Math. Program., Ser. B (2009) 117:163–194
DOI 10.1007/s10107-007-0160-2

FULL LENGTH PAPER

Generalized Nash equilibrium problems and Newton methods

Francisco Facchinei · Andreas Fischer · Veronica Piccialli

Received: 17 February 2006 / Accepted: 10 April 2007 / Published online: 19 July 2007

Abstract The generalized Nash equilibrium problem, where the feasible sets of the players may depend on the other players' strategies, is emerging as an important modeling tool. However, its use is limited by its great analytical complexity. We consider several Newton methods, analyze their features and compare their range of applicability. We illustrate in detail the results obtained by applying them to a model for internet switching.

Keywords Generalized Nash equilibrium · Semismooth Newton method · Levenberg–Marquardt method · Nonisolated solution · Internet switching

Mathematics Subject Classification (2000) 90C30 · 91A10 · 91A80 · 49M05

Dedicated to Stephen M. Robinson on the occasion of his 65th birthday, in honor of his fundamental contributions to mathematical programming.

The work of the first and third authors has been partially supported by MIUR-PRIN 2005 Research Program n.2005017083 "Innovative Problems and Methods in Nonlinear Optimization".

F. Facchinei (✉) · V. Piccialli
Department of Computer and System Sciences "A. Ruberti", "Sapienza" Università di Roma,
via Ariosto 25, 00185 Roma, Italy
e-mail: facchinei@dis.uniroma1.it

V. Piccialli
e-mail: piccialli@dis.uniroma1.it

A. Fischer
Institute of Numerical Mathematics, Technische Universität Dresden,
01062 Dresden, Germany
e-mail: Andreas.Fischer@tu-dresden.de

1 Introduction

In this paper we consider the Generalized Nash Equilibrium Problem (GNEP for short). The GNEP extends the classical Nash equilibrium problem by assuming that each player's feasible set can depend on the rival players' strategies. There are N players, and each player ν controls the variables $x^\nu \in \mathbb{R}^{n_\nu}$. We denote by $\boldsymbol{x}$ the vector formed by all these decision variables

$$\boldsymbol{x} := \begin{pmatrix} x^1 \\ \vdots \\ x^N \end{pmatrix},$$

which has dimension $n := \sum_{\nu=1}^N n_\nu$ and by $\boldsymbol{x}^{-\nu}$ the vector formed by all the players' decision variables except those of player ν. To emphasize the ν-th player's variables within $\boldsymbol{x}$ we sometimes write $(x^\nu, \boldsymbol{x}^{-\nu})$ instead of $\boldsymbol{x}$. As a general mnemonic rule we note that if a denotes a vector attached to a single player, we denote by $\boldsymbol{a}$ the vector comprising the a of all (or of a certain subset of) the players.

Each player's strategy must belong to a set $X_\nu(\boldsymbol{x}^{-\nu}) \subseteq \mathbb{R}^{n_\nu}$ that depends on the rival players' strategies. The aim of player ν, given the other players' strategies $\boldsymbol{x}^{-\nu}$, is to choose a strategy x^ν that solves the minimization problem

$$\text{minimize}_{x^\nu}\ \theta_\nu(x^\nu, \boldsymbol{x}^{-\nu}) \quad \text{subject to} \quad x^\nu \in X_\nu(\boldsymbol{x}^{-\nu}), \tag{1}$$

where $-\theta_\nu$ is often called payoff function of player ν. For any $\boldsymbol{x}^{-\nu}$, the solution set of problem (1) is denoted by $\mathcal{S}_\nu(\boldsymbol{x}^{-\nu})$. The GNEP is the problem of finding a vector $\bar{\boldsymbol{x}}$ such that

$$\bar{x}^\nu \in \mathcal{S}_\nu(\bar{\boldsymbol{x}}^{-\nu}) \qquad \text{for all } \nu.$$

Such a point $\bar{\boldsymbol{x}}$ is called a (generalized) Nash equilibrium or, more simply, a solution of the GNEP.

There are many interesting issues related to this kind of problem, some arising from its mathematical challenges some from its typical applications. When GNEPs are used to establish "behavioral rules" for example, modelers often want the solution to be unique, so that the study of this problem has an important role, even if uniqueness is a very strong condition. When a manifold of solutions exists, one might be interested in computing a selection of solutions that in some sense approximates the set of all solutions. In other applications it can be important to find a solution that satisfies additional, desirable properties. For example, in some cases it is sensible to look for a "normalized equilibrium" (see for example Sect. 3.2 and [18]). In other cases a more general approach can be envisaged where a Mathematical Program with Equilibrium Constraints may allow the modeler to find a generalized Nash equilibrium that minimizes a certain additional criterion.

The main aim of this paper is algorithmic. We focus on the study of several Newton methods for the computation of one generalized Nash equilibrium (of possibly infinitely many existing ones). Using the Karush–Kuhn–Tucker (KKT) systems for the player's optimization problems, one can show that a necessary condition for a point $\boldsymbol{x}$ to be a solution of the GNEP is that it satisfies, together with suitable multipliers, a structured mixed complementarity problem. To this system we apply appropriate semismooth Newton-type methods requiring, at each iteration, the solution of a linear system of equations. In spite of the many similarities to optimization/variational inequalities (VI) problems, the GNEP presents challenging peculiarities that make its analysis especially demanding. The natural extension of standard conditions and assumptions normally used in optimization/VI theory may turn out to be inappropriate in many cases and care must be exercised so that realistic assumptions are made in dealing with GNEPs. For example, in large and interesting classes of GNEP local uniqueness of the solutions is not likely to be encountered. Therefore, classical conditions and techniques for the development of Newton methods must be abandoned in favor of more sophisticated ones. We not only develop a local convergence theory for several cases of the GNEP, but also identify specific structures that are likely to occur in GNEPs and analyze their properties and peculiarities. We hope that this study of a largely uncharted territory may be useful to other researchers and will stimulate further interest in GNEPs.

Since Arrow and Debreu's 1954 classical paper [1] on the existence of equilibria for a competitive economy, GNEPs have been the subject of a constant if not intense interest. There have been further studies on existence [2,22,30], and connections to quasi-variational inequalities have been highlighted [3,18]. Furthermore the GNEP has been used to model a host of interesting problems arising in economy and, more recently, computer science, telecommunications, and deregulated markets. However, probably due to the daunting difficulty of the problem, advancements on the algorithmic side have been rather scarce, and essentially only amount to the development of the so-called relaxation algorithm (see [4,21,31]) based on the Nikaido–Isoda function [24].

With a few notable exceptions (see [18,28,29]), the interest of the mathematical programming community in generalized Nash equilibrium problems is recent, see [14,17,25,26], and stems principally from the desire to attack some very hard problems describing complex competition situations, especially in the energy markets, see for example [5,6,8,17] and references therein. In turn, the development of efficient numerical methods for this kind of problems rests on recent advancements in the study of variational inequalities, semismooth methods, and mathematical programs with equilibrium constraints.

The paper is organized as follows. In the next section we recall some preliminary and basic facts and definitions. In Sect. 3 three approaches to the development of Newton methods for the solution of the GNEP are described, i.e., we discuss the setting in which each method can be applied and introduce the required assumptions. The latter are discussed and compared in detail in Sect. 4. Finally, in Sect. 5 we illustrate the various methods and conditions on an interesting application coming from computer science.

For a continuously differentiable function $H : \mathbb{R}^s \to \mathbb{R}^s$ the Jacobian of H at $y \in \mathbb{R}^s$ is denoted by $JH(y)$ and its transposed by $\nabla H(y)$. Throughout the paper $\|\cdot\|$ denotes the Euclidean norm and $\mathbb{B}(y, \delta)$ the closed Euclidean ball with center y and radius δ. For a nonempty set $\Omega \subseteq \mathbb{R}^s$ the Euclidean distance of y to Ω is defined by $\operatorname{dist}[y, \Omega] := \inf_{z\in\Omega} \|z - y\|$. Let $M = (M_{ij})$ be an $s \times s$ matrix. Then, for index sets $I, J \subseteq \{1, \dots s\}$, $M_{I,J}$ denotes the $|I| \times |J|$ submatrix of M consisting of elements M_{ij}, $i \in I$, $j \in J$. For $w \in \mathbb{R}^s$, w_J is the subvector with components w_j, $j \in J$. $\mathbf{I}_s$

denotes the $s \times s$ identity matrix, whereas $\mathbf{0}_s$ is the $s \times s$ matrix and $\mathbf{0}_{s\times t}$ the $s \times t$ matrix with zero entries only.

2 Basic facts, definitions and assumptions

In this section we will introduce a system that is naturally associated with the GNEP and discuss the relations between these two problems. This system is then reformulated as a nonsmooth system of equations, and this will be the basis of many of the developments in the paper.

In practical applications the feasible set $X_\nu(\boldsymbol{x}^{-\nu})$ of player ν is defined by a finite number of constraints. Let $g^\nu : \mathbb{R}^n \to \mathbb{R}^{m_\nu}$ denote the constraint mapping associated with player ν (where we recall that $n = \sum_{\nu=1}^N n_\nu$), the feasible set of player ν is then given by

$$X_\nu(\boldsymbol{x}^{-\nu}) := \{x^\nu \in \mathbb{R}^{n_\nu} \; : \; g^\nu(x^\nu, \boldsymbol{x}^{-\nu}) \leq 0\}, \tag{2}$$

where $g^\nu(\boldsymbol{x}) \leq 0$ is understood componentwise. We denote by m the total number of constraints in the GNEP, i.e., $m := \sum_{\nu=1}^N m_\nu$. Throughout the paper we make the following blanket assumption:

Smoothness assumption. For each $\nu = 1, \ldots, N$ the functions $\theta_\nu : \mathbb{R}^n \to \mathbb{R}$ and $g^\nu : \mathbb{R}^n \to \mathbb{R}^{m_\nu}$ are twice differentiable with locally Lipschitz continuous second order derivatives.

Remark 1 If the local Lipschitz continuity of the second order derivatives is replaced by their simple continuity, the convergence results in the paper still hold with superlinear convergence instead of a quadratic rate.

Suppose that $\bar{\boldsymbol{x}}$ is a solution of the GNEP. Then, if for player ν a suitable constraint qualification holds (for example the Mangasarian–Fromovitz or the Slater constraint qualification), there is a vector $\bar{\lambda}^\nu \in \mathbb{R}^{m_\nu}$ of multipliers so that the classical KKT conditions

$$\nabla_{x^\nu} L_\nu(x^\nu, \bar{\boldsymbol{x}}^{-\nu}, \lambda^\nu) = 0$$
$$0 \leq \lambda^\nu \perp -g^\nu(x^\nu, \bar{\boldsymbol{x}}^{-\nu}) \geq 0$$

are satisfied by $(\bar{x}^\nu, \bar{\lambda}^\nu)$, where $L_\nu(\boldsymbol{x}, \lambda^\nu) := \theta_\nu(\boldsymbol{x}) + g^\nu(\boldsymbol{x})^\top \lambda^\nu$ is the Lagrangian associated with the ν-th player's optimization problem. Concatenating these N KKT systems, we obtain that if $\bar{\boldsymbol{x}}$ is a solution of the GNEP and if a suitable constraint qualification holds for all players, then a multiplier $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^m$ exists that together with $\bar{\boldsymbol{x}}$ satisfies the system

$$\begin{aligned} \mathbf{L}(\boldsymbol{x}, \boldsymbol{\lambda}) &= 0 \\ 0 \leq \boldsymbol{\lambda} \perp -\mathbf{g}(\boldsymbol{x}) &\geq 0, \end{aligned} \tag{3}$$

where

$$\boldsymbol{\lambda} := \begin{pmatrix} \lambda^1 \\ \vdots \\ \lambda^N \end{pmatrix}, \quad \mathbf{g}(\boldsymbol{x}) := \begin{pmatrix} g^1(\boldsymbol{x}) \\ \vdots \\ g^N(\boldsymbol{x}) \end{pmatrix}, \text{ and } \mathbf{L}(\boldsymbol{x}, \boldsymbol{\lambda}) := \begin{pmatrix} \nabla_{x^1} L_1(\boldsymbol{x}, \lambda^1) \\ \vdots \\ \nabla_{x^N} L_N(\boldsymbol{x}, \lambda^N) \end{pmatrix}.$$

For simplicity no distinction will be made between

$$z := (\boldsymbol{x}, \boldsymbol{\lambda}) \quad \text{and} \quad \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{\lambda} \end{pmatrix}.$$

Moreover, we indicate by $\boldsymbol{Z}$ the set of all solutions of system (3).

Under a suitable constraint qualification system (3) can be regarded as a first order necessary condition for the GNEP and indeed system (3) is akin to a KKT system. However, its structure is different from that of a classical KKT system. Under further convexity assumptions it can be easily seen that the $\boldsymbol{x}$-part of a solution of system (3) solves the GNEP so that (3) then turns out to be a sufficient condition as well. To formulate this result we first introduce some further terminology. Let $f_\nu : \mathbb{R}^n \to \mathbb{R}$ be a function attached to player ν and depending on all players' variables. We say that f_ν is *player convex* if, for every fixed $\boldsymbol{x}^{-\nu}$, the function $f_\nu(\cdot, \boldsymbol{x}^{-\nu})$ is convex in x^ν. If, instead, f_ν is convex with respect to $\boldsymbol{x}$, f_ν is called *jointly convex.*

Let a GNEP be given where each player's minimization problem is defined by (1) with the feasible set given by (2). We call this GNEP player convex if each player's objective function θ_ν and constraint functions g_i^ν, for $i = 1, \dots, m_\nu$, are player convex. Note that if a GNEP is player convex then, given $\boldsymbol{x}^{-\nu}$, the minimization problem of player ν is convex. Therefore, the minimum principle applied to every player readily yields the following result.

Proposition 1 *Let the GNEP be player convex. Then, for each solution* $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ *of system* (3) *the vector* $\bar{\boldsymbol{x}}$ *solves the GNEP.*

Player convexity is the standard setting under which GNEPs are usually investigated in the literature. With the exception of Sect. 3.2 we will not make use of any convexity assumptions for the functions defining the GNEP.

