Als Manuskript gedruckt

Technische Universität Dresden
Institut für Numerische Mathematik

A New Error Bound Result for Generalized Nash Equilibrium Problems and its Algorithmic Application

A. Dreves, F. Facchinei, A. Fischer, and M. Herrich

MATH–NM–1–2013

January 2013

A New Error Bound Result for Generalized Nash Equilibrium Problems and its Algorithmic Application [1] [2]

Axel Dreves
Institute of Mathematics
Department of Aerospace Engineering
Universität der Bundeswehr München
Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany
Axel.Dreves@unibw.de

Francisco Facchinei

Department of Computer and
System Sciences Antonio Ruberti
Universita di Roma La Sapienza
Via Ariosto 25, 00185 Rome, Italy
facchinei@dis.uniroma1.it

Andreas Fischer and **Markus Herrich**

Institute of Numerical Mathematics
Department of Mathematics
Technische Universität Dresden
01062 Dresden, Germany
{Andreas.Fischer,Markus.Herrich}@tu-dresden.de

January 18, 2013

Abstract. We present a new algorithm for the solution of Generalized Nash Equilibrium Problems. This hybrid method combines the robustness of a potential reduction algorithm and the local quadratic convergence rate of the LP-Newton method. We base our local convergence theory on an error bound and provide a new sufficient condition for it to hold that is weaker than known ones. In particular, this condition implies neither local uniqueness of a solution nor strict complementarity. We also report promising numerical results.

Keywords. Generalized Nash Equilibrium Problem, potential reduction algorithm, LP-Newton method, global convergence, local quadratic convergence, local error bound condition

[1] This work is supported in a part by the German Research Foundation (DFG) in the Collaborative Research Center 912 "Highly Adaptive Energy-Efficient Computing".

[2] The paper was partially completed while the last three authors were visiting the Institute for Mathematical Sciences, National University of Singapore in 2012.

1 Introduction

We consider the Generalized Nash Equilibrium Problem (GNEP) where, given N players $\nu = 1, \ldots, N$, every player aims at minimizing

$$\min_{x^\nu} \theta_\nu(x^\nu, x^{-\nu}) \quad \text{s.t.} \quad g^\nu(x^\nu, x^{-\nu}) \leq 0 \tag{1}$$

by controlling his own variables $x^\nu \in \mathbb{R}^{n_\nu}$. Here, $\theta_\nu : \mathbb{R}^n \to \mathbb{R}$ is the objective function of player ν, $g^\nu : \mathbb{R}^n \to \mathbb{R}^{m_\nu}$ represents his constraints, and $n := n_1 + \cdots + n_N$ denotes the total number of variables. As usual in the context of GNEPs, we often write $(x^\nu, x^{-\nu})$ instead of $x \in \mathbb{R}^n$, where the vector $x^{-\nu}$ is defined by $x^{-\nu} := (x^\mu)_{\mu=1,\mu\neq\nu}^N$. The total number of inequality constraints will be denoted by $m := m_1 + \ldots + m_N$. We make the blanket assumption that the functions θ_ν and g^ν are twice continuously differentiable with locally Lipschitz continuous second-order derivatives for all $\nu = 1, \ldots, N$.

Our interest is in the computation of a Nash equilibrium (which we also simply call a solution of a GNEP). We recall that a point $x^* \in \mathbb{R}^n$ is a Nash equilibrium if, for $\nu = 1, \ldots, N$, it holds that $g^\nu(x^{*,\nu}, x^{*,-\nu}) \leq 0$ and

$$\theta_\nu(x^{*,\nu}, x^{*,-\nu}) \leq \theta_\nu(x^\nu, x^{*,-\nu}) \quad \text{for all } x^\nu \text{ satisfying } g^\nu(x^\nu, x^{*,-\nu}) \leq 0,$$

i.e., if $x^{*,\nu}$ is an optimal solution of the ν-th player's problem, given $x^{*,-\nu}$. We refer the reader to [11, 15] and references therein for a survey on GNEPs and also to [1, 3, 4, 5, 6, 12, 14, 16, 17, 18, 20, 23, 24, 25, 26] for more recent (methodological) results. GNEPs, in their full generality, are extremely hard problems and few sound algorithms have been proposed so far. In particular, as far as we are aware of, there exists no algorithm for *general* GNEPs that is both globally convergent to a solution and locally superlinearly or quadratically convergent. Our main aim in this paper is to fill this gap. To this end we make two main contributions:

- a new local error bound result for the Karush-Kuhn-Tucker (KKT) system of a GNEP;
- a hybrid algorithm combining the two recently proposed methods from [5] and [8] in a nontrivial way and resulting in a globally convergent and locally quadratically convergent algorithm.

The error bound result, beside being of independent interest, is a key to the analysis of the hybrid algorithm. Note that because of the peculiarities of the KKT system of a GNEP, standard error bound results, see [7, 15, 22] for example, are not easily applicable and new techniques are needed for the analysis of the KKT system. The hybrid algorithm follows the long tradition of combining two algorithms with different properties; however in our case we have the additional difficulty that usually solutions of GNEPs are not locally unique, see [2, 10]. In the remaining part of this introduction we give a few more details on our approach.

