, since Ψ is locally Lipschitz-continuous, which in turn is implied by the piecewise smoothness of Ψ.

Assumption 2 in [9] is the error bound property for (4.5), and it follows from our Theorem 3.1, as we demonstrate next.

Corollary 4.1 *Under the assumptions of Theorem 3.1, for the solution set $\bar{\Omega}$ of the equation*

$$\Psi(u, v) = 0$$

with $\Psi : (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m) \times \mathbb{R}^m \to (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}) \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m$ defined in (4.4), the error bound

$$\text{dist}((u, v), \bar{\Omega}) = O(\|\Psi(u, v)\|)$$

holds as $u = (x^1, x^2, \mu^1, \mu^2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m$ tends to $\bar{u}$, and as $v \in \mathbb{R}^m$ tends to $-g(\bar{x}^1, \bar{x}^2)$.

Proof. Employing the Mean-Value Theorem, we obtain the existence of $\ell > 0$ such that

$$
\begin{aligned}
\text{dist}((u, v), \bar{\Omega}) &= \inf_{\tilde{u}=(\tilde{x}^1, \tilde{x}^2, \tilde{\mu}^1, \tilde{\mu}^2)\in \bar{U}} \|(u, v) - (\tilde{u}, -g(\tilde{x}^1, \tilde{x}^2))\| \\
&\leq \inf_{\tilde{u}=(\tilde{x}^1, \tilde{x}^2, \tilde{\mu}^1, \tilde{\mu}^2)\in \bar{U}} (\|u - \tilde{u}\| + \|g(x^1, x^2) - g(\tilde{x}^1, \tilde{x}^2))\|) + \|g(x^1, x^2) + v\| \\
&\leq \inf_{\tilde{u}=(\tilde{x}^1, \tilde{x}^2, \tilde{\mu}^1, \tilde{\mu}^2)\in \bar{U}} (\|u - \tilde{u}\| + \ell\|(x^1, x^2) - (\tilde{x}^1, \tilde{x}^2)\|) + \|g(x^1, x^2) + v\| \\
&\leq (1 + \ell) \inf_{\tilde{u}\in \bar{U}} \|u - \tilde{u}\| + \|g(x^1, x^2) + v\| \\
&= (1 + \ell) \text{dist}(u, \bar{U}) + \|g(x^1, x^2) + v\|
\end{aligned}
$$

for any u and v. Observe that according to (3.5) and (4.4)

$$\|\Phi(u)\| \leq \|\Psi(u, v)\| + 2\|g(x^1, x^2) + v\|_1,$$

15

and therefore, employing (3.7), it holds that

$$\text{dist}((u,\, v),\, \bar{\Omega}) = O(\|\Phi(u)\|) + \|g(x^1,\, x^2) + v\| = O(\|\Psi(u,\, v)\|)$$

as $u \to \bar{u}$ and $v \to -g(\bar{x}^1,\, \bar{x}^2)$. ∎

It remains to verify Assumption 3 in [9]. This assumption is rather technical, and it again does not seem reasonable to state here in its original form. Instead, we shall mention that [9, Corollary 3] shows that it is satisfied provided the error bound holds for the set defined by the system of equations corresponding to each smooth piece of Ψ active at the solution. It is the latter property that we establish next.

Proposition 4.1 *Under the assumptions of Theorem 3.1, for any partitions $I^1 = (I_1^1,\, I_2^1)$ of A_0^1, and $I^2 = (I_1^2,\, I_2^2)$ of A_0^2, for the solution set $\bar{\Omega}_{I^1,\, I^2}$ of the system of equations*

$$\Psi_{I^1,\, I^2}(u,\, v) = 0 \tag{4.7}$$

with $\Psi_{I^1,\, I^2} : (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m) \times \mathbb{R}^m \to (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}) \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m$,