The key to our approach in this paper is a reformulation of system (3) as a possibly nonsmooth system of equations by using complementarity functions. A complementarity function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is a function such that

$$\phi(a, b) = 0 \iff a \geq 0,\ b \geq 0,\ ab = 0.$$

If ϕ is a complementarity function then (3) can be reformulated as the system

$$\boldsymbol{\Phi}(z) := \begin{pmatrix} \mathbf{L}(z) \\ \boldsymbol{\phi}(-\mathbf{g}(\boldsymbol{x}), \boldsymbol{\lambda}) \end{pmatrix} = 0, \tag{4}$$

where $\boldsymbol{\phi} : \mathbb{R}^{m+m} \to \mathbb{R}^m$ is defined, for all $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^m$, by

$$\boldsymbol{\phi}(\boldsymbol{a}, \boldsymbol{b}) := \begin{pmatrix} \phi(a_1, b_1) \\ \vdots \\ \phi(a_m, b_m) \end{pmatrix}.$$

Many complementarity functions are known (see, e.g., [13]). In this paper we will always use the min function and therefore set

$$\phi(a, b) := \min\{a, b\} \qquad \text{for all } a, b \in \mathbb{R}.$$

Since ϕ is not everywhere differentiable the mapping $\mathbf{\Phi}$ and equation (4) are called nonsmooth. The use of this complementarity function usually leads to the definition of Newton type methods that can be proven to have a fast local convergence under conditions that are among the weakest possible (see for example [13]). However, we emphasize that in principle another complementarity function could be used leading to a different Newton method. As a pointer to this fact we write ϕ instead of min. The classical Newton method cannot be applied to the solution of the equation $\mathbf{\Phi}(z) = 0$ because of its possible nonsmoothness at a solution. However methods have been developed in the past 20 years to cope with several kinds of nonsmoothness. One method we are interested in is the renowned semismooth Newton method. We refer the reader to [13] for a more complete exposition and for historical background. Among the many papers on the implementation of semismooth methods and their practical application we highlight [7,9,12,23].

By Rademacher's theorem a locally Lipschitzian function $H : \mathbb{R}^s \to \mathbb{R}^s$ is differentiable almost everywhere. Let $D_H \subseteq \mathbb{R}^s$ indicate the set where H is differentiable. Then,

$$\operatorname{Jac} H(y) := \left\{ V : V = \lim_{k\to\infty} JH(y^k) \text{ with } \{y^k\} \subset D_H,\ \lim_{k\to\infty} y^k = y \right\}.$$

defines the limiting Jacobian (or B-subdifferential) of H at y. The convex hull of $\operatorname{Jac} H(y)$ is known as Clarke generalized Jacobian, denoted by $\partial H(y)$.

Definition 1 Let $H : \mathbb{R}^s \to \mathbb{R}^s$ be Lipschitzian around $y \in \mathbb{R}^s$ and directionally differentiable at y. H is said to be *strongly semismooth* at y if for any $V \in \partial H(y+d)$,

$$Vd - H'(y; d) = O(\|d\|^2),$$

where $H'(y; d)$ is the directional derivative of H in y along the direction d.

In the study of algorithms for locally Lipschitzian systems of equations, the following regularity condition plays a role similar to that of the nonsingularity of the Jacobian in the study of algorithms for smooth systems of equations.

Definition 2 Let $H : \mathbb{R}^s \to \mathbb{R}^s$ be Lipschitzian around y. H is said to be *BD-regular* at y if all the elements in $\operatorname{Jac} H(y)$ are nonsingular. If $\bar{y}$ is a solution of the system $H(y) = 0$ and H is BD-regular at $\bar{y}$ then $\bar{y}$ is called a BD-regular solution of this system.

A generalized Newton method for the solution of a locally Lipschitzian system of equations $H(y) = 0$ can be defined as

$$y^{k+1} := y^k - (V^k)^{-1} H(y^k), \qquad V^k \in \operatorname{Jac} H(y^k); \tag{5}$$

(V^k can be any element in Jac $H(y^k)$). The following result holds.

Theorem 1 *Let $H : \mathbb{R}^s \to \mathbb{R}^s$ be strongly semismooth at a BD-regular solution $\bar{y}$ of $H(y) = 0$. Then, a neighborhood $\mathcal{N}$ of $\bar{y}$ exists so that if $y^0 \in \mathcal{N}$ then the iteration method (5) is well defined and generates a sequence $\{y^k\}$ that converges Q-quadratically to $\bar{y}$, i.e., there is a $C > 0$ so that*

$$\|y^{k+1} - \bar{y}\| \leq C\|y^k - \bar{y}\|^2 \quad \textit{for } k = 0, 1, 2, \ldots$$

It is well known that both the min function and any differentiable function with a locally Lipschitzian derivative are strongly semismooth. Since the composition of strongly semismooth functions is again strongly semismooth our general smoothness assumption on the GNEP implies that $\boldsymbol{\Phi}$ is strongly semismooth everywhere. However, referring to Sect. 3.3 we note that Theorem 1 is not the only source of results on quadratic convergence.

3 Newton methods for GNEPs: general description

In this section we examine three settings in which we develop Newton methods and show their fast local convergence. The settings represent, we believe, meaningful situations that are likely to be encountered often in practical applications. While we describe, for each setting, a corresponding Newton method and the assumptions needed for their analysis, a detailed investigation of the assumptions and of their relations is postponed to Sect. 4.

3.1 Semismooth Newton method for system (4)

The first and simplest Newton method we consider is nothing else than the semismooth Newton method applied to the reformulation (4) of system (3). Starting with an initial point $z^0 = (\boldsymbol{x}^0, \boldsymbol{\lambda}^0)$, this method generates a sequence $\{z^k\}$ according to the iteration

Newton Method I

$$z^{k+1} := z^k + \boldsymbol{d}^k,$$

where, for some $\boldsymbol{V}^k \in \text{Jac}\, \boldsymbol{\Phi}(z^k)$, $\boldsymbol{d}^k$ solves the linear system

$$\boldsymbol{V}^k \boldsymbol{d} = -\boldsymbol{\Phi}(z^k) \tag{6}$$

According to Theorem 1 and the discussion we made after it, this algorithm is well defined and has a quadratic convergence rate if z^0 is sufficiently close to a BD-regular solution $\bar{z}$ of (4). Therefore, we introduce the following definition.

Definition 3 A point $\bar{z} = (\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ is called *quasi-regular* if $\boldsymbol{\Phi}(\bar{z})$ is BD-regular at $\bar{z}$, i.e., if all the matrices in Jac $\boldsymbol{\Phi}(\bar{z})$ are nonsingular.

Theorem 1 now immediately gives us the following result.

Theorem 2 *Let* $\bar{z} = (\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ *be a* quasi-regular *solution of system* (3). *Then, a neighborhood* $\mathcal{N}$ *of* $\bar{z}$ *exists so that if* $z^0 \in \mathcal{N}$ *then Newton Method I is well defined and generates a sequence* $\{z^k\}$ *that converges Q-quadratically to* $\bar{z}$.

In the case of optimization problems, quasi-regularity is a rather weak condition [10]. However, as we shall analyze in detail in Sect. 4, in the case of system (3) quasi-regularity is a rather stringent assumption that might not be satisfied for several important classes of problems (see also Remark 3). In particular, we will see that quasi-regularity is never satisfied if even just two players share a constraint that is active.

3.2 Shared constraints: the "common multipliers" case

In this subsection we start a deeper investigation of what can happen when the players share some constraints, a most common circumstance. Then, as we mentioned at the end of the previous subsection, the quasi-regularity assumption is not likely to be satisfied. Below we suggest an alternative and simple Newton method that can be used under a set of conditions that, although somewhat restrictive, are often met in practice. Specifically, following [30], we assume that the feasible sets of the players are defined as

$$X_\nu(\boldsymbol{x}^{-\nu}) := \{x^\nu \in \mathbb{R}^{n_\nu} : s(\boldsymbol{x}) \leq 0, \;\; h^\nu(x^\nu) \leq 0\}. \tag{7}$$

Here $s : \mathbb{R}^n \to \mathbb{R}^{m_0}$ defines those constraints that are shared by *all* players and that can depend on all variables. We call the constraints $s(\boldsymbol{x}) \leq 0$ the *shared* constraints. In other words, the constraints $s(\boldsymbol{x}) \leq 0$ are the same for all players. Instead, the h^ν represent constraints that depend only on the variables of a single player. Note that (7) is a most common case in practice. In particular, it often happens that each player has its own constraints h^ν depending on his own decisions only plus additional constraints that represent the use of some common resource (a transmission channel for electricity or information, for example) that has a certain capacity. Then, the constraints $s(\boldsymbol{x}) \leq 0$ would simply be (linear) constraints that express that the capacity of the shared resources is limited (see Sect. 5 for an example of this type).

For any given $\boldsymbol{x}^{-\nu}$ the ν-th player's KKT conditions can be rewritten as

$$\begin{aligned} \nabla_{x^\nu}\theta_\nu(x^\nu, \boldsymbol{x}^{-\nu}) + \nabla_{x^\nu}h^\nu(x^\nu)\sigma^\nu + \nabla_{x^\nu}s(x^\nu, \boldsymbol{x}^{-\nu})\mu^\nu = 0 \\ 0 \leq \sigma^\nu \perp -h^\nu(x^\nu) \geq 0 \\ 0 \leq \mu^\nu \perp -s(x^\nu, \boldsymbol{x}^{-\nu}) \geq 0. \end{aligned}$$

Concatenating these KKT conditions for all players, we reobtain system (3), just with a different notation to take into account the specific structure (7) of the sets $X_\nu(\boldsymbol{x}^{-\nu})$. Assume further that a solution $(\bar{\boldsymbol{x}}, \bar{\sigma}^1, \ldots, \bar{\sigma}^N, \bar{\mu}^1, \ldots, \bar{\mu}^N)$ of the concatenated KKT systems satisfies $\bar{\mu}^1 = \cdots = \bar{\mu}^N := \bar{\mu}$, i.e., the multipliers of the shared constraints are equal for all players. This might appear as a "strange" requirement. We will see shortly that under appropriate conditions this is not so. For the time being we just accept the existence of such a solution. It is now easy to verify that $(\bar{\boldsymbol{x}}, \bar{\sigma}^1, \ldots, \bar{\sigma}^N, \bar{\mu})$ solves

$$\begin{array}{r} \mathbf{F}(\boldsymbol{x}) + \sum_{\nu=1}^N \nabla_{\boldsymbol{x}} h^\nu(\boldsymbol{x})\sigma^\nu + \nabla_{\boldsymbol{x}} s(\boldsymbol{x})\mu = 0 \\ 0 \leq \sigma^\nu \perp -h^\nu(\boldsymbol{x}) \geq 0 \\ 0 \leq \mu \perp -s(\boldsymbol{x}) \geq 0 \end{array} \tag{8}$$

with $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$ defined by $\mathbf{F}(\boldsymbol{x}) := \left(\nabla_{x^i}\theta_i(\boldsymbol{x})\right)_{i=1}^N$. For uniformity of notation, we wrote $h^\nu(\boldsymbol{x})$ instead of $h^\nu(x^\nu)$. Following the usual pattern we can now reformulate this system using the (min) complementarity function ϕ and rewrite it as the square system

$$\boldsymbol{\Phi}_{\text{VI}}(\boldsymbol{x}, \boldsymbol{\sigma}, \mu) := \begin{pmatrix} \mathbf{F}(\boldsymbol{x}) + \sum_{\nu=1}^N \nabla_{\boldsymbol{x}} h^\nu(\boldsymbol{x})\sigma^\nu + \nabla_{\boldsymbol{x}} s(\boldsymbol{x})\mu \\ \boldsymbol{\phi}(-h^1(x^1), \sigma^1) \\ \vdots \\ \boldsymbol{\phi}(-h^N(x^N), \sigma^N) \\ \boldsymbol{\phi}(-s(\boldsymbol{x}), \mu) \end{pmatrix} = 0. \tag{9}$$

At this point we can proceed as in the previous subsection and apply the semismooth Newton method to solve system $\boldsymbol{\Phi}_{\text{VI}}(\boldsymbol{x}, \boldsymbol{\sigma}, \mu) = 0$. Thus, starting with an initial point $\boldsymbol{w}^0 = (\boldsymbol{x}^0, \boldsymbol{\sigma}^0, \mu^0)$, the semismooth Newton method generates a sequence $\{\boldsymbol{w}^k\}$ according to the iteration

Newton Method II

$$\boldsymbol{w}^{k+1} := \boldsymbol{w}^k + \boldsymbol{d}^k,$$

where, for some $\boldsymbol{V}_{\text{VI}}^k \in \text{Jac}\, \boldsymbol{\Phi}_{\text{VI}}(\boldsymbol{w}^k)$, $\boldsymbol{d}^k$ solves the linear system

$$\boldsymbol{V}_{\text{VI}}^k \boldsymbol{d} = -\boldsymbol{\Phi}_{\text{VI}}(\boldsymbol{w}^k) \tag{10}$$

Newton Methods I and II result from the application of the same semismooth Newton method to two different systems of equations. To ensure that the Newton Method II is locally well defined and converges Q-quadratically to a solution of (9) we introduce the following assumption.

Definition 4 A point $\bar{\boldsymbol{w}} = (\bar{\boldsymbol{x}}, \bar{\boldsymbol{\sigma}}, \bar{\mu})$ is called *VI-quasi-regular* if all the matrices in $\text{Jac}\, \boldsymbol{\Phi}_{\text{VI}}(\bar{\boldsymbol{w}})$ are nonsingular.

Theorem 1 immediately gives us:

Theorem 3 *Let* $\bar{\boldsymbol{w}} = (\bar{\boldsymbol{x}}, \bar{\boldsymbol{\sigma}}, \bar{\mu})$ *be a* VI-quasi-regular *solution of system* (9). *Then, a neighborhood* $\mathcal{N}$ *of* $\bar{\boldsymbol{w}}$ *exists so that if* $\boldsymbol{w}^0 \in \mathcal{N}$ *the Newton Method II is well defined and generates a sequence* $\{\boldsymbol{w}^k\}$ *that converges Q-quadratically to* $\bar{\boldsymbol{w}}$.