In order to solve the GNEP, we solve its KKT system, i.e., the system obtained by concatenating the KKT systems of the N players' optimization problems. It is well known that under some standard, mild constraint qualification, a solution x^* of the GNEP must satisfy this system. The reverse is true if, for example, one assumes that $\theta_\nu(\cdot, x^{-\nu})$ and $g_i^\nu(\cdot, x^{-\nu})$ are convex functions for every ν, i, and $x^{-\nu}$. With the Lagrange function of the ν-th player

$$L^\nu(x, \lambda^\nu) := \theta_\nu(x^\nu, x^{-\nu}) + \sum_{i=1}^{m_\nu} \lambda_i^\nu g_i^\nu(x^\nu, x^{-\nu})$$

and

$$\lambda := (\lambda^\nu)_{\nu=1}^N, \quad F(x,\lambda) := (\nabla_{x^\nu} L^\nu(x,\lambda^\nu))_{\nu=1}^N, \quad g(x) := (g^\nu(x))_{\nu=1}^N,$$

the concatenated KKT system can be written as

$$F(x,\lambda) = 0, \qquad \lambda \geq 0, \qquad g(x) \leq 0, \qquad \lambda^\top g(x) = 0. \tag{2}$$

By introducing slack variables $w \in \mathbb{R}^m$ and using the Hadamard product notation, $(w \circ \lambda)_i := w_i \lambda_i$ for all $i \in \mathcal{J} := \{1, \ldots, m\}$, we can reformulate the concatenated KKT system (2) as the following constrained system of equations

$$H(z) := \begin{pmatrix} F(x,\lambda) \\ g(x) + w \\ w \circ \lambda \end{pmatrix} = 0 \quad \text{s.t.} \quad z := (x,\lambda,w) \in \Omega := \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^m_+. \tag{3}$$

Obviously, a point (x,λ) is a solution of the KKT system (2) of the GNEP if and only if $w := -g(x)$ and (x,λ) solve (3). We denote the solution set of the constrained system (3) by

$$S_\Omega := \{z \in \Omega \mid H(z) = 0\}$$

and assume throughout the paper that S_Ω is nonempty.

In [5] a very robust, globally convergent interior-point algorithm for the solution of (3) was analyzed. However, this method does not have fast local convergence properties and usually does not permit to compute solutions with high precision.

To develop locally fast solution methods one can reformulate the concatenated KKT system as a possibly nonsmooth system of equations. However, the application of a local (nonsmooth) Newton method to such a system normally requires some nonsingularity assumption on the (generalized) Jacobian. But this nonsingularity condition is almost never met in the case of system (3) because it is very common to have the same constraint shared by several players, see [10]. Therefore, Newton methods for GNEPs usually require quite strong conditions that cannot be expected to hold for a general GNEP (see, e.g., [4, 10, 18]). Thus, they are applicable to a restricted class of problems only (the computation of a variational solution of a jointly convex GNEP, for example). In order to overcome this kind of difficulties, we propose to use the LP-Newton method which was suggested in [8] for the solution of constrained systems of equations. This method enjoys local quadratic convergence properties under mild assumptions that include a local error bound condition and allows nonisolated solutions and nonsmoothness of equations, see also [9] for a whole family of such methods.

In this paper we analyze the applicability of the LP-Netwon method to the constrained system (3) and show that we can expect the key error bound to hold for a broad class of GNEPs, thus covering many previously untractable games. We then combine the local LP-Newton method with the globally convergent potential reduction method in [5] and so obtain a robust and efficient solution method that theoretically seems to be superior to existing methods and numerically, as shown by our tests, behaves in a very satisfactory way.

The paper is organized as follows. In the next section we establish the error bound result. In Section 3, after recalling the methods in [5] and [8], we present our hybrid algorithm and analyze its convergence properties. Finally, in Subsection 3.4 we report numerical results showing that the new algorithm is promising.

Throughout the paper, $\|\cdot\|$ denotes the Euclidean vector norm and $\|\cdot\|_\infty$ the infinity norm. $JH : \mathbb{R}^q \to \mathbb{R}^{q\times q}$ denotes the Jacobian of H, where $q := n + 2m$ is the total

number of variables and equations in (3). For $\delta > 0$ and a given point $z \in \mathbb{R}^q$ we set $\mathcal{B}(z, \delta) := \{y \in \mathbb{R}^q \mid \|y - z\| \leq \delta\}$, i.e., $\mathcal{B}(z, \delta)$ is the closed ball around z with radius δ. Moreover, we indicate by $\operatorname{dist}[z, A]$ the distance of a point $z \in \mathbb{R}^q$ to a nonempty set $A \subseteq \mathbb{R}^q$, i.e., $\operatorname{dist}[z, A] := \inf\{\|y - z\| \mid y \in A\}$.

2 An error bound result for the KKT system

In this section we establish an error bound for the constrained system of equations (3). More precisely, given $z^* \in S_\Omega$, we prove that under a reasonable condition positive constants δ and ℓ exist such that

$$\operatorname{dist}[z, S_\Omega] \leq \ell \|H(z)\| \qquad \text{for all } z \in \mathcal{B}(z^*, \delta) \cap \Omega.$$