$$\Psi_{I^1,\, I^2}(u,\, v) = \begin{pmatrix} F(u) \\ g(x^1,\, x^2) + v \\ \left(\mu_{I_1^1 \cup N}^1,\, v_{A_+^1 \cup I_2^1} \right) \\ \left(\mu_{I_1^2 \cup N}^2,\, v_{A_+^2 \cup I_2^2} \right) \end{pmatrix},$$

the error bound

$$\text{dist}\left((u,\, v),\, \bar{\Omega}_{I^1,\, I^2}\right) = O\left(\|\Psi_{I^1,\, I^2}(u,\, v)\|\right) \tag{4.8}$$

holds as $u = (x^1,\, x^2,\, \mu^1,\, \mu^2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m$ tends to $\bar{u}$, and as $v \in \mathbb{R}^m$ tends to $-g(\bar{x}^1,\, \bar{x}^2)$.

Proof. System (4.7) may contain duplicated equations of the form $v_i = 0$, $i \in A_+^1 \cap A_+^2$ or $i \in I_2^1 \cap I_2^2$. Removing the duplicated copies of these equations, we transform (4.7) into the equivalent system

$$\tilde{\Psi}_{I^1,\, I^2}(u,\, v) = 0$$

with $\tilde{\Psi}_{I^1,\, I^2} : (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \times \mathbb{R}^m) \times \mathbb{R}^m \to (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}) \times \mathbb{R}^{|I_1^1 \cup N|} \times \mathbb{R}^{|I_1^2 \cup N|} \times \mathbb{R}^{|A|}$,

$$\tilde{\Psi}_{I^1,\, I^2}(u,\, v) = \begin{pmatrix} F(u) \\ g(x^1,\, x^2) + v \\ \mu_{I_1^1 \cup N}^1 \\ \mu_{I_1^2 \cup N}^2 \\ v_{A_+ \cup ((I_2^1 \cup I_2^2) \cap A_0)} \end{pmatrix}.$$

The error bound

$$\text{dist}\left((u,\, v),\, \bar{\Omega}_{I^1,\, I^2}\right) = O\left(\|\tilde{\Psi}_{I^1,\, I^2}(u,\, v)\|\right)$$

16

(evidently implying the needed error bound (4.8)) holds if the Jacobian $\tilde{\Psi}'_{I^1, I^2}(\bar{u}, -g(\bar{x}^1, \bar{x}^2))$ has full row rank. It can be directly verified that the latter holds if the matrix

$$
\begin{pmatrix}
\dfrac{\partial^2 L_1}{\partial x^1 \partial x^1}(\bar{x}^1, \bar{x}^2, \bar{\mu}^1) & \dfrac{\partial^2 L_1}{\partial x^1 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^1) & \left(\dfrac{\partial g_{A^1_+ \cup I^1_2}}{\partial x^1}(\bar{x}^1, \bar{x}^2)\right)^{\mathrm{T}} & 0 \\[2ex]
\dfrac{\partial^2 L_2}{\partial x^1 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^2) & \dfrac{\partial^2 L_2}{\partial x^2 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^2) & 0 & \left(\dfrac{\partial g_{A^2_+ \cup I^2_2}}{\partial x^2}(\bar{x}^1, \bar{x}^2)\right)^{\mathrm{T}} \\[2ex]
\dfrac{\partial g_{A_+}}{\partial x^1}(\bar{x}^1, \bar{x}^2) & \dfrac{\partial g_{A_+}}{\partial x^2}(\bar{x}^1, \bar{x}^2) & 0 & 0 \\[1ex]
\dfrac{\partial g_{I^1_2 \cap A_0}}{\partial x^1}(\bar{x}^1, \bar{x}^2) & \dfrac{\partial g_{I^1_2 \cap A_0}}{\partial x^2}(\bar{x}^1, \bar{x}^2) & 0 & 0 \\[2ex]
\dfrac{\partial g_{I^2_2 \cap A_0}}{\partial x^1}(\bar{x}^1, \bar{x}^2) & \dfrac{\partial g_{I^2_2 \cap A_0}}{\partial x^2}(\bar{x}^1, \bar{x}^2) & 0 & 0
\end{pmatrix}
$$

has full row rank. It remains to observe that the latter is automatic if the matrix in (3.6) has full row rank. ∎

Combining the properties verified above with [9, Theorem 1], we obtain local quadratic convergence of the LP-Newton method described by (4.6) under the assumptions of our Theorem 3.1.