Obviously, an important question must be answered before we can consider this as an acceptable approach: when does a solution of the GNEP exist such that the multipliers of the shared constraints are the same? To answer this question and also to better understand the nature of system (9) we study a particular setting that is common in practical applications. Let

$$\boldsymbol{X} := \{\boldsymbol{x} \in \mathbb{R}^n : s(\boldsymbol{x}) \leq 0,\ h^\nu(x^\nu) \leq 0,\ \nu = 1, \ldots, N\}$$

and assume that $s_1, \ldots, s_{m_0}$ are jointly convex while the h^ν ($\nu = 1, \ldots, N$) are componentwise convex so that $\boldsymbol{X}$ is convex. Then, it is easy to check that system (9) is nothing else than the KKT system of the VI($\boldsymbol{X}$, $\mathbf{F}$), that is the problem of finding an $\boldsymbol{x} \in \boldsymbol{X}$ such that

$$\mathbf{F}(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) \geq 0 \quad \text{for all } \boldsymbol{y} \in \boldsymbol{X}.$$

Based on these elements, it is possible to show the following result (see [11] for details).

Theorem 4 *Suppose that, for every player ν, the function θ_ν is player convex and the set $X_\nu(\boldsymbol{x}^{-\nu})$ is defined by (7) with a componentwise convex function h^ν. Moreover, assume that $s_1, \ldots, s_{m_0}$ are jointly convex. Then, every solution $\bar{\boldsymbol{x}}$ of the VI($\boldsymbol{X}$, $\mathbf{F}$) is a solution of the GNEP. Furthermore, if $\bar{\boldsymbol{x}}$ satisfies (9) for some multipliers $(\bar{\sigma}^1, \ldots, \bar{\sigma}^N, \bar{\mu})$ then $\bar{\boldsymbol{x}}$ is a solution of the GNEP such that the multipliers of the shared constraints are identical.*

An immediate consequence of this result is that if the VI($\boldsymbol{X}$, $\mathbf{F}$) has a solution satisfying its KKT conditions, the existence of a solution of the GNEP with identical multipliers for the shared constraint is guaranteed. To ensure that the VI($\boldsymbol{X}$, $\mathbf{F}$) has a solution, we can apply, in principle, any standard condition for the existence of a solution of VI. However, one must pay attention to the fact that the VI($\boldsymbol{X}$, $\mathbf{F}$) has a special structure and so many classical conditions for the existence of a solution of a VI could not be satisfied in practice. We do not dwell on these issues here, but simply observe that a reasonable condition is the compactness of the set $\boldsymbol{X}$ (see e.g., [13]). We now summarize the discussion so far in the following proposition.

Proposition 2 *Assume that the sets $X_\nu(\boldsymbol{x}^{-\nu})$ are defined by (7) with componentwise convex functions h^ν and jointly convex functions $s_1, \ldots, s_{m_0}$. If $\bar{\boldsymbol{x}}$ is a solution of the VI($\boldsymbol{X}$, $\mathbf{F}$) it also solves the GNEP. If any constraint qualification holds at $\bar{\boldsymbol{x}}$ the multipliers of the shared constraints are identical.*

The existence of a solution with identical multipliers for the shared constraints in the case of a compact $\boldsymbol{X}$ has long been known. In the seminal paper [30], Rosen analyzed (although from a rather different perspective) GNEPs with feasible sets defined by (7) with s jointly convex and h^ν componentwise convex. Among many interesting results he proved the existence of solutions with identical multipliers for the shared constraints and called them *normalized equilibria*. Rosen's paper has been very influential and his setting has been used in many subsequent works. Theorem 4, together with Theorem 3, shows that we can develop Newton methods in Rosen's setting. The search of a normalized equilibrium is meaningful from the practical point of view since this solution has a special interest in many applications, see [18].

Example 1 [1] To illustrate the developments in this subsection we consider

$$\begin{array}{rcl} \min_x (x-1)^2 & \qquad & \min_y (y-\frac{1}{2})^2 \\ x+y \le 1 & \qquad & x+y \le 1. \end{array}$$

The optimal solution sets of the two players are given by

$$\mathcal{S}(y) = \begin{cases} 1 & \text{if } y \le 0 \\ 1-y & \text{if } y \ge 0 \end{cases} \quad \text{and} \quad \mathcal{S}(x) = \begin{cases} 1/2 & \text{if } x < 1/2 \\ 1-x & \text{if } x \ge 1/2. \end{cases}$$

It is easy to check that this GNEP has infinitely many solutions given by $(\alpha, 1-\alpha)$ for $\alpha \in [1/2, 1]$. Since at each solution the linear independence constraint qualification holds, for each such solution there are unique multipliers $\lambda(\alpha)$, for the first player, and $\mu(\alpha)$, for the second player, that together with $(\alpha, 1-\alpha)$ satisfy the KKT conditions. Setting the gradients of the Lagrangians of the two problems to zero, we get $\lambda(\alpha) = 2-2\alpha$, $\mu(\alpha) = 2\alpha - 1$. Thus, only one solution exists with $\lambda(\alpha) = \mu(\alpha)$ which is obtained for $\alpha = 3/4$ with $(\bar{x}, \bar{y}) = (3/4, 1/4)$, $\bar{\lambda} = 1/2 = \bar{\mu}$. Consider now the VI$(\boldsymbol{X}, \mathbf{F})$, where

$$\boldsymbol{X} := \{(x, y) \in \mathbb{R}^2 : x + y \le 1\}, \qquad \mathbf{F}(x, y) := \begin{pmatrix} 2x-2 \\ 2y-1 \end{pmatrix}.$$

$\mathbf{F}$ is clearly strictly monotone. Thus, the VI has a unique solution which is given by $(3/4, 1/4)$; just check by using the definition of a VI. Furthermore, if we write down the KKT conditions for this VI, we see that the multiplier corresponding to the sole constraint that defines $\boldsymbol{X}$ is $1/2$ (i.e., the common value of the multipliers of the generalized Nash game in the solution $(3/4, 1/4)$). We see then that the original GNEP has infinitely many solutions while the VI$(\boldsymbol{X}, \mathbf{F})$ has only one solution: the solution of the generalized game for which the shared constraint has equal multipliers.

3.3 Shared constraints: the hard case

In Sect. 3.2 we discussed Newton methods for normalized equilibria. Unfortunately, the favorable feature of the existence of a solution with "common" multipliers can get lost as soon as any of the assumptions of Proposition 2 is violated. We illustrate this point by three examples. The first has shared constraints that are not jointly convex. In other words, the set $\boldsymbol{X}$ is not convex, although the sets $X_\nu(\boldsymbol{x}^{-\nu}) = \{x^\nu \in \mathbb{R}^{n_\nu} : (x^\nu, \boldsymbol{x}^{-\nu}) \in \boldsymbol{X}\}$ are convex.

[1] In order to make this and the following examples more readable we deviate from the general notation adopted and indicate the variables of the first player with x, those of the second with y and so forth. Similarly, the multipliers of the first player are λ, those of the second μ and so on.

Example 2 Consider the game with two players:

$$\begin{array}{rcl} \min_x & -x \\ xy & \leq & 1 \\ g(x) & \leq & 0 \end{array} \qquad \begin{array}{rcl} \min_y & (2y-1)^2 \\ xy & \leq & 1 \\ 0 \leq y & \leq & 1 \end{array} \qquad \text{with} \quad g(x) := \begin{cases} (x-1)^2 & \text{if } x < 1 \\ 0 & \text{if } 1 \leq x \leq 2 \\ (x-2)^2 & \text{if } x > 2. \end{cases}$$

Obviously, the feasible sets of the players are convex and can be expressed in the form of (7). Moreover, $\boldsymbol{X} = \{(x, y) : xy \leq 1,\ g(x) \leq 0,\ 0 \leq y \leq 1\}$ is compact but not convex. With the solution sets of the players given by:

$$\mathcal{S}(y) = \begin{cases} 2 & \text{if } y < 1/2 \\ 1/y & \text{if } 1/2 \leq y \leq 1 \\ \emptyset & \text{if } y > 1 \end{cases} \qquad \text{and} \qquad \mathcal{S}(x) = \begin{cases} 1/2 & \text{if } x < 2 \\ 1/x & \text{if } x \geq 2 \end{cases}$$

it can be seen that the GNEP has only one solution: $(2, 1/2)$. Equalling the gradients of the Lagrangians to zero we find that the multipliers of the shared constraint are $\lambda = 2$ for the first player and $\mu = 0$ for the second. Thus, if a shared constraint is not convex with respect to the variables of all players, common multipliers might not exist even if the feasible set $\boldsymbol{X}$ is compact.

In the next example the set $\boldsymbol{X}$ is convex, but the third player does not share a constraint that is shared by the first two players.

Example 3 For the game with three players

$$\begin{array}{r} \min_x \ -x \\ z \leq x + y \leq 1 \\ x \geq 0 \end{array} \qquad \begin{array}{r} \min_y \ (2y-1)^2 \\ z \leq x + y \leq 1 \\ y \geq 0 \end{array} \qquad \begin{array}{r} \min_z \ (2z-3x)^2 \\ 0 \leq z \leq 2 \end{array}$$

the solution sets are given by

$$\mathcal{S}(y, z) = \begin{cases} 1 - y & \text{if } y \leq 1,\ z \leq 1 \\ \emptyset & \text{otherwise,} \end{cases} \qquad \mathcal{S}(x, z) = \begin{cases} 1/2 & \text{if } z - x \leq 1/2,\ x \leq 1/2 \\ 1 - x & \text{if } 1/2 \leq x \leq 1,\ z \leq 1 \\ \emptyset & \text{otherwise,} \end{cases}$$

$$\text{and} \quad \mathcal{S}(x, y) = \begin{cases} 3x/2 & \text{if } 0 \leq x \leq 4/3 \\ 2 & \text{if } x > 4/3 \\ 0 & \text{if } x < 0. \end{cases}$$

This GNEP has infinitely many solutions given by $(\alpha, 1-\alpha, (3/2)\alpha)$ for $\alpha \in [1/2, 2/3]$. Let the multipliers of the constraints $x + y \leq 1$ and $z \leq x + y$ be denoted by λ_1 and λ_2 for the first player and by μ_1 and μ_2 for the second player. We want to check that $(\lambda_1, \lambda_2) = (\mu_1, \mu_2)$ never occurs. In fact, equalling the gradient of the Lagrangian of the first player to zero, we obtain $\lambda_1 = 1 + \lambda_2$ (with $\lambda_2 = 0$ if $\alpha < 2/3$), while we get $\mu_1 = 2\alpha - 1 + \mu_2$ (with $\mu_2 = 0$ if $\alpha < 2/3$) for the second player. From these relations it is trivial to see that for no value of $\alpha \in [1/2, 2/3]$ we can have $\lambda_1 = \mu_1$ and $\lambda_2 = \mu_2$.

Remark 2 A key issue in the previous example is that the constraint $z \le x + y$ shared only by the first two players depends also on the variables of the third player. If this was not the case we could append the constraint shared by the first two players to those of the third one without changing his feasible set, thus reducing the problem to the setting of Sect. 3.2. Generalizing this observation, we see that the setting of Sect. 3.2 covers also those cases where if a constraint $s_i(\boldsymbol{x}) \le 0$ actually depends on the variables of a subset of players, then s_i is shared by all the players in this subset.

The last example shows that even when the GNEP has jointly convex constraints and player convex objective functions, if $\boldsymbol{X}$ is not bounded then a solution to (3) with common multipliers may not exists.

Example 4 For the game

$$\begin{array}{cc} \min_x \ -x & \min_y \ -2y \\ x + y \le 1 & x + y \le 1 \end{array}$$

the players' solution sets are $\mathcal{S}(y) = 1 - y$ and $\mathcal{S}(x) = 1 - x$. Moreover, each solution of the GNEP is given by $(\alpha, 1-\alpha)$ with $\alpha \in \mathbb{R}$. For all solutions the common constraint is active. The multipliers are 1 for the first player and 2 for the second player and do not depend on the solution.

Examples 2–4 convincingly show that the approach of Sect 3.2 cannot be easily extended to other cases. We are then left with the question: what can we do if we have shared constraints but (possibly) no "common multiplier"?

To motivate our approach, we first show that a solution $\bar{\boldsymbol{x}}$ of the GNEP at which a shared constraint is active is not likely to be an isolated solution (see also Examples 1, 3, and 4). To this end suppose that $\bar{z} = (\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ is a solution of system (3) at which strict complementarity (see Assumption 1 below) holds. Then, $\boldsymbol{\Phi}$ is continuously differentiable around $\bar{z}$, so that $\partial\boldsymbol{\Phi}(\bar{z})$ reduces to the singleton $J\boldsymbol{\Phi}(\bar{z})$. If a constraint is active and shared by more than one player, this entails the singularity of $J\boldsymbol{\Phi}(\bar{z})$ (see the next section for a formal proof). Consider the function $\boldsymbol{\Phi}_{\bar{z}}$ obtained from $\boldsymbol{\Phi}$ by removing all the rows corresponding to this repeated and active constraint except one; repeat the procedure for all the repeated and active constraints. This amounts to leaving in the system $\boldsymbol{\Phi}_{\bar{z}}(z) = 0$ only one copy for each active repeated constraint. We call $\boldsymbol{\Phi}_{\bar{z}}(z) = 0$ the reduced system (to be described in more detail in Sect. 4.3). Due to the strict complementarity assumption it is not difficult to see that in a neighborhood of $\bar{z}$, a point z is a solution of system (3) if and only if $\boldsymbol{\Phi}_{\bar{z}}(z) = 0$ (see Lemma 1 for details). The Jacobian $J\boldsymbol{\Phi}_{\bar{z}}$ has more columns than rows. Assume now that $J\boldsymbol{\Phi}_{\bar{z}}(\bar{z})$ has full row rank. This condition appears to be rather natural and favorable. However, by the implicit function theorem we see that the solution $\bar{z}$ is not locally unique. Therefore any standard Newton method will have serious difficulties in this case. In the field of Nonlinear Programming there are developments that deal with Newton methods for cases when nonisolated KKT points occur. However, these developments are mainly concerned with the nonisolatedness of the multiplier part of the KKT system, whereas the primal solution of the minimization problem is locally unique. In our case we might often expect that even the primal part is not locally unique. In fact, if in the setting we are considering in this paragraph the LICQ (Linear Independence Constraint Qualification) holds at $\bar{z}$ for every player, then for any fixed $\bar{\boldsymbol{x}}$ there is a unique multiplier $\bar{\boldsymbol{\lambda}}$. Thus, if $\bar{z} = (\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ is a nonisolated solution of (3), the $\boldsymbol{x}$-part must be nonisolated. We now summarize the discussion.