The interest of this result is twofold. On the one hand it is the key element in the development of our hybrid solution method to be discussed in the next section. On the other hand, to the best of our knowledge, this is the first error bound result for GNEPs that requires neither the local uniqueness of the solution and of the multipliers nor a strong strict complementarity condition. Indeed, error bounds for (KKT systems derived from) GNEPs are rare. In fact standard results require conditions that are not likely satisfied by realistic GNEPs; some error bound results for GNEPs were developed in [10], but they require some relatively strong assumptions, including strict complementarity. For later use we define the set

$$S := \{z \in \mathbb{R}^q \mid H(z) = 0\}$$

and denote by $z^* = (x^*, \lambda^*, w^*) \in S_\Omega$ an arbitrary but fixed solution of (3). Note that obviously $S_\Omega \subseteq S$ so that, since we assumed S_Ω to be nonempty, also S is nonempty. For a given point $x \in W := \{x \in \mathbb{R}^n \mid g(x) \leq 0\}$ let

$$I_=(x) := \{i \in \mathcal{J} \mid g_i(x) = 0\}$$

and

$$I_<(x) := \{i \in \mathcal{J} \mid g_i(x) < 0\}$$

denote the index set of active and inactive constraints, respectively. Moreover, let $Q(x) \leq m$ denote the maximal number of different active constraints, i.e.,

$$Q(x) := \max\{\, |\alpha| \mid \alpha \subseteq I_=(x), g_i \not\equiv g_j \text{ for all } i, j \in \alpha,\, i \neq j\}\,.$$

Then, we can define the set $\mathcal{I}_=(x)$ of all subsets of $I_=(x)$, where each repeated constraint is included exactly once, i.e.,

$$\mathcal{I}_=(x) := \{\alpha \subseteq I_=(x) \mid |\alpha| = Q(x), g_i \not\equiv g_j \text{ for all } i, j \in \alpha,\, i \neq j\}\,.$$

Now we are in the position to introduce the central assumption we are going to use in order to show the new error bound result. This assumption is tailored to take into account the main difficulty in the analysis of KKT systems of GNEPs: the possibility of shared constraints among the players, giving rise to duplicated constraint components in g.

Assumption 1. There is an index set $\alpha \in \mathcal{I}_=(x^*)$ such that $\lambda_i^* > 0$ for all $i \in \alpha$, and the matrix

$$\begin{pmatrix} J_x F(x^*, \lambda^*) & E_\alpha(x^*) \\ J g_\alpha(x^*) & 0 \end{pmatrix}$$

is nonsingular, where $E(x) := \operatorname{blockdiag}\left(\nabla_{x^1} g^1(x), \ldots, \nabla_{x^N} g^N(x)\right)$ and $E_\alpha(x)$ contains all columns of $E(x)$ whose indices belong to α.

Assumption 1 requires that for all different active constraints the multiplier of at least one player is positive. In particular, if all constraints for all players are different this is strict complementarity. However, if there are repeated active constraints, which is a typical situation for GNEPs, this assumption is weaker than strict complementarity. The nonsingularity assumption particularly requires that all different active constraints are linearly independent, which is LICQ in the case of non repeated constraints. Let us illustrate in an example, that the assumption above is weaker than the commonly used ones.

Example 1. Consider the following small 2-player game, which is example A11 from the test library in [3].

$$\begin{aligned} &\text{Player 1:} \quad \min_{x_1} (x_1 - 1)^2 \qquad \text{s.t.} \quad x_1 + x_2 \leq 1, \\ &\text{Player 2:} \quad \min_{x_2} \left(x_2 - \frac{1}{2}\right)^2 \quad \text{s.t.} \quad x_1 + x_2 \leq 1. \end{aligned}$$

One can show that the solution set of the corresponding KKT system (3) is given by

$$S_\Omega = \left\{ (t, 1-t, 2-2t, 2t-1, 0, 0)^\top \middle| t \in \left[\frac{1}{2}, 1\right] \right\}.$$

Obviously, for every solution $z^* = (x^*, \lambda^*, w^*) \in S_\Omega$ we obtain $\lambda_1^* = 2 - 2t > 0$ or $\lambda_2^* = 2t - 1 > 0$. Therefore, we can find an index set $\alpha \in \mathcal{I}_=(x^*) = \{\{1\}, \{2\}\}$ such that $\lambda_\alpha^* > 0$. Since the matrices

$$\begin{pmatrix} J_x F(x^*, \lambda^*) & E_1(x^*) \\ Jg_1(x^*) & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} J_x F(x^*, \lambda^*) & E_2(x^*) \\ Jg_2(x^*) & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

are nonsingular for all $t \in [\frac{1}{2}, 1]$, Assumption 1 is satisfied for any solution z^*. Note that the solutions corresponding to $t = \frac{1}{2}$ and $t = 1$ both violate the strict complementarity condition. Furthermore, the Jacobian of the function H is singular at all solutions $(x^*, \lambda^*, w^*) \in S_\Omega$. Therefore, locally fast convergence of classical Newton-like methods, e.g., those analyzed in [10], cannot be expected here. □

Next we provide a technical lemma on local error bounds for certain unconstrained problems. To this end, for every index set $\gamma \subseteq \mathcal{J}$, let us define the function

$$H_\gamma(x, \lambda, w) := \begin{pmatrix} F(x, \lambda) \\ g(x) + w \\ (\lambda_i w_i)_{i \notin \gamma} \\ (\lambda_i)_{i \in \gamma} \end{pmatrix}.$$

and the set

$$S_\gamma := \{z \in \mathbb{R}^q \mid H_\gamma(z) = 0\}.$$

Lemma 1. *Suppose Assumption 1 is satisfied with some set $\alpha \in \mathcal{I}_=(x^*)$. Then, there are constants $\bar{\delta} > 0$ and $\bar{\ell} > 0$ such that, for any $\gamma \subseteq \mathcal{J} \setminus \alpha$ with $z^* \in S_\gamma$,*

$$\operatorname{dist}[z, S_\gamma] \leq \bar{\ell} \|H_\gamma(z)\|$$

holds for all $z \in \mathcal{B}(z^*, \bar{\delta}) \cap \Omega$.