Theorem 4.1 *In addition to the assumptions of Theorem 3.1, assume that the second derivatives of f_1, f_2 and g are locally Lipschitz-continuous at $(\bar{x}^1, \bar{x}^2)$.*

Then for any starting point $(u^0, v^0) \in \Omega$ close enough to $(\bar{u}, -g(\bar{x}^1, \bar{x}^2))$ the iterative scheme (4.6) defines a sequence $\{(u^k, v^k)\}$ which converges quadratically to $(u_*, -g(x^1_*, x^2_*))$ with some $u_* \in \bar{U}$, where $\bar{U}$ is the solution set of system (1.2).

Remark 4.1 Since the extension in [10] requires the same four assumptions as [9], our conclusions apply to those methods as well. In particular, to the inexact nonsmooth constrained Levenberg–Marquardt method.

Remark 4.2 Subsequently to the original version of this paper, there appeared an independent technical report [7] concerned with similar issues, namely, primal-dual error bounds and Newton-type methods for GNEP.

In our setting (joint constraints only) Assumption 1 in [7] required to obtain an error bound consists of saying that there exists a partition (I_1, I_2) of A (i.e., $I_1 \cup I_2 = A$, $I_1 \cap I_2 = \emptyset$) such that $I_1 \subset A^1_+$, $I_2 \subset A^2_+$, and the square matrix

$$
\begin{pmatrix}
\dfrac{\partial^2 L_1}{\partial x^1 \partial x^1}(\bar{x}^1, \bar{x}^2, \bar{\mu}^1) & \dfrac{\partial^2 L_1}{\partial x^1 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^1) & \left(\dfrac{\partial g_{I_1}}{\partial x^1}(\bar{x}^1, \bar{x}^2)\right)^{\mathrm{T}} & 0 \\[2ex]
\dfrac{\partial^2 L_2}{\partial x^1 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^2) & \dfrac{\partial^2 L_2}{\partial x^2 \partial x^2}(\bar{x}^1, \bar{x}^2, \bar{\mu}^2) & 0 & \left(\dfrac{\partial g_{I_2}}{\partial x^2}(\bar{x}^1, \bar{x}^2)\right)^{\mathrm{T}} \\[2ex]
\dfrac{\partial g_A}{\partial x^1}(\bar{x}^1, \bar{x}^2) & \dfrac{\partial g_A}{\partial x^2}(\bar{x}^1, \bar{x}^2) & 0 & 0
\end{pmatrix}
$$

17

is nonsingular. This evidently implies the full rank condition in our Theorem 3.1. The converse is not true, even assuming strict complementarity, as demonstrated by Example 4.1 below. Moreover, as mentioned in Remark 3.1, our line of analysis allows for further relaxations of this full rank assumption. Finally, the two error bounds are of different nature, as explained next. The error bound in [7] is valid only on the set including all nonnegativity conditions, and on this set the residual used in the error bound is smooth. Our residual is nonsmooth, and the error bound is on the entire space.

As a consequence of nonsmoothness of our residual, application of the LP-Newton method, developed in [9] specifically for nonsmooth constrained equations, is natural in our setting. In [7], since the residual there is smooth, in principle the usual Levenberg–Marquardt methods for smooth constrained equations could be used instead of the LP-Newton method. Also, the methods in this paper and in [7] are different, and each may have some practical advantages (as well as disadvantages, of course). For instance, our approach requires less auxiliary variables: in our setting the dimension of v in (4.4) is that of the number of constraints, while for the method in [7] it would be twice as much. And most importantly, as explained above, our assumptions for local quadratic convergence of Newtonian methods are weaker than those in [7].