Proposition 3 *Let a GNEP be given as described in Sect. 2. Assume that a constraint is shared by at least two players. Let $\bar{z} = (\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ be a solution of (3) at which this shared constraint is active and assume that strict complementarity holds at $\bar{z}$ for all constraints of all players. Assume further that the matrix $J\boldsymbol{\Phi}_{\bar{z}}(\bar{z})$ has full row rank and that the LICQ holds for every player. Then the solution $\bar{\boldsymbol{x}}$ and the KKT point $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ are nonisolated.*

Now it is clear that to develop a Newton method in the case of repeated, non jointly convex constraints a non standard approach is needed: the choice is very restricted. We consider a recently developed Levenberg–Marquardt type method [32] that can deal with the nonisolatedness of solutions. Since we are interested in the local convergence behavior, a solution $\bar{z}$ of (3) is fixed throughout this subsection. The following two assumptions will be needed.

Assumption 1 Strict complementarity holds at $\bar{z} = (\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$, i.e., $g_i^\nu(\bar{\boldsymbol{x}}) = 0$ implies $\bar{\lambda}_i^\nu > 0$ for arbitrary $\nu = 1, \dots, N$ and $i = 1, \dots, m_\nu$.

This assumption guarantees the differentiability of $\boldsymbol{\Phi}$ and the Lipschitz continuity of $J\boldsymbol{\Phi}$ in a neighborhood of $\bar{z}$. By now, in principle, such a smoothness condition is needed for proving quadratic convergence of Levenberg–Marquardt type methods in the case of nonisolated solutions, see [15,16,32].

Assumption 2 There are $c, \delta > 0$ so that $\|\boldsymbol{\Phi}(z)\| \geq c \operatorname{dist}[z, \boldsymbol{Z}]$ for all $z \in \mathbb{B}(\bar{z}, \delta)$.

We recall that $\boldsymbol{Z}$ is the solution set of system (3). Assumption 2 requires that $\|\boldsymbol{\Phi}\|$ be a local error bound. Conditions under which this assumption is satisfied will be discussed in Sect. 4.3. The Levenberg–Marquardt method starts from some z^0 and generates a sequence $\{z^k\}$ as follows:

Newton Method III

$$z^{k+1} := z^k + \boldsymbol{d}^k,$$

where, with $\alpha(z^k) := \|\boldsymbol{\Phi}(z^k)\|$, $\boldsymbol{d}^k$ solves the linear system

$$J\boldsymbol{\Phi}(z^k)^\top \boldsymbol{\Phi}(z^k) + (J\boldsymbol{\Phi}(z^k)^\top J\boldsymbol{\Phi}(z^k) + \alpha(z^k)I)\boldsymbol{d} = 0 \tag{11}$$

Newton Method III is well defined (as long as z^k does not belong to the solution set $\boldsymbol{Z}$) since the subproblems (11) always have a unique solution if $\alpha(z^k) > 0$.

Theorem 5 *Let $\bar{z}$ be a solution of the system (3) at which Assumptions 1 and 2 hold. If z^0 is sufficiently close to $\bar{z}$, then the sequence $\{z^k\}$ produced by the Levenberg–Marquardt either generates a solution in $\mathbf{Z}$ after a finite number of steps or converges Q-quadratically to some $\hat{z} \in \mathbf{Z}$.*

Proof If $\alpha(z)$ in Newton Method III is replaced by $\|\boldsymbol{\Phi}(z)\|^2$ the quadratic convergence follows from [32, Theorem 2.1]. For $\alpha(z) = \|\boldsymbol{\Phi}(z)\|$ as in Newton Method III the quadratic convergence is shown in [15, Theorem 2.2] and [16, Theorem 10]. The result in [16] makes use of $\tilde{\alpha}(z) := \|J\boldsymbol{\Phi}(z)^T\boldsymbol{\Phi}(z)\|$. However, by [16, Theorem 9] it can be easily seen that $\kappa_0\alpha(z) \le \tilde{\alpha}(z) \le \kappa_1\alpha(z)$ holds in a neighborhood of $\bar{z}$ for some $\kappa_0, \kappa_1 > 0$. □

The Assumption 2 that $\|\boldsymbol{\Phi}\|$ is a local error bound seems to be the crucial assumption in Theorem 5. Unfortunately, not much is known of error bounds for a generalized Nash equilibrium problem and the related system (3). In the next section we will undertake a preliminary study of this issue.

4 Newton methods for GNEPs: Analysis of the assumptions

In this section we analyze in detail the conditions used in the Sect. 3 to establish the convergence rates of the three Newton methods. We also discuss how these conditions are related and compare them to results in the literature.

4.1 Quasi-regularity

Our first task is to calculate the limiting Jacobian of $\boldsymbol{\Phi}$ at a solution $\bar{z}$ of (4), where ϕ is the min-function. Standard nonsmooth calculus yields

$$\operatorname{Jac}\boldsymbol{\Phi}(\bar{z}) = \{H(\boldsymbol{\gamma}) : \boldsymbol{\gamma} \in \boldsymbol{\Gamma}\} \quad \text{with} \quad H(\boldsymbol{\gamma}) = \begin{pmatrix} A & B \\ C(\boldsymbol{\gamma}) & D(\boldsymbol{\gamma}) \end{pmatrix}. \tag{12}$$

Here A and B are fixed matrices given by:

$$A := J_{\boldsymbol{x}}\mathbf{L}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}}) \quad \text{and} \quad B := J_{\boldsymbol{\lambda}}\mathbf{L}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}}) = \operatorname{diag}\left(J_{x^\nu} g^\nu(\bar{\boldsymbol{x}})^\top\right).$$

The matrix A can be expanded in the following form

$$A := \begin{pmatrix} A_{11} & \cdots & A_{1\nu} & \cdots & A_{1N} \\ \vdots & & \vdots & & \vdots \\ A_{N1} & \cdots & A_{N\nu} & \cdots & A_{NN} \end{pmatrix}$$

with

$$A_{\nu_1\nu_2} \;=\; J_{x^{\nu_2}} \nabla_{x^{\nu_1}} L_{\nu_1}(\bar{\boldsymbol{x}}, \bar{\lambda}^{\nu_1}) \;=\; J_{x^{\nu_2}}\left(\nabla_{x^{\nu_1}}\theta_{\nu_1}(\bar{\boldsymbol{x}}) + \sum_{\ell=1}^{m_{\nu_1}} \nabla_{x^{\nu_1}} g_\ell^{\nu_1}(\bar{\boldsymbol{x}})\bar{\lambda}_\ell^{\nu_1}\right).$$

To describe instead the matrices $C(\boldsymbol{\gamma})$ and $D(\boldsymbol{\gamma})$, that depend on a parameter γ, let us first define for each player ν the sets of active, strongly active, degenerate, and non active indices by:

$$\begin{array}{ll} I_0^\nu \;:=\; \{\, i \in \{1, \ldots m_\nu\} : g_i^\nu(\bar{\boldsymbol{x}}) = 0\,\}, & I_+^\nu \;:=\; \{\, i \in I_0^\nu : \bar{\lambda}_i^\nu > 0\,\}, \\ I_{00}^\nu \;:=\; \{\, i \in I_0^\nu : \bar{\lambda}_i^\nu = 0\,\}, & I_<^\nu \;:=\; \{\, i \in \{1, \ldots m_\nu\} : g_i^\nu(\bar{\boldsymbol{x}}) < 0\,\}. \end{array}$$

For any subset $\gamma^\nu \subseteq I_{00}^\nu$ (empty set included), we set

$$\alpha^\nu := I_+^\nu \cup \gamma^\nu \qquad \text{and} \qquad \beta^\nu := I_<^\nu \cup (I_{00}^\nu \setminus \gamma^\nu).$$

Then, for any

$$\boldsymbol{\gamma} \in \boldsymbol{\Gamma} := \{(\gamma^1, \ldots, \gamma^N) \;:\; \gamma^\nu \subseteq I_{00}^\nu \text{ for } \nu = 1, \ldots, N\},$$

the matrices $C(\boldsymbol{\gamma})$ and $D(\boldsymbol{\gamma})$ are given by

$$C(\boldsymbol{\gamma}) := -\begin{pmatrix} J_{x^1} g_{\alpha^1}^1(\bar{\boldsymbol{x}}) & \cdots & J_{x^\nu} g_{\alpha^1}^1(\bar{\boldsymbol{x}}) & \cdots & J_{x^N} g_{\alpha^1}^1(\bar{\boldsymbol{x}}) \\ \mathbf{0}_{|\beta^1|\times n_1} & \cdots & \mathbf{0}_{|\beta^1|\times n_\nu} & \cdots & \mathbf{0}_{|\beta^1|\times n_N} \\ \vdots & & \vdots & & \vdots \\ J_{x^1} g_{\alpha^N}^N(\bar{\boldsymbol{x}}) & \cdots & J_{x^\nu} g_{\alpha^N}^N(\bar{\boldsymbol{x}}) & \cdots & J_{x^N} g_{\alpha^N}^N(\bar{\boldsymbol{x}}) \\ \mathbf{0}_{|\beta^N|\times n_1} & \cdots & \mathbf{0}_{|\beta^N|\times n_\nu} & \cdots & \mathbf{0}_{|\beta^N|\times n_N} \end{pmatrix} \tag{13}$$

and

$$D(\boldsymbol{\gamma}) := \operatorname{diag}(D^\nu(\boldsymbol{\gamma})),$$

where $D^\nu(\boldsymbol{\gamma})$ is an $m_\nu \times m_\nu$ matrix with

$$D^\nu(\boldsymbol{\gamma}) \;:=\; \begin{pmatrix} \mathbf{0}_{|\alpha^\nu|} & \mathbf{0}_{|\alpha^\nu|\times|\beta^\nu|} \\ \mathbf{0}_{|\beta^\nu|\times|\alpha^\nu|} & \mathbf{I}_{|\beta^\nu|} \end{pmatrix}. \tag{14}$$

This completes the description of formula (12). To facilitate the understanding of the results that follow it is useful to show the structure of the matrices $H(\boldsymbol{\gamma})$ in more detail; for simplicity the dependence on $\bar{\boldsymbol{x}}$ is suppressed.

$$\left(\begin{array}{ccc|cccc} J_{x^1}\nabla_{x^1}L_1 & \cdots & J_{x^N}\nabla_{x^1}L_1 & J_{x^1}[g^1_{\alpha^1}]^\top \; J_{x^1}[g^1_{\beta^1}]^\top & & \\ \vdots & & \vdots & & \ddots & \\ J_{x^1}\nabla_{x^N}L_N & \cdots & J_{x^N}\nabla_{x^N}L_N & & & J_{x^N}[g^N_{\alpha^N}]^\top \; J_{x^N}[g^N_{\beta^N}]^\top \\ \hline -J_{x^1}g^1_{\alpha^1} & \cdots & -J_{x^N}g^1_{\alpha^1} & \mathbf{0}_{|\alpha^1|} \; \mathbf{0}_{|\alpha^1|\times|\beta^1|} & & \\ \mathbf{0}_{|\beta^1|\times n_1} & & \mathbf{0}_{|\beta^1|\times n_N} & \mathbf{0}_{|\beta^1|\times|\alpha^1|} \; \mathbf{I}_{|\beta^1|} & & \\ \vdots & & \vdots & & \ddots & \\ -J_{x^1}g^N_{\alpha^N} & & -J_{x^N}g^N_{\alpha^N} & & & \mathbf{0}_{|\alpha^N|} \; \mathbf{0}_{|\alpha^N|\times|\beta^N|} \\ \mathbf{0}_{|\beta^N|\times n_1} & \cdots & \mathbf{0}_{|\beta^N|\times n_N} & & & \mathbf{0}_{|\beta^N|\times|\alpha^N|} \; \mathbf{I}_{|\beta^N|} \end{array}\right). \tag{15}$$

Remark 3 A look at this matrix, in particular to the two lower blocks C and D, clearly shows that if a shared constraint is active at $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ there are some matrices in Jac $\boldsymbol{\Phi}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ that are singular. These are those matrices that correspond to $\boldsymbol{\gamma} \in \boldsymbol{\Gamma}$ for which there are at least two players ν_a and ν_b with a shared constraint for which the (index of the) shared constraint belongs to α^{ν_a} and α^{ν_b}. This will generate two identical rows in the matrix $H(\boldsymbol{\gamma})$ (the gradient of the shared constraints with respect to $\boldsymbol{x}$ followed by zeros).

To better understand quasi-regularity, it is useful to study some characterizations. To this end we define a family of *reduced* matrices $H_R(\boldsymbol{\gamma})$, that are obtained from $H(\boldsymbol{\gamma})$ by deleting the rows and the columns corresponding to the identity matrices in $D(\boldsymbol{\gamma})$. Note therefore that the matrices $H_R(\boldsymbol{\gamma})$ may have different dimensions according to the cardinality of the sets γ^ν. Due to the Laplace formula for the calculation of the determinant of a matrix the following proposition holds.

Proposition 4 *Let $\bar{\boldsymbol{z}} = (\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ be a solution of system* (3). *The point $\bar{\boldsymbol{z}}$ is quasi-regular if and only if all the matrices $H_R(\boldsymbol{\gamma})$ are nonsingular.*

Consider the case of a GNEP with only one player, i.e., an optimization problem. Then, Proposition 4 shows that our definition of quasi-regularity is an extension of the one introduced in [10] (see also [13]) for the KKT triples of a VI. The definition of quasi-regularity in [10], in fact, is phrased directly in terms of what we call here reduced matrices $H_R(\gamma)$. Note also that the definition of quasi-regularity in [10] refers to a KKT point of any VI.