Proof. As a first step, we show that the mapping $\hat{H}_\gamma$ that is obtained from H_γ by dropping the components $\lambda_i w_i$ for $i \in \eta := I_=(x^*) \setminus \{\gamma \cup \alpha\}$ has locally the same set of zeros $\hat{S}_\gamma := \{z \in \mathbb{R}^q \mid \hat{H}_\gamma(z) = 0\}$ as H_γ. In other words, we show that

$$\hat{S}_\gamma \cap \mathcal{B}(z^*, \tilde{\delta}) = S_\gamma \cap \mathcal{B}(z^*, \tilde{\delta}) \tag{4}$$

holds for $\tilde{\delta} > 0$ sufficiently small. Since $S_\gamma \subseteq \hat{S}_\gamma$, only $\hat{S}_\gamma \cap \mathcal{B}(z^*, \tilde{\delta}) \subseteq S_\gamma \cap \mathcal{B}(z^*, \tilde{\delta})$ has to be shown. To this end, we choose $\tilde{\delta} > 0$ small enough so that, for any $z = (x, \lambda, w) \in \mathcal{B}(z^*, \tilde{\delta})$, we have $\lambda_i > 0$ for all $i \in \alpha$ which is possible due to Assumption 1. Moreover, for any $(x, \lambda, w) \in \hat{S}_\gamma \cap \mathcal{B}(z^*, \tilde{\delta})$, this implies $g_i(x) = w_i = 0$ for all $i \in \alpha$. Since, for every index $i \in \eta \subseteq I_=(x^*) \setminus \alpha$ there is an index $j \in \alpha$ with $g_i \equiv g_j$, we have that $(x, \lambda, w) \in \hat{S}_\gamma \cap \mathcal{B}(z^*, \tilde{\delta})$ yields $0 = g_j(x) = g_i(x)$ and, thus, $\lambda_i w_i = 0$. Hence, (4) holds true.

By definition the function $\hat{H}_\gamma$ is differentiable and its Jacobian is locally Lipschitz continuous. With $\beta := I_<(x^*) \setminus \gamma$, the Jacobian of $\hat{H}_\gamma$ at z^* is after some row and column permutations given by

$$J := \begin{pmatrix} J_x F(x^*, \lambda^*) & E_\alpha(x^*) & E_\gamma(x^*) & E_\beta(x^*) & E_\eta(x^*) & 0 & 0 & 0 \\ Jg_\alpha(x^*) & 0 & 0 & 0 & 0 & I & 0 & 0 \\ Jg_\gamma(x^*) & 0 & 0 & 0 & 0 & 0 & I & 0 \\ Jg_\beta(x^*) & 0 & 0 & 0 & 0 & 0 & 0 & I \\ 0 & 0 & 0 & 0 & 0 & \operatorname{diag}(\lambda^*_\alpha) & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \operatorname{diag}(w^*_\beta) & 0 & 0 & 0 & 0 \end{pmatrix},$$

where $w^*_\alpha = 0$, $\lambda^*_\beta = 0$ was used. Dropping the columns with $E_\eta(x^*)$ and using $w^*_\beta > 0$, $\lambda^*_\alpha > 0$ we get that the resulting matrix is nonsingular if and only if

$$\begin{pmatrix} J_x F(x^*, \lambda^*) & E_\alpha(x^*) \\ Jg_\alpha(x^*) & 0 \end{pmatrix}$$

is nonsingular. Therefore, Assumption 1 guarantees that J, and hence also the Jacobian of $\hat{H}_\gamma$ at z^*, has full row rank. By [10, Lemma 2], there are $\delta_\gamma \in \left(0, \frac{1}{2}\tilde{\delta}\right]$ and $\ell_\gamma > 0$ such that

$$\ell_\gamma \|\hat{H}_\gamma(x, \lambda, w)\| \geq \operatorname{dist}[z, \hat{S}_\gamma] \overset{(4)}{=} \operatorname{dist}[z, S_\gamma]$$

holds for all $z \in \mathcal{B}(z^*, \delta_\gamma) \cap \Omega$. Finally, with

$$\bar{\delta} := \min_\gamma \delta_\gamma > 0 \quad \text{and} \quad \bar{\ell} := \max_\gamma \ell_\gamma > 0$$

the assertion of the lemma follows from $\|\hat{H}_\gamma(z)\| \leq \|H_\gamma(z)\|$. □

Now we are in the position to prove the new local error bound result for (3).

Theorem 1. *Let Assumption 1 be satisfied. Then, there are $\delta > 0$ and $\ell > 0$ such that*

$$\operatorname{dist}[z, S_\Omega] \leq \ell \|H(z)\|$$

holds for all $z \in \mathcal{B}(z^*, \delta) \cap \Omega$.