Example 4.1 Consider the GNEP

$$
\begin{array}{ll}
\text{minimize}_{x^1} \quad -x^1 & \qquad \text{minimize}_{x^2} \quad -x^2 \\
\text{subject to} \quad x^1 + x^2 \le 0, & \qquad \text{subject to} \quad x^1 + x^2 \le 0.
\end{array}
$$

The primal-dual solution set of this problem has the form

$$
\bar{U} = \{u = (x^1,\, x^2,\, 1,\, 1) \mid x^1 + x^2 = 0\}
$$

(in particular, all the equilibria are in fact variational).

At any point of $\bar{U}$, it holds that $A_0^1 = A_0^2 = \emptyset$, and the matrix in (3.6) has the form

$$
\begin{pmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
1 & 1 & 0 & 0
\end{pmatrix}.
\tag{4.9}
$$

This matrix has full row rank. Hence, Theorem 3.1, Corollary 3.1, and Theorem 4.1 apply.

A the same time, in this case Assumption 1 in [7] consists of saying that at least one of the matrices obtained by eliminating one of the last two columns in (4.9) is nonsingular, which is not the case.

References

[1] R. Andreani, G. Haeser, M.L. Schuverdt, and P.J.S. Silva. A relaxed constant positive linear dependence constraint qualification and applications. Math. Program. 135 (2012), 255–273.

[2] S.C. Billups. A homotopy-based algorithm for mixed complementarity problems. SIAM J. Optim. 12 (2002), 583–605.

[3] J.F. Bonnans and A. Shapiro. Perturbation Analysis of Optimization Problems. Springer-Verlag, New York, 2000.

[4] A.N. Daryina, A.F. Izmailov, and M.V. Solodov. A class of active-set Newton methods for mixed complementarity problems. SIAM J. Optim. 15 (2004), 409–429.

[5] A.N. Daryina, A.F. Izmailov, and M.V. Solodov. Mixed complementary problems: regularity, estimates of the distance to the solution, and Newton's methods. Comput. Math. Math. Phys. 44 (2004), 45–61.

[6] A.N. Daryina, A.F. Izmailov, and M.V. Solodov. Numerical results for a globalized active-set Newton method for mixed complementarity problems. Comp. Appl. Math. 24 (2005), 293–316.

[7] A. Dreves, F. Facchinei, A. Fischer, and M. Herrich. A new error bound result for generalized Nash equilibrium problems and its algorithmic application. Technical Report MATH-NM-1-2013. Technische Universität Dresden, Institute für Mathematik. January 2013. Available at Optimization Online.

[8] A. Dreves, F. Facchinei, C. Kanzow, and S. Sagratella. On the solution of the KKT conditions of generalized Nash equilibrium problems. SIAM J. Optim. 21 (2011), 1082–1108.

[9] F. Facchinei, A. Fischer, and M. Herrich. An LP-Newton method: nonsmooth equations, KKT systems, and nonisolated solutions. Math. Program., DOI 10.1007/s10107-013-0676-6, 2013.

[10] F. Facchinei, A. Fischer, and M. Herrich. A family of Newton methods for nonsmooth constrained systems with nonisolated solutions. Math. Methods Oper. Res., DOI 10.1007/s00186-012-0419-0, 2012.

[11] F. Facchinei, A. Fischer, and V. Piccialli. On generalized Nash games and variational inequalities. Oper. Res. Lett. 35 (2007), 159–164.

[12] F. Facchinei, A. Fischer, and V. Piccialli. Generalized Nash equilibrium problems and Newton methods. Math. Program. 117 (2009), 163–194.