Having defined the notion of quasi-regularity for system (3), it is in order to understand the relation between this condition and what can be considered as the most standard and fundamental regularity condition in the field of variational inequalities: Robinson's strong regularity. The following theorem extends to system (3) a result that is well known in the case of a KKT system of a VI (or of an optimization problem), see [13]. To facilitate its proof, (3) is considered as a mixed complementarity problem MiCP($\mathbf{K}$, $\boldsymbol{G}$) with

$$\boldsymbol{G}(\boldsymbol{z}) := \begin{pmatrix} \mathbf{L}(\boldsymbol{z}) \\ -\mathbf{g}(\boldsymbol{x}) \end{pmatrix} \quad \text{and} \quad \mathbf{K} := \mathbb{R}^n \times \mathbb{R}^m_+ \tag{16}$$

(recall that the MiCP($\mathbf{K}$, $\boldsymbol{G}$) is the VI($\mathbf{K}$, $\boldsymbol{G}$), the name mixed complementarity being normally used for this variational inequality with the set $\mathbf{K}$ having the special structure in (16)). Note that z is a solution of system (3) if and only if it is a solution of the MiCP($\mathbf{K}$, $\boldsymbol{G}$). This point of view will be used also in some other technical developments later in this section.

Theorem 6 *Let $\bar{z}$ be a solution of system* (3). *Then, $\bar{z}$ is strongly regular if and only if all the matrices $H_R(\boldsymbol{\gamma})$, $\boldsymbol{\gamma} \in \boldsymbol{\Gamma}$, have the same nonzero determinantal sign.*

Proof We will make use of Theorem 5.3.24 in [13] which requires some complicated matrices and sets which are defined in [13] just before Theorem 5.3.24. Below we review these definitions and the relevant part of Theorem 5.3.24 taking into account that in our case, due to the simple structure of the MiCP($\mathbf{K}$, $\boldsymbol{G}$), things simplify considerably.

First note that the constraints ($\boldsymbol{\lambda} \geq 0$) of the MiCP($\mathbf{K}$, $\boldsymbol{G}$) are linear and their gradients are linearly independent. Therefore the KKT conditions for the MiCP($\mathbf{K}$, $\boldsymbol{G}$) hold at $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ and, by the linear independence of the active gradients, there is a unique corresponding multiplier denoted by $\bar{\boldsymbol{\mu}} \in \mathbb{R}^m$. Let

$$\mathbf{I}_0 := \{ i \, : \, \bar{\boldsymbol{\mu}}_i > 0 = \bar{\boldsymbol{\lambda}}_i \}, \quad \mathbf{I}_{00} := \{ i \, : \, \bar{\boldsymbol{\mu}}_i = 0 = \bar{\boldsymbol{\lambda}}_i \}$$

and define the family of index sets $\mathbf{J} := \{J : \mathbf{I}_0 \subseteq J \subseteq \mathbf{I}_0 \cup \mathbf{I}_{00}\}$ (this is, in our setting, what is called $\mathcal{J}(\lambda)$ on p. 470 of [13]). Consider now the family of matrices B_J given, for each $J \in \mathbf{J}$, by the $|J| \times (n+m)$ matrix whose rows are $-e_i{}^{\top}$, $i \in J$, where the components of $e_i \in \mathbb{R}^n$ are 0 except the i-th one which is equal to 1; in other words, e_i is the gradient with respect to z of the i-th constraint in J (the set of all matrices B_J forms what is called $\mathcal{B}^{\lambda}_{\text{bas}}(\mathcal{C})$ on p. 470 of [13]). Under our conditions, Theorem 5.3.24 in [13] states (among other things) that $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ is strongly stable if and only if the matrices

$$\begin{pmatrix} J_z \boldsymbol{G}(\bar{z}) & B_J^{\top} \\ -B_J & 0 \end{pmatrix} \tag{17}$$

have the same nonzero determinantal sign for all $J \in \mathbf{J}$.

From the KKT conditions for MiCP($\mathbf{K}$, $\boldsymbol{G}$) one has $\bar{\boldsymbol{\mu}} = -\mathbf{g}(\bar{\boldsymbol{x}})$. This implies $\mathbf{I}_0 = \cup_{\nu=1}^N I_<^{\nu}$ and $\mathbf{I}_{00} = \cup_{\nu=1}^N I_{00}^{\nu}$. Therefore, applying Laplace's formula to a column of $B_J^{\top}$ and to the corresponding row of $-B_J$, recursively until all the columns of $B_J^{\top}$ and all the rows of $-B_J$ have been eliminated, one can check that the matrices (17) having the same nonzero determinantal sign is equivalent to the matrices $H_R(\boldsymbol{\gamma})$, $\boldsymbol{\gamma} \in \boldsymbol{\Gamma}$, having the same nonzero determinantal sign. Since the equivalence of strong stability and strong regularity is a standard result (see for example [13, p. 447]) this concludes the proof. □

This result, together with Proposition 4 shows that quasi-regularity is implied by strong regularity. In Sect. 5 we will also give an example showing that the reverse is not true, so that quasi-regularity is weaker than strong regularity.

4.2 VI-quasi-regularity

VI-quasi-regularity is in some sense easier to analyze, because it is just quasi-regularity of a KKT system of the VI$(\boldsymbol{X}, \mathbf{F})$ and as such has been already analyzed in [10,13]. The interesting thing we are left to do is to establish how VI-quasi-regularity relates to quasi-regularity. Since the comparison is meaningful only under the conditions for which VI-quasi-regularity can be defined we use the notation of Sect. 3.2, assume that every player in the GNEP has the feasible set defined by (7), and that $\bar{\boldsymbol{x}}$ is a normalized equilibrium of the GNEP with multipliers $\bar{\boldsymbol{\sigma}}$ and $\bar{\mu}$ so that $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\sigma}}, \bar{\mu})$ solves (8).

In order to perform this comparison we have to write down the explicit expression of Jac $\boldsymbol{\Phi}_{\mathrm{VI}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\sigma}}, \bar{\mu})$. To this end let us define some set of indices:

$$\begin{aligned}
&I_0^\nu := \{i \in \{1, \dots m_\nu\} : h_i^\nu(\bar{x}^\nu) = 0\}, && I_+^\nu := \{i \in I_0^\nu : \bar{\sigma}_i^\nu > 0\},\\
&I_{00}^\nu := \{i \in I_0^\nu : \bar{\sigma}_i^\nu = 0\}, && I_<^\nu := \{i \in \{1, \dots m_\nu\} : h_i^\nu(\bar{x}^\nu) < 0\},\\
&I_0^s := \{i \in \{1, \dots m_0\} : s_i(\bar{\boldsymbol{x}}) = 0\}, && I_+^s := \{i \in I_0^s : \bar{\mu}_i > 0\},\\
&I_{00}^s := \{i \in I_0^s : \bar{\mu}_i = 0\}, && I_<^s := \{i \in \{1, \dots m_0\} : s_i(\bar{\boldsymbol{x}}) < 0\}.
\end{aligned}$$

For any subset $\gamma^\nu \subseteq I_{00}^\nu$ we set $\alpha^\nu := I_+^\nu \cup \gamma^\nu$ and $\beta^\nu := I_<^\nu \cup (I_{00}^\nu \setminus \gamma^\nu)$. In a similar way, for any subset $\gamma^s \subseteq I_{00}^s$ we set $\alpha^s := I_+^s \cup \gamma^s$ and $\beta^s := I_<^s \cup (I_{00}^s \setminus \gamma^s)$. Then, for all possible $\boldsymbol{\gamma} := (\gamma^1, \dots, \gamma^N)$, and γ^s, the matrix

$$\mathcal{H}(\boldsymbol{\gamma}, \gamma^s) := \begin{pmatrix} \mathcal{A} & \mathcal{B} & \mathcal{E} \\ \mathcal{C}(\boldsymbol{\gamma}) & \mathcal{D}(\boldsymbol{\gamma}) & \mathbf{0} \\ \mathcal{F}(\gamma^s) & \mathbf{0} & \mathcal{M}(\gamma^s) \end{pmatrix} \tag{18}$$

belongs to Jac $\boldsymbol{\Phi}_{\mathrm{VI}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\sigma}}, \bar{\mu})$. Vice versa, any matrix in this limiting Jacobian can be obtained by (18) for suitably chosen $(\boldsymbol{\gamma}, \gamma^s)$. To point out the differences between the matrices in (12) and those in (18) we use slightly different notations. The matrices in (18) are given by

$$\begin{aligned}
&\mathcal{A} := J_{\boldsymbol{x}} \boldsymbol{\mathcal{L}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\sigma}}, \bar{\mu}) \ \text{ with } \ \boldsymbol{\mathcal{L}}(\boldsymbol{x}, \boldsymbol{\sigma}, \mu) := \mathbf{F}(\boldsymbol{x}) + \sum_{\nu=1}^N \nabla_{\boldsymbol{x}} h^\nu(\boldsymbol{x}) \sigma^\nu + \nabla_{\boldsymbol{x}} s(\boldsymbol{x}) \mu,\\
&\mathcal{B} := J_{\boldsymbol{\sigma}} \boldsymbol{\mathcal{L}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\sigma}}, \bar{\mu}) = \operatorname{diag}(J_{x^\nu} h^\nu(\bar{x}^\nu)^\top),\\
&\mathcal{E} := J_{\mu} \boldsymbol{\mathcal{L}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\sigma}}, \bar{\mu}) = J_{\boldsymbol{x}} s(\bar{\boldsymbol{x}})^\top,\\
&\mathcal{C}(\boldsymbol{\gamma}) := \operatorname{diag}(\mathcal{C}^\nu(\boldsymbol{\gamma})) \qquad \text{with} \qquad \mathcal{C}^\nu(\boldsymbol{\gamma}) := -\begin{pmatrix} J_{x^\nu} h_{\alpha^\nu}^\nu(x^\nu) \\ \mathbf{0}_{|\beta^\nu| \times n_\nu} \end{pmatrix},\\
&\mathcal{D}(\boldsymbol{\gamma}) := \operatorname{diag}(\mathcal{D}^\nu(\boldsymbol{\gamma})) \qquad \text{with} \qquad \mathcal{D}^\nu(\boldsymbol{\gamma}) := \begin{pmatrix} \mathbf{0}_{|\alpha^\nu|} & \mathbf{0}_{|\alpha^\nu| \times |\beta^\nu|} \\ \mathbf{0}_{|\beta^\nu| \times |\alpha^\nu|} & \mathbf{I}_{|\beta^\nu|} \end{pmatrix},\\
&\mathcal{F}(\gamma^s) := -\begin{pmatrix} J_{\boldsymbol{x}} s_{\alpha^s}(\bar{\boldsymbol{x}}) \\ \mathbf{0}_{|\beta^s| \times n} \end{pmatrix} \qquad \text{and} \qquad \mathcal{M}(\gamma^s) := \begin{pmatrix} \mathbf{0}_{|\alpha^s|} & \mathbf{0}_{|\alpha^s| \times |\beta^s|} \\ \mathbf{0}_{|\beta^s| \times |\alpha^s|} & \mathbf{I}_{|\beta^s|} \end{pmatrix}.
\end{aligned}$$

We are now able to show the fact we already hinted at above, that if we are in a solution of system (8), then VI-quasi-regularity is weaker than quasi-regularity. In fact, suppose that we are in the setting of this subsection and suppose also that a solution $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\sigma}}, \bar{\mu}^1, \dots, \bar{\mu}^N)$ is quasi-regular (for system (3)). This obviously implies that no shared constraint is active (see Remark 3), and hence $\bar{\mu}^1 = \cdots = \bar{\mu}^N = 0$. Then, we get $I^s_{00} = I^s_+ = \emptyset$ and $\gamma^s = \emptyset$ and for the matrices in (18) we have $\mathcal{M}(\gamma^s) = \mathbf{I}_{m_0}$ and $\mathcal{F}(\gamma^s) = \mathbf{0}_{m_0 \times n}$. Thus, the nonsingularity of the matrices (18) is equivalent to the nonsingularity of

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C}(\boldsymbol{\gamma}) & \mathcal{D}(\boldsymbol{\gamma}) \end{pmatrix}. \tag{19}$$

It is easy to show that the nonsingularity of all matrices (15) (suitably reduced due to the absence of shared active constraints) implies the nonsingularity of the matrices (19). Vice versa, it is rather intuitive that the inverse implication cannot hold. We illustrate this by considering again Example 1.

Example 1 (continued) As shown in Sect. 3.2 this game has infinitely many solutions given by $(\alpha, 1-\alpha)$ for $\alpha \in [1/2, 1]$. At none of them quasi-regularity holds. This can be verified directly or deduced by observing that quasi-regularity implies the local uniqueness of the solution, which is not satisfied. Consider now the solution $(\bar{x}, \bar{y}, \bar{\mu}) := (3/4, 1/4, 1/2)$ of the VI-KKT system (8). Since each player has exactly one constraint (the shared one) and by $\bar{\mu} = 1/2 > 0$ it holds that $I^s_{00} = \emptyset$. Therefore, VI-quasi-regularity at $(\bar{x}, \bar{y}, \bar{\mu})$ simply amounts to the nonsingularity of the matrix

$$\begin{pmatrix} \mathcal{A} & \mathcal{E} \\ \mathcal{F}(\emptyset) & \mathcal{M}(\emptyset) \end{pmatrix} = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ -1 & -1 & 0 \end{pmatrix}$$

which obviously holds. Note that at the solution $(3/4, 1/4, 1/2, 1/2)$ of the KKT system (3), $\gamma^\nu = \emptyset$ for $\nu = 1, 2$, and the corresponding matrix $H_R(\boldsymbol{\gamma})$ with $\boldsymbol{\gamma} = (\emptyset, \emptyset)$ is given by

$$H_R(\boldsymbol{\gamma}) = H(\boldsymbol{\gamma}) = \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C}(\boldsymbol{\gamma}) & \mathcal{D}(\boldsymbol{\gamma}) \end{pmatrix} = \begin{pmatrix} 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ -1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix}$$

which is obviously singular.

In Sect. 5.2 we give another example where VI-quasi-regularity is satisfied while there are shared constraints that are active so that quasi-regularity cannot hold. Therefore we can conclude that the following theorem is valid.