Proof. The proof is divided into the parts a) and b). Part a) provides some preliminary facts and the components of an induction including the idea of the proof, whereas part b) is the induction itself.

a) Let $\bar{\delta} > 0$ and $\bar{\ell} > 0$ be taken from Lemma 1. Then, we choose $\hat{\delta} \in (0, \bar{\delta}]$ small enough so that $z = (x, \lambda, w) \in \mathcal{B}(z^*, \hat{\delta})$ implies

$$w_i > 0 \text{ for all } i \in I_<(x^*) \quad \text{and} \quad \lambda_i > 0 \text{ for all } i \in \Lambda_+, \tag{5}$$

where $\Lambda_+ := \{i \in \mathcal{J} \mid \lambda_i^* > 0\}$. Since H is locally Lipschitz continuous there is $L > 0$ so that

$$\|H(y) - H(y')\| \le L\|y - y'\| \qquad \text{for all } y, y' \in \mathcal{B}(z^*, \bar{\delta}). \tag{6}$$

In the following let the finite sequence $\{\delta_k\}_{k=0}^{m+1}$ be defined by

$$\delta_k := \hat{\delta}/2^{m+1-k} \quad \text{for } k = 0, \ldots, m+1.$$

For $\delta := \delta_0$, let $z^0 \in \mathcal{B}(z^*, \delta) \cap \Omega$ be arbitrarily chosen and define $\gamma_0 := \varnothing$. Then, denoting by $\hat{z}^0 \in S_{\gamma_0}$ a point with $\operatorname{dist}[z^0, S_{\gamma_0}] = \|z^0 - \hat{z}^0\|$, we have

$$\|z^0 - \hat{z}^0\| = \operatorname{dist}[z^0, S_{\gamma_0}] \le \|z^0 - z^*\| \le \delta_0 \tag{7}$$

since $z^* \in S_{\gamma_0}$ and further

$$\|\hat{z}^0 - z^*\| \le \|\hat{z}^0 - z^0\| + \|z^0 - z^*\| \le 2\delta_0 = \delta_1. \tag{8}$$

Moreover, by Lemma 1 and $S_{\gamma_0} = S$, it follows that

$$\operatorname{dist}[z^0, S_{\gamma_0}] = \|z^0 - \hat{z}^0\| \le \bar{\ell}\|H_{\gamma_0}(z^0)\|. \tag{9}$$

In part b) we will show by induction that the following objects

$$\gamma_k \subseteq \mathcal{J}, \quad z^k,\, \hat{z}^k \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m, \quad l_k, \ell_k > 0$$

can be constructed for $k = 0, \ldots, m$ so that

$$\begin{aligned}
z^k \in \Omega \text{ and } \|z^k - z^*\| &\le \delta_k, && (10)\\
\hat{z}^k \in S_{\gamma_k} \text{ and } \|\hat{z}^k - z^*\| &\le \delta_{k+1}, && (11)\\
\|z^k - \hat{z}^k\| &\le \delta_k, && (12)\\
\|z^k - \hat{z}^k\| &\le l_k\|H(z^0)\|, && (13)\\
\|z^0 - \hat{z}^k\| &\le \ell_k\|H(z^0)\|, && (14)\\
\hat{w}^k &\in \mathbb{R}^m_+ && (15)
\end{aligned}$$

are satisfied and exactly one of the two cases

$$\begin{aligned}
&\text{i)} \quad \hat{z}^k \in S_\Omega \text{ or}\\
&\text{ii)} \quad \gamma_k \subsetneq \gamma_{k+1} \subseteq \mathcal{J} \setminus \Lambda_+ \quad \text{with} \quad \hat{\lambda}_i^k < 0 \text{ if and only if } i \in \gamma_{k+1} \setminus \gamma_k
\end{aligned} \tag{16}$$

occurs. Before performing the induction suppose that i) is true for some k. Then, due to (14),

$$\operatorname{dist}[z^0, S_\Omega] \le \|z^0 - \hat{z}^k\| \le \ell_k\|H(z^0)\|$$

follows, i.e., the theorem would be proven with $\ell := \ell_k$. Since in ii) the index set γ_{k+1} is always strictly larger than γ_k case ii) can occur a finite number of times only. Hence, case i) will eventually occur for some $k \leq m$ and, more importantly, the theorem will be proved after the induction below.

b) Obviously, due to $\gamma_0 = \varnothing$, $z^0 \in \mathcal{B}(z^*, \delta_0) \cap \Omega$, $\hat{z}^0 \in S_{\gamma_0}$, (7), (8), and (9), we have that (10) – (14) hold for $k = 0$ with $\ell_0 := l_0 := \bar{\ell}$. Having in mind the second part of (11) for $k = 0$, we get from (5) that $\hat{w}_i^0 > 0$ for $i \in I_<(x^*)$ and

$$\hat{\lambda}_i^0 > 0 \quad \text{for all } i \in \Lambda_+. \tag{17}$$

By $\hat{z}^0 \in S_{\gamma_0}$, this implies $\hat{w}_i^0 = 0$ for all $i \in \alpha$ and, thus, $g_i(\hat{x}^0) = \hat{w}_i^0 = 0$ for all $i \in I_=(x^*)$. So, (15) is satisfied for $k = 0$. Suppose that $\hat{z}^0$ does not belong to S_Ω. Then, since $\hat{z}^0 \in S_{\gamma_0} = S$ and by (15) for $k = 0$ and (17), a nonempty index set γ_1 must exist so that ii) in (16) is valid for $k = 0$. Hence, in (16) either case i) or case ii) occurs.