[13] F. Facchinei and C. Kanzow. Penalty methods for the solution of generalized Nash equilibrium problems. SIAM J. Optim. 20 (2010), 2228–2253.

[14] F. Facchinei and C. Kanzow. Generalized Nash equilibrium problems. Ann. Oper. Res. 175 (2010), 177–211.

[15] F. Facchinei and J.-S. Pang. Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer-Verlag, New York, 2003.

[16] J. Fan and Y. Yuan. On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption. Computing 74 (2005), 23–39.

[17] M.C. Ferris, C. Kanzow, and T.S. Munson. Feasible descent algorithms for mixed complementarity problems. Math. Program. 86 (1999), 475–497.

[18] A. Fischer. Local behavior of an iterative framework for generalized equations with nonisolated solutions. Math. Program. 94 (2002), 91–124.

[19] M. Fukushima and J.-S. Pang. Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games. Comput. Manag. Science 2 (2005), 21–56. [Erratum, ibid. 6 (2009), 373–375].

[20] P.T. Harker. Generalized Nash games and quasi-variational inequalities. Eur. J. Oper. Res. 54 (1991), 81–94.

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

[22] R. Janin. Directional derivative of the marginal function in nonlinear programming. Math. Program. Stud. 21 (1984), 110–126.

[23] C. Kanzow. Strictly feasible equation-based methods for mixed complementarity problems. Numer. Math. 89 (2001), 135–160.

[24] C. Kanzow and M. Fukushima. Solving box constrained variational inequality problems by using the natural residual with D-gap function globalization. Oper. Res. Letters 23 (1998), 45–51.

[25] C.T. Kelley. Iterative Methods for Optimization. SIAM, Philadelphia, 1999.

[26] K. Kubota and M. Fukushima. Gap function approach to the generalized Nash equilibrium problem. J. Optim. Theory Appl. 144 (2010), 511–531.

[27] A.A. Kulkarni and U.V. Shanbhag. Revisiting generalized Nash games and variational inequalities. J. Optim. Theory Appl., DOI 10.1007/s10957-011-9981-5, 2012.

[28] J.P. Luna, C. Sagastizábal, and M. Solodov. A class of Dantzig–Wolfe type decomposition methods for variational inequality problems. Math. Program., DOI 10.1007/s10107-012-0599-7, 2012.

[29] Z.-Q. Luo, J.-S. Pang, and D. Ralph. Mathematical Programs with Equilibrium Constraints. Cambridge Univ. Press, Cambridge, 1996.

[30] O.L. Mangasarian and S. Fromovitz. The Fritz John optimality conditions in the presence of equality and inequality constraints. J. Math. Anal. Appl. 17 (1967), 37–47.

[31] L. Minchenko and S. Stakhovski. On relaxed constant rank regularity condition in mathematical programming. Optimization 60 (2011), 429–440.

[32] K. Nabetani, P. Tseng, and M. Fukushima. Parametrized variational inequality approaches to generalized Nash equilibrium problems with shared constraints. Comput. Optim. Appl. 48 (2011), 423–452.

[33] J.-S. Pang. Error bounds in mathematical programming. Math. Program. 79 (1997), 299–332.

[34] L. Qi and Z. Wei. On the constant positive linear dependence condition and its application to SQP methods. SIAM J. Optim. 10 (2000), 963–981.

[35] J.B. Rosen. Existence and uniqueness of equilibrium points for concave N-person games. Econometrica 33 (1965), 52–534.

[36] M.V. Solodov. Constraint qualifications. In Wiley Encyclopedia of Operations Research and Management Science, James J. Cochran, et al. (editors), John Wiley & Sons, Inc., 2010.

[37] N. Yamashita and M. Fukushima. On the rate of convergence of the Levenberg-Marquardt method. In Topics in Numerical Analysis, pp. 239–249, Comput. Suppl., 15, Springer, Vienna, 2001.