Theorem 7 *Assume that for every player the feasible set is defined by* (7)*, and let* $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\sigma}}, \bar{\mu})$ *be a solution of system* (8)*. Then, if the corresponding solution* $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ *of system* (3) *with* $\bar{\boldsymbol{\lambda}} := (\bar{\boldsymbol{\sigma}}, \bar{\mu}, \dots, \bar{\mu})$ *is quasi-regular,* $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\sigma}}, \mu)$ *is VI-quasi-regular. On the contrary, if* $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\sigma}}, \bar{\mu})$ *is VI-quasi-regular then* $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ *is a solution of* (3) *which is not necessarily quasi-regular.*

4.3 Error bounds

Our aim here is not to perform an exhaustive analysis, which would be impossible, but rather to show that there are important cases where the error bound condition of Assumption 2 holds, thus indicating the reasonability of the assumption itself. Our first result regards a special, but interesting case.

Theorem 8 *Suppose that the* $\boldsymbol{G} : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ (*see* (16)) *is an affine map. Then, Assumption 2 holds for any* $\bar{z} \in \boldsymbol{Z}$.

Proof If $\boldsymbol{G}$ is affine then, by Corollary 6.2.2 in [13], the mixed complementarity problem MiCP($\mathbf{K}, \boldsymbol{G}$) has a local error bound with the natural residual (denoted $\boldsymbol{G}_{\mathbf{K}}^{\text{nat}}$) if $\boldsymbol{Z}$ is nonempty. More in detail, there are $\hat{c}, \hat{\kappa} > 0$ so that

$$\|\boldsymbol{G}_{\mathbf{K}}^{\text{nat}}(z)\| \geq \hat{c}\, \text{dist}[z, \boldsymbol{Z}]$$

holds for all z with $\|\boldsymbol{G}_{\mathbf{K}}^{\text{nat}}(z)\| \leq \hat{\kappa}$. Since in our setting $\boldsymbol{G}_{\mathbf{K}}^{\text{nat}} = \boldsymbol{\Phi}$ the continuity of $\boldsymbol{\Phi}$ yields that there are $c, \delta > 0$ so that Assumption 2 is satisfied for any $\bar{z} \in \boldsymbol{Z}$. □

Theorem 8 covers the class of GNEP that have quadratic objective function and affine constraints for each player. This is an important class of problems with applications in microeconomics and that could also be used to approximate more complex problems, see Sect. 4.6 in [2]. The ν-th player of such GNEPs has the following minimization problem

$$\text{minimize}_{x^\nu} \ \frac{1}{2}\boldsymbol{x}^\top Q^\nu \boldsymbol{x} + (d^\nu)^\top \boldsymbol{x} \qquad \text{subject to} \quad A^\nu \boldsymbol{x} \leq c^\nu$$

with appropriate dimensions of vectors and matrices.

We now consider a more general case. For a given $\bar{z}$ in $\boldsymbol{Z}$, assume that at least one constraint is shared by two players and is active. Let us first make the discussion preceding Theorem 3 more notationally precise. To this end we define the mapping $\boldsymbol{\Phi}_{\bar{z}} : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m-\ell}$ by

$$\boldsymbol{\Phi}_{\bar{z}}(z) := \begin{pmatrix} \mathbf{L}(z) \\ \phi(-g^1_{R_1}(\boldsymbol{x}), \lambda^1_{R_1}) \\ \vdots \\ \phi(-g^N_{R_N}(\boldsymbol{x}), \lambda^N_{R_N}) \end{pmatrix},$$

where R_ν is an index set satisfying $I^\nu_<(\bar{\boldsymbol{x}}) \subseteq R_\nu \subseteq \{1, \dots, m_\nu\}$. Moreover, an index $i \in I^\nu_0(\bar{\boldsymbol{x}}) = \{1, \dots, m_\nu\} \setminus I^\nu_<(\bar{\boldsymbol{x}})$ belongs to R_ν if and only if, for any player $\mu = 1, \dots, \nu - 1$, there is no $j \in I^\mu_0(\bar{\boldsymbol{x}})$ with $g^\nu_i = g^\mu_j$. Thus, to obtain $\boldsymbol{\Phi}_{\bar{z}}$ from $\boldsymbol{\Phi}$, consider a constraint that is active at $\bar{\boldsymbol{x}}$ and that is shared by at least two players. Except for the first of these players we remove all those complementarity functions ϕ from $\boldsymbol{\Phi}$ that belong to this constraint. Obviously, $\ell = \sum_{\nu=1}^N (m_\nu - |R_\nu|)$ is the number of those complementarity functions ϕ that have been dropped from $\boldsymbol{\Phi}$. The next lemma states that under strict complementarity at $\bar{z}$ the solution sets of the systems $\mathbf{\Phi}(z) = 0$ and $\mathbf{\Phi}_{\bar{z}}(z) = 0$ are the same if z is sufficiently close to $\bar{z}$.

Lemma 1 *Let $\bar{z}$ be a solution of system (3) and suppose that Assumption 1 holds at $\bar{z}$. Then, there is $\delta_0 > 0$ so that*

$$\{z \in \mathbb{B}(\bar{z}, \delta_0) \mid \mathbf{\Phi}_{\bar{z}}(z) = 0\} \;=\; \mathbf{Z} \cap \mathbb{B}(\bar{z}, \delta_0).$$

Proof Choose any player $\nu \in \{1, \ldots, N\}$ and an index $i \in \{1, \ldots, m_\nu\} \setminus R_\nu$, i.e., the equation $\phi(-g_i^\nu(\boldsymbol{x}), \lambda_i^\nu) = 0$ does not appear in the system $\Phi_{\bar{z}}(z) = 0$. For $\delta_0 > 0$ sufficiently small we know by Assumption 1 and by the continuity of ϕ and g_i^ν that

$$\phi(-g_i^\nu(\boldsymbol{x}), \lambda_i^\nu) = \min\{-g_i^\nu(\boldsymbol{x}), \lambda_i^\nu\} = -g_i^\nu(\boldsymbol{x})$$

holds for all $z \in \mathbb{B}(\bar{z}, \delta_0)$. Since, by the definition of R_ν, there is a $\mu \in \{1, \ldots, \nu - 1\}$ and $j \in I_0^\mu(\bar{\boldsymbol{x}})$ so that $g_i^\nu = g_j^\mu$ we either have that $j \in R_\mu$ or, if not, we can repeat the reasoning above a finite number of times and eventually will end up with some smaller μ and a certain $j \in R_\mu$ so that $g_i^\nu = g_j^\mu$. Then, using Assumption 1 again we see that

$$\phi(-g_j^\mu(\boldsymbol{x}), \lambda_j^\mu) = \min\{-g_j^\nu(\boldsymbol{x}), \lambda_j^\mu\} = -g_j^\mu(\boldsymbol{x}) = -g_i^\nu(\boldsymbol{x}) \;\text{ for all } z \in \mathbb{B}(\bar{z}, \delta_0).$$

Moreover, it follows that equation $\phi(-g_j^\mu(\boldsymbol{x}), \lambda_j^\mu) = 0$ appears in the system $\Phi_{\bar{z}}(z) = 0$ and is the same as $\phi(-g_i^\nu(\boldsymbol{x}), \lambda_i^\nu) = 0$ if z is restricted to $\mathbb{B}(\bar{z}, \delta_0)$. Thus, for any $z \in \mathbb{B}(\bar{z}, \delta_0)$, we have $\mathbf{\Phi}_{\bar{z}}(z) = 0$ if and only if $\mathbf{\Phi}(z) = 0$. □

Due to Lemma 1 it is reasonable to seek for a condition under which an error bound holds for the mapping $\mathbf{\Phi}_{\bar{z}}$ around $\bar{z}$. The next lemma presents such a condition. Since the result is valid for a general differentiable mapping and not only for $\mathbf{\Phi}_{\bar{z}}$, we state it at a more general level than that needed here.

Lemma 2 *Let $H : \mathbb{R}^p \to \mathbb{R}^q$ with $p \geq q$ be given and let z^* denote a solution of $H(z) = 0$. Assume that, in some neighborhood of z^*, H is continuously differentiable and ∇H is locally Lipschitz continuous. Moreover, suppose that there is a partition (u, v) of the vector z so that $u \in \mathbb{R}^q$, $v \in \mathbb{R}^{p-q}$, and the matrix $\nabla_u H(z^*) \in \mathbb{R}^{q \times q}$ is nonsingular. Then, denoting with Z the solution set of the system $H(z) = 0$, there are $c_1, \delta_1 > 0$ so that, for all $z \in \mathbb{B}(z^*, \delta_1)$,*

$$\|H(z)\| \geq c_1 \operatorname{dist}[z, Z].$$

Proof By the classical implicit function theorem there is $\delta_1 > 0$ so that a continuously differentiable function $u(\cdot) : \mathbb{B}(v^*, \delta_1) \to \mathbb{R}^q$ exists with $u(v^*) = u^*$ and

$$H(u(v), v) = 0 \quad \text{for all } v \in \mathbb{B}(v^*, \delta_1). \tag{20}$$

Without loss of generality $\delta_1 > 0$ is assumed to be small enough so that $\nabla_u H(u(v), v)^{-1}$ exists on $\mathbb{B}(v^*, \delta_1)$. Then, since $\nabla_u H(u(v), v)$ depends continuously on v on the ball

$\mathbb{B}(v^*, \delta_1)$ the same is true for $\nabla_u H(u(v), v)^{-1}$. Thus, there is $C_0 > 0$ so that

$$\|\nabla_u H(u(v), v)^{-1}\| \leq C_0 \quad \text{for all } v \in \mathbb{B}(v^*, \delta_1). \tag{21}$$

Due to the local Lipschitz continuity of ∇H, $u(v^*) = u^*$, and the continuity of $u(\cdot)$ in $\mathbb{B}(v^*, \delta_1)$ there is $L_0 > 0$ so that

$$\|\nabla_u H(u(v) + t(u - u(v)), v) - \nabla_u H(u(v), v)\| \leq t L_0 \|u - u(v)\| \tag{22}$$

for all $(u, v) \in \mathbb{B}(z^*, \delta_1)$ and all $t \in [0, 1]$.

Now, for any $v \in \mathbb{B}(v^*, \delta_1)$, a Taylor expansion of $H(\cdot, v)$ at $u(v)$ yields

$$\begin{aligned} H(u, v) = {} & H(u(v), v) + \nabla H_u(u(v), v)^\top (u - u(v)) \\ & + \int_0^1 (\nabla_u H(u(v) + t(u - u(v)), v) - \nabla H_u(u(v), v))^\top (u - u(v)) dt. \end{aligned}$$

for all $u \in \mathbb{R}^q$. Therefore, using (20), (21), and (22) we obtain

$$\|u(v) - u\| \leq C_0 \|H(u, v)\| + \frac{1}{2} L_0 C_0 \|u - u(v)\|^2 \tag{23}$$

for all $(u, v) \in \mathbb{B}(z^*, \delta_1)$. For $\delta_1 > 0$ sufficiently small it follows by $u(v^*) = u^*$ and the continuity of $u(\cdot)$ that $\|u - u(v)\| \leq 2/L_0$ and, with (23), that

$$\text{dist}[(u, v), Z] \leq \|(u, v) - (u(v), v)\| = \|u - u(v)\| \leq 2C_0 \|H(u, v)\|$$

for all $(u, v) \in \mathbb{B}(z^*, \delta_1)$. Setting $c_1 := 2C_0$ completes the proof. □

We are now able to give a sufficient condition for Assumption 2 to hold.

Theorem 9 *Suppose that $\bar{z}$ is a solution of system (3) that satisfies Assumption 1 and $J\boldsymbol{\Phi}_{\bar{z}}(\bar{z})$ has full row rank. Then, Assumption 2 holds.*

Proof By Assumption 1 there is some neighborhood where the mapping $\boldsymbol{\Phi}_{\bar{z}}$ is differentiable with locally Lipschitz continuous derivative. Then, we apply Lemma 2 to $H := \boldsymbol{\Phi}_{\bar{z}}$ and $z^* := \bar{z}$. Taking into account the full rank assumption and Lemma 1 we get

$$\|\boldsymbol{\Phi}(z)\| = \|\boldsymbol{\Phi}_{\bar{z}}(z)\| \geq c_1 \,\text{dist}[z, \boldsymbol{Z}]$$

for all z sufficiently close to $\bar{z}$. □

Note that the map $\boldsymbol{\Phi}_{\bar{z}}$ is used only for analysis purposes and it is never necessary to actually calculate it (which would be obviously impossible).

Remark 4 It is possible to further weaken the assumptions in Theorem 9. To this end let $\bar{z}$ again be a solution of system (3) and $\tilde{\mathbf{\Phi}} : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m-l}$ be any continuously differentiable mapping so that $J\tilde{\mathbf{\Phi}}(\bar{z})$ has full row rank and

$$\{z \in \mathbb{B}(\bar{z}, \delta_0) \mid \tilde{\mathbf{\Phi}}(z) = 0\} \;=\; \mathbf{Z} \cap \mathbb{B}(\bar{z}, \delta_0).$$

Then, obviously, Assumption 2 holds.

Consider now a quasi-regular solution; then, we are able to get quadratic convergence for the classical semismooth method, see Sect. 3.1. However, while quasi-regularity ensures the error bound condition (see [13]), strict complementarity might not be satisfied, and thus the convergence properties of the Levenberg–Marquardt method, as studied here, may be in jeopardy. However, quasi-regularity ensures local uniqueness and this makes the situation much simpler: in fact [12, Theorem 3] provides the following result.

Theorem 10 *Let $\bar{z}$ be a quasi-regular solution of $\mathbf{\Phi}(z) = 0$. Then, there is a neighborhood $\mathcal{N}$ of $\bar{z}$ so that, for any starting point $z^0 \in \mathcal{N}$, the sequence $\{z^k\}$ generated by the Newton Method III converges to $\bar{z}$ Q-quadratically.*

5 Analysis of a model for internet switching

In this section we illustrate the theory developed so far by considering a model proposed by Kesselman et al. [20]. To better exemplify our results we also propose several extensions of the basic model.