By assuming that (10) – (16) are satisfied for some k we are now going to prove that (10) – (16) are valid for $k+1$ instead of k. To this end, we construct $z^{k+1} := (\hat{x}^k, \lambda^{k+1}, \hat{w}^k)$ by

$$\lambda_i^{k+1} := \begin{cases} 0, & \text{if } i \in \gamma_{k+1}, \\ \hat{\lambda}_i^k, & \text{if } i \notin \gamma_{k+1}. \end{cases},$$

which implies

$$\lambda^{k+1} \geq 0 \tag{18}$$

and $|\lambda_i^{k+1} - \lambda_i^k| \leq |\hat{\lambda}_i^k - \lambda_i^k|$ for all i. Hence, we get

$$\|z^{k+1} - z^k\| \leq \|\hat{z}^k - z^k\| \tag{19}$$

and, using (12) and (10),

$$\|z^{k+1} - z^*\| \leq \|z^{k+1} - z^k\| + \|z^k - z^*\| \leq 2\delta_k = \delta_{k+1}. \tag{20}$$

Moreover, since $z^{k+1} = (\hat{x}^k, \lambda^{k+1}, \hat{w}^k)$, (15) and (18) yield

$$z^{k+1} \in \Omega. \tag{21}$$

By the definition of λ^{k+1}, we have $\lambda_i^{k+1} = 0$ for all $i \in \gamma_{k+1}$ so that

$$H_{\gamma_{k+1}}(z^{k+1}) = H(z^{k+1}) \tag{22}$$

follows. Due to (5) and ii) of (16) we know that $\lambda_i^* = 0$ for all $i \in \gamma_{k+1} \subseteq \mathcal{J} \setminus \Lambda_+$. Therefore,

$$z^* \in S_{\gamma_{k+1}} \tag{23}$$

and we can choose

$$\hat{z}^{k+1} \in S_{\gamma_{k+1}} \tag{24}$$

so that $\text{dist}[z^{k+1}, S_{\gamma_{k+1}}] = \|z^{k+1} - \hat{z}^{k+1}\|$. Then, by (23) and (20), we have

$$\|z^{k+1} - \hat{z}^{k+1}\| \leq \|z^{k+1} - z^*\| \leq \delta_{k+1}. \tag{25}$$

and

$$\|\hat{z}^{k+1} - z^*\| \leq \|\hat{z}^{k+1} - z^{k+1}\| + \|z^{k+1} - z^*\| \leq 2\delta_{k+1} = \delta_{k+2}. \tag{26}$$

From ii) of (16) it follows that $\gamma_{k+1} \subseteq \mathcal{J} \setminus \Lambda_+ \subseteq \mathcal{J} \setminus \alpha$. Thus, by (20), we can apply Lemma 1 and obtain with (22)

$$\|z^{k+1} - \hat{z}^{k+1}\| = \operatorname{dist}[z^{k+1}, S_{\gamma_{k+1}}] \leq \bar{\ell}\|H_{\gamma_{k+1}}(z^{k+1})\| = \bar{\ell}\|H(z^{k+1})\|. \tag{27}$$

By $z^0 \in \mathcal{B}(z^*, \delta_0)$ and (20), the local Lipschitz continuity of H given by (6) yields

$$\begin{array}{rcl} \|H(z^{k+1})\| & \leq & \|H(z^{k+1}) - H(z^0)\| + \|H(z^0)\| \\ & \leq & L\|z^{k+1} - z^0\| + \|H(z^0)\|. \end{array} \tag{28}$$

Using (19), (13), and (14), we get

$$\begin{array}{rcl} \|z^{k+1} - z^0\| & \leq & \|z^{k+1} - z^k\| + \|z^k - \hat{z}^k\| + \|\hat{z}^k - z^0\| \\ & \leq & 2\|\hat{z}^k - z^k\| + \|\hat{z}^k - z^0\| \\ & \leq & (2l_k + \ell_k)\|H(z^0)\| \end{array}$$

so that (28) leads to

$$\|H(z^{k+1})\| \leq (L(2l_k + \ell_k) + 1)\|H(z^0)\|.$$

With $l_{k+1} := \bar{\ell}(L(2l_k + \ell_k) + 1)$, this and (27) imply

$$\|z^{k+1} - \hat{z}^{k+1}\| \leq l_{k+1}\|H(z^0)\|. \tag{29}$$

Thus, setting $\ell_{k+1} := l_{k+1} + (2l_k + \ell_k)$, we obtain

$$\begin{array}{rcl} \|z^0 - \hat{z}^{k+1}\| & \leq & \|z^0 - z^{k+1}\| + \|z^{k+1} - \hat{z}^{k+1}\| \\ & \leq & \ell_{k+1}\|H(z^0)\|. \end{array} \tag{30}$$

By (26) we get from (5) that $\hat{w}_i^{k+1} > 0$ for all $i \in I_<(x^*)$ and

$$\hat{\lambda}^{k+1} > 0 \qquad \text{for all } i \in \Lambda_+. \tag{31}$$

Due to (24), the latter implies $\hat{w}_i^{k+1} = 0$ for all $i \in \alpha \subseteq \Lambda_+$ and, thus, $g_i(\hat{x}^{k+1}) = \hat{w}_i^{k+1} = 0$ for all $i \in I_=(x^*)$. Therefore, we have

$$\hat{w}^{k+1} \in \mathbb{R}^m_+. \tag{32}$$

Now suppose that $\hat{z}^{k+1} \notin \Omega$. Then, taking into account (31) and (32), and $\hat{\lambda}_i^{k+1} = 0$ for $i \in \gamma_{k+1}$, there is at least one index $i \in \mathcal{J} \setminus (\Lambda_+ \cup \gamma_{k+1})$ with $\hat{\lambda}_i^{k+1} < 0$. Thus, defining

$$\gamma_{k+2} := \gamma_{k+1} \cup \{i \in \mathcal{J} \setminus \Lambda_+ \mid \hat{\lambda}^{k+1} < 0\}$$

guarantees

$$\gamma_{k+1} \subsetneq \gamma_{k+2} \subseteq \mathcal{J} \setminus \Lambda_+ \quad \text{with} \quad \hat{\lambda}_i^{k+1} < 0 \text{ if and only if } i \in \gamma_{k+2} \setminus \gamma_{k+1}.$$