5.1 The basic model and quasi-regularity

The model proposed in [20] analyzes the problem of internet switching where traffic is generated by selfish users. The model concerns the behavior of users sharing a first-in-first-out buffer with bounded capacity. The utility of each user depends on its transmission rate and the congestion level. More precisely we assume that there are N users and the buffer capacity is B. Each user ν controls the amount of his "packets" in the buffer; denote by $x^\nu \in [0, \infty)$ this number (for simplicity we assume that x^ν can take any nonnegative real value). It is assumed that the buffer is managed with drop-tail policy, which means that if the buffer is full, further packets are lost and should be resent. The utility of user ν is given by

$$\theta_\nu(\boldsymbol{x}) \;:=\; \begin{cases} \dfrac{x^\nu}{\sum_{\nu=1}^N x^\nu}\left(1 - \dfrac{\sum_{\nu=1}^N x^\nu}{B}\right) & \text{if } \sum_{\nu=1}^N x^\nu \;>\; 0 \\ 0 & \text{if } \sum_{\nu=1}^N x^\nu \;=\; 0 \end{cases}$$

which is to be maximized. The term $(x^\nu)/(\sum_{\nu=1}^N x^\nu)$ represents the *transmission rate* of user ν; the utility of the user increases with the increase of his transmission rate. The term $(\sum_{\nu=1}^N x^\nu)/B$ is the *congestion level* of the buffer and therefore the term $1-(\sum_{\nu=1}^N x^\nu)/B$ in the utility of the user weights the decrease in the utility of the user as the congestion level increases. Note that if the buffer is full ("congestion collapse") the utility of each user is zero. Taking into account the drop-tail policy we can see that the ν-th user's problem is

$$\begin{aligned} &\text{minimize}_{x^\nu} \ -\theta_\nu(\boldsymbol{x}) \\ &\text{subject to} \quad x^\nu \geq 0, \quad \textstyle\sum_{\nu=1}^N x^\nu \leq B. \end{aligned} \tag{24}$$

This is the model dealt with in [20][2]. We also consider the following variant:

$$\begin{aligned} &\text{minimize}_{x^\nu} \ -\theta_\nu(\boldsymbol{x}) \\ &\text{subject to} \quad x^\nu \geq l_\nu, \quad \textstyle\sum_{\nu=1}^N x^\nu \leq B, \end{aligned} \tag{25}$$

where $l_\nu \geq 0$ for each ν. It is clear that if $l_\nu = 0$ for each ν, (25) coincides with (24). The problem (25) however also models the case in which users do not enter the game if they don't have a minimal amount of data to send ($l_\nu > 0$). In the sequel we will always assume that $\sum_{\nu=1}^N l_\nu < B$, i.e., we exclude the uninteresting cases in which the feasible set is either empty or reduces to a singleton.

It is shown in [20] that the model (24) has a unique solution $\bar{\boldsymbol{x}}$ which is given by $\bar{x}^\nu = B(N-1)/N^2$. We therefore see that the solution of the generalized Nash game (24) is unconstrained. Taking into account that each player controls only one variable and the objective function has the same structure for each player, it can be checked that $\gamma^\nu = \emptyset$ for every ν and there is only one matrix H in the family $\{H_R(\boldsymbol{\gamma}) : \boldsymbol{\gamma} \in \boldsymbol{\Gamma}\}$, this matrix is given by

$$H = J_{\boldsymbol{x}}\mathbf{L}(\bar{\boldsymbol{x}}) = -P/\mathcal{X}^3 \tag{26}$$

with

$$P := \begin{pmatrix} -2\mathcal{X}^{-1} & \bar{x}^1 - \mathcal{X}^{-1} & \cdots & \bar{x}^1 - \mathcal{X}^{-1} \\ \bar{x}^2 - \mathcal{X}^{-2} & -2\mathcal{X}^{-2} & \cdots & \bar{x}^2 - \mathcal{X}^{-2} \\ \vdots & & \ddots & \vdots \\ \bar{x}^N - \mathcal{X}^{-N} & \bar{x}^N - \mathcal{X}^{-N} & \cdots & -2\mathcal{X}^{-N} \end{pmatrix},$$

$$\mathcal{X} := \textstyle\sum_{\nu=1}^N \bar{x}^\nu, \quad \text{and} \quad \mathcal{X}^{-\nu} := \sum_{\substack{j=1 \\ j\neq\nu}}^N \bar{x}^j,$$

where we set for notational convenience $x_1^\nu := x^\nu$ (i.e., x^ν is the unique variable controlled by the ν-th player).

In order to verify that quasi-regularity is satisfied at $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ with $\bar{\boldsymbol{\lambda}} = 0 \in \mathbb{R}^{N+1}$, the nonsingularity of the matrix P has to be shown. This is done by means of the following known formula for the determinant of the sum of a matrix M and of a diagonal matrix D:

$$\det(D + M) = \sum_\alpha \det D_{\alpha\alpha} \det M_{\bar{\alpha}\bar{\alpha}},$$

[2] The constraint $\sum_{\nu=1}^N x^\nu \leq B$ is not considered explicitly in [20], but rather dealt with implicitly. For the sake of clarity we have included this constraint explicitly.

where the summation ranges over all subsets of indices α with complement $\bar{\alpha}$. Included in the summation are the two extreme cases corresponding to α being the empty set and the full set; the convention for these cases is that the determinant of an empty matrix is set equal to 1. We can write

$$P = \begin{pmatrix} \bar{x}^1 - \mathcal{X}^{-1} & \bar{x}^1 - \mathcal{X}^{-1} & \cdots & \bar{x}^1 - \mathcal{X}^{-1} \\ \bar{x}^2 - \mathcal{X}^{-2} & \bar{x}^2 - \mathcal{X}^{-2} & \cdots & \bar{x}^2 - \mathcal{X}^{-2} \\ \vdots & \vdots & & \vdots \\ \bar{x}^N - \mathcal{X}^{-N} & \bar{x}^N - \mathcal{X}^{-N} & \cdots & \bar{x}^N - \mathcal{X}^{-N} \end{pmatrix} + \begin{pmatrix} -\mathcal{X} & & & \\ & -\mathcal{X} & & \\ & & \ddots & \\ & & & -\mathcal{X} \end{pmatrix}. \tag{27}$$

Therefore, applying the determinantal formula above, we get

$$\begin{aligned} \det P &= (-\mathcal{X})^N + \sum_{\nu=1}^{N} (-\mathcal{X})^{N-1}(\bar{x}^\nu - \mathcal{X}^{-\nu}) = (-\mathcal{X})^{N-1}\left[-\mathcal{X} + \sum_{\nu=1}^{N}(\bar{x}^\nu - \mathcal{X}^{-\nu})\right] \\ &= (-\mathcal{X})^{N-1}\sum_{\nu=1}^{N}\mathcal{X}^{-\nu} = -(N-1)(-\mathcal{X})^{N-1}\mathcal{X} = (N-1)(-\mathcal{X})^N. \end{aligned}$$

We then conclude that quasi-regularity and (according to Theorem 6) strong regularity are satisfied at the solution $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$.

Remark 5 It is important to note that given any nonempty principal submatrix of the matrix P, say $P_{\beta\beta}$, of order $q := |\beta|$ we can still calculate the determinant of $P_{\beta\beta}$ along the same lines, thus getting:

$$\det P_{\beta\beta} = (-\mathcal{X})^{q-1}\left[(q+1)(-\mathcal{X}) + 2\sum_{\nu\in\beta}\bar{x}^\nu\right].$$

We now consider the more interesting case of the constrained problem (25).

Theorem 11 *The GNEP (25) has at least one solution. For every solution $\bar{\boldsymbol{x}}$ there is $\bar{\boldsymbol{\lambda}}$ so that $\bar{\boldsymbol{z}} := (\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ satisfies system (3) associated with the GNEP and quasi-regularity holds at $\bar{\boldsymbol{z}}$.*

Proof The only case of interest is when at least one l_ν is positive since otherwise the model reduces to the previous one. Consider the set

$$\boldsymbol{X} = \{\boldsymbol{x} \in \mathbb{R}^N : x^\nu \ge l_\nu, \nu = 1, \ldots, N, \sum_{\nu=1}^{N} x^\nu \le B\}.$$

Given the other players' variables $\boldsymbol{x}^{-\nu}$, the ν-th player's feasible set is given by $X_\nu(\boldsymbol{x}^{-\nu}) = \{x^\nu \in \mathbb{R} : (x^\nu, \boldsymbol{x}^{-\nu}) \in \boldsymbol{X}\}$. Since we assumed that at least one l_ν is positive, the objective functions of all players are continuous on $\boldsymbol{X}$, which is obviously convex and compact. The existence of a solution then follows from [2, Theorem 4.4]. Since the constraints in (25) are linear, the KKT conditions hold at a solution $\bar{\boldsymbol{x}}$. Let

$\bar{z} := (\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ be a KKT point for the GNEP (25). The constraint $\sum_{\nu=1}^{N} x^\nu \leq B$ can not be active at $\bar{\boldsymbol{x}}$ since otherwise there is at least one player for which $x^\nu > l_\nu$ (because we assumed that $\sum_{\nu=1}^{N} l_\nu < B$) that can therefore improve his objective function by decreasing the value of his variable. Then, each player can have at most one active constraint. This implies that the linear independence assumption holds for each player, so that there is a unique multiplier $\bar{\boldsymbol{\lambda}}$ associated with $\bar{\boldsymbol{x}}$.

Since each player has only one variable and at most one active constraint, we simplify the notation introduced in Sect. 4.1 and write

$$\begin{aligned} I_0 &:= \{ i \in \{1, \dots N\} : \bar{x}^i = l_i \}, & I_+ &:= \{ i \in I_0 : \bar{\lambda}^i > 0 \}, \\ I_{00} &:= \{ i \in I_0 : \bar{\lambda}^i = 0 \}, & I_< &:= \{ i \in \{1, \dots N\} : \bar{x}^i > l_i \}, \end{aligned}$$

where $\bar{\lambda}^i$ is the multiplier associated with the constraint $x^i \geq l_i$. If $I_0 = \emptyset$ then quasi-regularity at $\bar{z}$ can be shown as in the case of lower bounds all equal to zero. Assume that $I_0 \neq \emptyset$. We further assume, without loss of generality, that the first $|I_0|$ players have the lower bound constraint active and that of these the first $|I_+|$ are strongly active while the remaining $|I_{00}|$ have zero multiplier. Quasi-regularity amounts to checking that the matrices $H_R(\boldsymbol{\gamma})$ are nonsingular. But these matrices have the following structure:[3]

$$\begin{array}{c} \\ \alpha \left\{ \right. \end{array} \begin{pmatrix} -P/\mathcal{X}^3 & \overbrace{\begin{matrix} -\mathbf{I}_t \\ \mathbf{0}_{N-t,t} \end{matrix}}^{\alpha} \\ \mathbf{I}_t \;\; \mathbf{0}_{t,N-t} & \mathbf{0}_t \end{pmatrix},$$

where we set $\alpha := I_+ \cup \boldsymbol{\gamma}$ with $\boldsymbol{\gamma} \subseteq I_{00}$ and $t := |\alpha|$. We now consider three cases. If $t = N$ then the two block matrices $\mathbf{0}_{t,N-t}$ and $\mathbf{0}_{N-t,t}$ vanish so that $H_R(\boldsymbol{\gamma})$ is obviously nonsingular because the columns of $H_R(\boldsymbol{\gamma})$ are linearly independent. If $0 < t < N$ we can calculate the determinant of $H(\boldsymbol{\gamma})$ by applying Laplace's rule iteratively to the t right columns and the t bottom rows. Setting $\beta := I_< \cup (I_{00} \setminus \boldsymbol{\gamma})$, $q := |\beta|$, taking into account Remark 5, and $t < N$ (which implies $|\beta| \geq 1$) we then get

$$\det H_R(\boldsymbol{\gamma}) = (-\mathcal{X})^{q-1} \left[(q+1)(-\mathcal{X}) + 2 \sum_{\nu \in \beta} \bar{x}^\nu \right] \neq 0.$$

The $\neq$ in the formula above can be derived in the following way. Because $\bar{\boldsymbol{x}}$ is a Nash equilibrium all the components of $\bar{\boldsymbol{x}}$ are positive. In fact, if the player ν had $\bar{x}^\nu = 0$ his objective value would be 0. But since we already know that $\sum_{\nu=1}^{N} l_\nu < B$, the ν-th player could increase $\bar{x}^\nu$ and improve his objective function. Thus, $\bar{\boldsymbol{x}} > 0$ and $\mathcal{X} > 0$ follows. We now consider the term $(q+1)(-\mathcal{X}) + 2 \sum_{\nu \in \beta} \bar{x}^\nu$. Since (a) $q \geq 1$, (b) β is a proper subset of all the players since $0 < |I_+ \cup \boldsymbol{\gamma}|$, and (c) $\mathcal{X} > 0$, we see that $(q+1)(-\mathcal{X}) + 2 \sum_{\nu \in \beta} x^\nu < 0$. This fully justifies that $\det H_R(\boldsymbol{\gamma}) \neq 0$. In the

[3] We assume that if we consider a subset $\boldsymbol{\gamma}$ of players/constraints in I_{00} the players/constraints in $\boldsymbol{\gamma}$ are placed immediately after the strongly active ones.

third case, if $I_+ \cup \boldsymbol{\gamma} = \emptyset$, it follows that $H_R(\boldsymbol{\gamma}) = -P/\mathcal{X}^3$ which is nonsingular as it has been already shown in this subsection. □

Remark 6 We note that the last part of the proof shows clearly that if $I_{00} \neq \emptyset$, i.e., if there are degenerate constraints, strong regularity can not be satisfied, In fact, in this case we can take two different γ in Γ, say γ' and γ'', such that $|\gamma'| = |\gamma''| + 1$ and $q' = q'' + 1$. But then it is clear from the last part of the proof that $(\det H_R(\gamma'))(\det H_R(\gamma'')) = -1$, so that by Theorem 6 strong regularity can not hold.

5.2 Avoiding saturation and the common multipliers case

Consider again the problem of internet switching and generalize it in the following way

$$\begin{aligned} &\text{minimize}_{x^\nu} \quad -\theta_\nu(\boldsymbol{x}) \\ &\text{subject to} \quad x^\nu \geq l_\nu, \quad \textstyle\sum_{\nu=1}^N x^\nu \leq \bar{B}. \end{aligned} \tag{28}$$

Note that the constant B in the definition of θ_ν is not replaced by $\bar{B}$. Therefore, problem (28) generalizes problem (25) by permitting the possibility to introduce a security margin $\bar{B} < B$ in order to avoid approaching the congestion collapse. This a common requirement in modelling this kind of situations. We will assume that $\sum_{\nu=1}^N l_\nu < \bar{B}$, i.e., we exclude the uninteresting cases in which the feasible set is either empty or reduces to a singleton. We want to show that the matrices (18) are nonsingular at a solution $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\sigma}}, \bar{\mu})$ of system (8), so that VI-quasi-regularity is satisfied at every such solution. As just discussed, the only case of interest is when the shared constraint is active. With the same notation and conventions introduced in the previous subsection the matrices (18) in this case have the form

$$\begin{array}{c} \\ \alpha\left\{\right. \\ \beta\left\{\right. \\ \\ \end{array}\begin{pmatrix} -P/\mathcal{X}^3 & \overbrace{-\mathbf{I}_N}^{\alpha\cup\beta} & \mathbf{e} \\ \mathbf{I}_t \;\; \mathbf{0}_{t,N-t} & \mathbf{0}_{t,N} & \mathbf{0}_{t,1} \\ \mathbf{0}_{q,N} & \mathbf{0}_{q,t} \;\; \mathbf{I}_q & \mathbf{0}_{q,1} \\ -a\,\mathbf{e}^T & \mathbf{0}_{1,N} & 1-a \end{pmatrix},$$

where $\mathbf{e} := (1, \ldots, 1)^T$ and t and q are defined as in the proof of Theorem 11. Moreover, $a = 1$ if the shared constraint has a positive multiplier and 0 otherwise. This matrix is nonsingular if and only if the $(q+1, q+1)$-matrix

$$M := \begin{pmatrix} 2\mathcal{X}^{-(t+1)}/\mathcal{X}^3 & \cdots & -(\bar{x}^{t+1} - \mathcal{X}^{-(t+1)})/\mathcal{X}^3 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ -(\bar{x}^N - \mathcal{X}^N)/\mathcal{X}^3 & \cdots & 2\mathcal{X}^{-N}/\mathcal{X}^3 & 1 \\ -a & \cdots & -a & 1-a \end{pmatrix} \tag{29}$$

is nonsingular. The matrix in the upper left corner of M is just a submatrix of $-P/\mathcal{X}^3$ (see (26)), namely $(-P/\mathcal{X}^3)_{\beta,\beta}$. If $a = 0$ the nonsingularity of (29) is readily seen to be equivalent to the nonsingularity of the matrix $(-P/\mathcal{X}^3)_{\beta,\beta}$. The nonsingularity of the latter matrix has already been shown in the last part of the proof of Theorem 11. Therefore (29) is nonsingular. Consider then the case $a = 1$. The matrix (29) is nonsingular if $My = 0$ implies $y = 0$. Suppose then that $y \in \mathbb{R}^{q+1}$ is such that $My = 0$. The last row of M shows that $\sum_{i=1}^{q} y_i = 0$. Recalling (27), the first q rows of the equation $My = 0$ can be written as:

$$\frac{-1}{\mathcal{X}^3}\begin{pmatrix} \bar{x}^{t+1} - \mathcal{X}^{-(t+1)} \\ \vdots \\ \bar{x}^N - \mathcal{X}^{-N} \end{pmatrix} \sum_{i=1}^{q} y_i + \frac{1}{\mathcal{X}^2}\begin{pmatrix} y_1 \\ \vdots \\ y_q \end{pmatrix} = -\begin{pmatrix} y_{q+1} \\ \vdots \\ y_{q+1} \end{pmatrix}$$

and, by using $\sum_{i=1}^{q} y_i = 0$, we get

$$\frac{1}{\mathcal{X}^2}\begin{pmatrix} y_1 \\ \vdots \\ y_q \end{pmatrix} = -\begin{pmatrix} y_{q+1} \\ \vdots \\ y_{q+1} \end{pmatrix}.$$

This yields $y_1 = y_2 = \cdots = y_q$. Thus, again by $\sum_{i=1}^{q} y_i = 0$, we get $y_1 = y_2 = \cdots = y_q = 0$, which, obviously, also implies $y_{q+1} = 0$ so that we conclude that $y = 0$ and hence M is nonsingular. Therefore, VI-quasi-regularity holds at every solution of system (8).

5.3 Quality of Service and the hard case

In this last subsection we consider one further modification of the problem we are dealing with that will allow us to illustrate also the hard case of Sect. 3.3. We assume that one user, say the first one, has a better treatment and is allowed to freely use the buffer up to $M < \bar{B}$, independent of what the remaining users do. Thus, while the optimization problems of users $2, 3, \ldots, N$ remain (28), the first user's problem becomes

$$\text{minimize}_{x^1} \ -\theta_1(\boldsymbol{x}) \quad \text{subject to} \quad 0 \le x^1 \le M. \tag{30}$$

This may be viewed as a simple example of the Quality of Service approach, where the users are divided in classes with different characteristics. Mathematically, the resulting GNEP falls in what we called the “hard case” since the group of players $\{2, \ldots, N\}$ shares a constraint that depends also on the variable of the first player who, however is not sharing this constraint. Our aim is to show that at any solution where strict complementarity holds, the sufficient full rank condition of Theorem 9 is verified. Actually we will not give a detailed proof of this fact, because an analysis of all possible situations would be simple but long. We prefer to show on a specific case the result and leave to the reader to examine, in a similar way, the remaining cases.

Assume then that $\bar{\boldsymbol{x}}$ is a solution of the GNEP just described at which strict complementarity holds. Obviously the interesting case to analyze is when the constraint

$\sum_{\nu=1}^{N} x^\nu \leq \bar{B}$ is active. Furthermore, we consider the case in which $x^1 = M$ and none of the variables $x^2, \ldots, x^N$ is at its lower bound. Under these conditions the "reduced" $\Phi_{\bar{z}}$ in Sect. 4.3 becomes

$$\boldsymbol{\Phi}_{\bar{z}}(z) = \begin{bmatrix} -\nabla_{\boldsymbol{x}}\theta(\boldsymbol{x}) - \sum_{\nu=1}^{N} e^\nu \lambda_1^\nu + \sum_{\nu=1}^{N} e^\nu \lambda_2^\nu \\ \lambda_1^1 \\ \vdots \\ \lambda_1^N \\ M - x^1 \\ \bar{B} - \sum_{\nu=1}^{N} x^\nu \end{bmatrix},$$

where the multipliers λ_1^ν correspond to the constraints $x^\nu \geq l_\nu$ and the multipliers λ_2^ν correspond to the second constraint of player ν. Moreover, there is exactly one multiplier vector $\bar{\boldsymbol{\lambda}}$ so that $\bar{z} := (\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ solves $\boldsymbol{\Phi}(z) = 0$. Differentiating $\boldsymbol{\Phi}_{\bar{z}}$ yields

$$J_z \boldsymbol{\Phi}_{\bar{z}}(\bar{z}) = \begin{pmatrix} -P/\mathcal{X}^3 & -\mathbf{I}_N & \mathbf{I}_N \\ \mathbf{0}_N & \mathbf{I}_N & \mathbf{0}_N \\ -1\ \mathbf{0}_{1,N-1} & \mathbf{0}_N & \mathbf{0}_N \\ -\mathbf{e}^T & \mathbf{0}_{1,N} & \mathbf{0}_{1,N} \end{pmatrix} \in \mathbb{R}^{(2N+2)\times 3N}$$

with $\mathbf{e} := (1, \ldots, 1)^T$. We can easily verify that the rank of this rectangular matrix is $2N+2$ by showing that the rows of $J_z \boldsymbol{\Phi}_{\bar{z}}(\bar{z})$ are linearly independent. Proceeding in a similar fashion we can analyze other cases (the constraint $x^1 \leq M$ is not active, some of the x^ν are at their lower bounds, etc.). This tedious but not difficult examinations lead us to conclude that the following result holds (we are also using the fact that the feasible set is compact, so a solution exists).

Proposition 5 *In the setting of this subsection, the GNEP we are considering always has a solution and any solution at which strict complementarity holds, satisfies the sufficient condition of Theorem 9, which implies that Assumption 2 is satisfied.*

6 Conclusions

In this paper we have analyzed in detail the applicability of semismooth Newton methods to the GNEP. To the best of our knowledge, the only other investigation of Newton's method for the GNEP has been carried out by Pang [26] who, under suitable convexity and regularity assumptions, begins with the same reduction to a structured mixed complementarity problem we also use. However, he then applies the Josephy–Newton method to this problem (see [13,19,27]), giving rise to an approach that requires the solution of a linearized mixed complementarity problem at every iteration.

We believe that the next big issue one has to study is the development of a globally convergent algorithm for the solution of a GNEP. To date, little is known in this field. It is interesting to note that the two provably globally convergent algorithms, see

[21,25,31] are applicable to subclasses of what we called the jointly convex GNEPs. Since we showed that these GNEPs can be solved by solving an appropriate VI, a problem for which a rich theory exists, we believe that a promising research direction will be to study the application of the theory of VIs to this specific VI reformulation of jointly convex GNEPs.

References

1. Arrow, K.J., Debreu, G.: Existence of an equilibrium for a competitive economy. Econometrica **22**, 265–290 (1954)
2. Başar, T., Olsder, G.J.: Dynamic Noncooperative Game Theory, 2nd edn. Academic, London/New York (1989). (Reprinted in SIAM Series "Classics in Applied Mathematics, 1999)
3. Bensoussan, A.: Points de Nash dans le cas de fonctionnelles quadratiques et jeux differentiels lineaires a N personnes. SIAM J. Control **12**, 460–499 (1974)
4. Berridge, S., Krawczyk, J.B.: Relaxation algorithms in finding Nash equilibria. Economic working papers archives (1997), http://econwpa.wustl.edu/eprints/comp/papers/9707/9707002.abs
5. Cardell, J.B., Cullen Hitt, C., Hogan, W.W.: Market power and strategic interaction in electricity networks. Resour. Energy Econ. **19**, 109–137 (1997)
6. Chen, Y., Hobbs, B.F., Leyffer, S., Munson, T.S.: Leader-Follower Equilibria for Electric Power and NO_x Allowances Markets. Preprint ANL/MCS-P1191-0804 (2004), Mathematics and Computer Science Division, Argonne National Laboratory, Argonne (2004)
7. Christensen, P.W.: A semi-smooth Newton method for elasto-plastic contact problems. Int. J. Solids Struct. **39**, 2323–2341 (2002)
8. Contreras, J., Klusch, M.K., Krawczyk, J.B.: Numerical solution to Nash-Cournot equilibria in coupled constraints electricity markets. IEEE T. Power Syst. **19**, 195–206 (2004)
9. De Luca, T., Facchinei, F., Kanzow, C.: A theoretical and numerical comparison of some semismooth algorithms for complementarity problems. Comput. Optim. Appl. **16**, 173–205 (2000)
10. Facchinei, F., Fischer, A., Kanzow, C.: On the accurate identification of active constraints. SIAM J. Optim. **8**, 14–32 (1998)
11. Facchinei, F., Fischer, A., Piccialli, V.: On generalized Nash games and variational inequalities. Oper. Res. Lett. **35**, 159–164 (2007)
12. Facchinei, F., Kanzow, C.: A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems. Math. Program. **76**, 195–206 (1997)
13. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer, New York (2003)
14. Facchinei, F., Pang, J.-S.: Exact penalty functions for generalized Nash problems. In: Di Pillo, G., Roma, M. (eds.) Large-scale nonlinear optimization, pp 115–126. Springer, Heidelberg (2006)
15. Fan, J., Yuan, Y.: On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption. Computing **74**, 23–39 (2005)
16. Fischer, A.: Local behavior of an iterative framework for generalized equations with nonisolated solutions. Math. Program. **94**, 91–124 (2002)
17. Fukushima, M., Pang, J.-S.: Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games. Comput. Manage. Sci. **2**, 21–56 (2005)
18. Harker, P.T.: Generalized Nash games and quasi-variational inequalities. Eur. J. Oper. Res. **54**, 81–94 (1991)
19. Josephy, N.H.: Newton's method for generalized equations. Technical Summary Report 1965, Mathematics Research Center, University of Wisconsin, Madison (1979)
20. Kesselman, A., Leonardi, S., Bonifaci, V.: Game-Theoretic Analysis of Internet Switching with Selfish Users. In: Proceedings of the First International Workshop on Internet and Network Economics, WINE 2005. Lectures Notes in Computer Science, vol. 3828, pp 236–245
21. Krawczyk, J.B., Uryasev, S.: Relaxation algorithms to find Nash equilibria with economic applications. Environ. Model. Assess. **5**, 63–73 (2000)
22. McKenzie, L.W.: On the existence of a general equilibrium for a competitive market. Econometrica **27**, 54–71 (1959)

23. Munson, T.S., Facchinei, F., Ferris, M.C., Fischer, A., Kanzow, C.: The semismooth algorithm for large scale complementarity problems. INFORMS J. Comput. **13**, 294–311 (2001)
24. Nikaido, H., Isoda, K.: Note on noncooperative convex games. Pac. J. Math. **5**, 807–815 (1955)
25. Outrata, J., Kočvara, M., Zowe, J.: Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Kluwer, Dordrecht (1998)
26. Pang, J.-S.: Computing Generalized Nash Equilibria. Math. Program (2002) (submitted)
27. Robinson, S.M.: Generalized equations. In: Bachem, A., Grötschel, M., Korte, B. (eds.) Mathematical Programming: The State of the Art, pp 346–367. Springer, Berlin (1983)
28. Robinson, S.M.: Shadow prices for measures of effectiveness. I.Linear Model Oper. Res. **41**, 518–535 (1993)
29. Robinson, S.M.: Shadow prices for measures of effectiveness. II. General Model Oper. Res. **41**, 536–548 (1993)
30. Rosen, J.B.: Existence and uniqueness of equilibrium points for concave n-person games. Econometrica **33**, 520–534 (1965)
31. Uryasev, S., Rubinstein, R.Y.: On relaxation algorithms in computation of noncooperative equilibria. IEEE T Automat. Control **3**, 1263–1267 (1994)
32. Yamashita, N., Fukushima, M.: On the rate of convergence of the Levenberg-Marquardt method. Computing **15**(Suppl.), 239–249 (2001)