Finally, with this and (20), (21), (24) – (26), (29), (30), and (32), the induction is successfully completed. □

The following example shows that the error bound condition in Theorem 1 may be too difficult to obtain if we relax Assumption 1.

Example 2. Consider the 2-player game

$$\begin{aligned} \text{Player 1:} \quad & \min_{x_1} \frac{1}{4}x_1^4 \quad \text{s.t.} \quad x_1 + x_2 \geq 1, \\ \text{Player 2:} \quad & \min_{x_2} \frac{1}{4}x_2^4 \quad \text{s.t.} \quad x_1 + x_2 \geq 1. \end{aligned}$$

The solution set of the corresponding system (3) is given by

$$S_\Omega = \left\{(t, 1-t, t^3, (1-t)^3, 0, 0)^\top \mid t \in [0,1]\right\}.$$

Consider the solution $z^* = (0,1,0,1,0,0)^\top$ and the sequence $\{z^k\} \subset \Omega$ with $z^k := \left(-\frac{1}{k}, 1+\frac{1}{k}, 0, 1+\frac{3}{k}, 0, 0\right)^\top$. Obviously, $\{z^k\}$ converges to z^* for $k \to \infty$. Now, on the one hand, we have

$$\begin{aligned} \|H(z^k)\|^2 &= \|H(x^k, \lambda^k, w^k)\|^2 = \left\| \begin{pmatrix} (x_1^k)^3 - \lambda_1^k \\ (x_2^k)^3 - \lambda_2^k \\ -x_1^k - x_2^k + 1 + w_1^k \\ -x_1^k - x_2^k + 1 + w_2^k \\ \lambda_1^k w_1^k \\ \lambda_2^k w_2^k \end{pmatrix} \right\|^2 \\ &= \left(\frac{1}{k^3}\right)^2 + \left(\frac{3}{k^2} + \frac{1}{k^3}\right)^2 = O\left(\frac{1}{k^4}\right). \end{aligned}$$

On the other hand, the special construction of the sequence $\{z^k\}$ yields

$$\operatorname{dist}[z^k, S_\Omega]^2 = \|z^* - z^k\|^2 = \frac{11}{k^2}.$$

Therefore, H does not provide a local error bound near z^*. There are two index sets $\alpha_1, \alpha_2 \in \mathcal{I}_=(x^*)$. For $\alpha_1 = \{1\}$ the matrix

$$\begin{pmatrix} J_x F(z^*) & E_{\alpha_1}(x^*) \\ Jg_{\alpha_1}(x^*) & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

is nonsingular, but $\lambda_1^* = 0$. Thus, Assumption 1 is not satisfied for α_1. For $\alpha_2 = \{2\}$ we have $\lambda_2^* = 1 > 0$, but

$$\begin{pmatrix} J_x F(z^*) & E_{\alpha_2}(x^*) \\ Jg_{\alpha_2}(x^*) & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

is singular. Altogether, the example shows that Assumption 1 in Theorem 1 can not be replaced by only one of its parts, the nonsingularity assumption, or $\lambda_\alpha^* > 0$. □

3 A globally and locally fast convergent hybrid algorithm

In this section we combine two recently proposed algorithms and define a new solution method for GNEPs that enjoys the best theoretical convergence properties among the algorithms in the literature. Our new method derives from a suitable combination of the

potential reduction method proposed in [5] for the solution of GNEPs and a very recent Newton type method studied in [8] for the solution of constrained systems of equations with locally fast convergence. The idea of combining a globally convergent algorithm with possibly poor local convergence rate with a second algorithm which is only locally convergent but enjoys a fast convergence rate is certainly not new and actually part of the folklore in optimization. However, the assumptions under which such combinations are shown to work properly usually imply the local uniqueness of the solution towards which the algorithm converges. Unfortunately, when dealing with GNEPs, local uniqueness of the solutions is rarely encountered, and this poses a serious challange. One of the main merits of our developments in this section is therefore to couple the global and local methods in a not totally standard way and to carry out a refined analysis so as to be able to show global and fast local convergence without resorting to any requirement about local uniqueness of the solutions.

In the next two subsections we briefly recall the methods in [5] and [8], respectively. Then, in Subsection 3.3, we introduce our hybrid algorithm and analyze its convergence properties. Finally, Subsection 3.4 reports numerical results.

3.1 A potential reduction algorithm for the solution of GNEPs

In [5], based on the algorithm first proposed in [21], an interior point method for the solution of (3) is suggested that is based on the minimization of a potential function. In this section we want to recall this algorithm and a convergence result proven in [5]. First, let us repeat some definitions. For a given real number $\zeta > m$ the function $P : \mathbb{R}^n \times \mathbb{R}^{2m}_{++} \to \mathbb{R}$ is defined by

$$P(u,v) := \zeta \log(\|u\|^2 + \|v\|^2) - \sum_{i=1}^{2m} \log(v_i).$$

Furthermore, let $Z_I := \{z = (x, \lambda, w) \in \mathbb{R}^n \times \mathbb{R}^m_{++} \times \mathbb{R}^m_{++} \mid g(x) + w > 0\}$ denote the set of all strictly feasible points for which the last $2m$ components of H are positive. The potential function $\psi : Z_I \to \mathbb{R}$ is defined by

$$\psi(z) := P(H(z)).$$

With $a^\top := (0_n^\top, 1_{2m}^\top)$ we can now present the algorithm from [5].

Algorithm 1: Potential Reduction Algorithm for GNEPs

(S.0): Choose $z^0 \in Z_I$ and $\beta, \eta \in (0,1), \zeta > m$. Set $k := 0$.

(S.1): If $H(z^k) = 0$ then STOP.

(S.2): Choose $\sigma_k \in [0,1)$ and compute a solution d^k of the linear system

$$JH(z^k)d = -H(z^k) + \sigma_k \frac{a^\top H(z^k)}{\|a\|^2}\, a.$$

(S.3): Compute a stepsize $t_k := \max\{\beta^i \mid i = 0, 1, 2, \ldots\}$ such that

$$z^k + t_k d^k \in Z_I \quad \text{and}$$
$$\psi(z^k + t_k d^k) \le \psi(z^k) + \eta t_k \nabla\psi(z^k)^\top d^k.$$

(S.4): Set $z^{k+1} := z^k + t_k d^k$, $k := k + 1$, and go to (S.1).

The next result follows from Theorem 4.3 in [5].

Theorem 2. *Assume that $JH(z)$ is nonsingular for all $z \in Z_I$ and that the sequence $\{\sigma_k\}$ satisfies the condition*

$$\limsup_{k\to\infty} \sigma_k < 1.$$

Let $\{z^k\}$ be any sequence generated by Algorithm 1. Then the following assertions hold:

(a) The sequence $\{H(z^k)\}$ is bounded.

(b) Any accumulation point of $\{z^k\}$ is a solution of (3).

The nonsingularity of $JH(z)$ on Z_I guarantees that Algorithm 1 is well defined and the iterates z^k are contained in Z_I, see [5] for details. Therefore, the case $H(z^k) = 0$ can (theoretically) never occur. In [5, 13] the interested reader can find some sufficient conditions for the nonsingularity of the Jacobian for all $z \in Z_I$.

3.2 The LP-Newton method for the solution of GNEPs

In this subsection we briefly recall the LP-Newton method from [8] as applied to our constrained system of equations (3). The subproblems of the LP-Newton method are optimization problems. For a given point $z^k = (x^k, \lambda^k, w^k) \in \mathbb{R}^q$ a solution $(z^{k+1}, \gamma^{k+1}) \in \mathbb{R}^q \times \mathbb{R}$ of the program

$$\begin{aligned} \min_{z,\gamma} \gamma \quad \text{s.t.} \quad & z \in \Omega, \\ & \|H(z^k) + JH(z^k)(z - z^k)\|_\infty \le \gamma \|H(z^k)\|_\infty^2, \\ & \|z - z^k\|_\infty \le \gamma \|H(z^k)\|_\infty \end{aligned} \tag{33}$$

has to be computed. Since in our case $\Omega = \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^m_+$ is a polyhedron and the infinity norm is used, (33) is a linear program. This program is solvable for any $z^k \notin S$ so that the following algorithm is well defined, see also [8, Proposition 1]. We recall the algorithm from [8].

Algorithm 2: LP-Newton Method for GNEPs

`(S.0)`: Choose $z^0 \in \Omega$ and set $k := 0$.

`(S.1)`: If $H(z^k) = 0$ then STOP.

`(S.2)`: Compute a solution (z^{k+1}, γ^{k+1}) of (33).

`(S.3)`: Set $k := k + 1$ and go to `(S.1)`.

In [8] four assumptions were used to prove local quadratic convergence of the LP-Newton method. However, due to the smoothness properties of H and in view of [8, Corollary 1], these assumptions essentially boil down to the requirement that H satisfies a local error bound condition.

Theorem 3. *Let Assumption 1 be satisfied. Then, there is $r > 0$ such that any infinite sequence $\{z^k\}$ generated by Algorithm 2 with starting point $z^0 \in \mathcal{B}(z^*, r) \cap \Omega$ converges Q-quadratically to some $\hat{z} = (\hat{x}, \hat{\lambda}, \hat{w}) \in S_\Omega$.*

Proof. Our smoothness assumptions imply local Lipschitz continuity of H. Therefore, Assumption 1 from [8] is valid. Due to Theorem 1 a local error bound holds for problem (3). Thus, Assumption 2 from [8] is satisfied and, in view of the smoothness properties of H and [8, Corollary 1] Assumptions 3 and 4 from [8] are satisfied as well. Hence, the assertion on the local Q-quadratic convergence follows from [8, Theorem 1]. □

We conclude this subsection by noting that due to the error bound provided in Theorem 1 and the smoothness of H we could also employ a constrained Levenberg-Marquardt method [19] or an inexact version [9] as local method. The numerical comparison of several local approaches is left to future work.

3.3 A hybrid algorithm for the solution of GNEPs

We are now ready to describe a hybrid algorithm enjoying both global and fast local convergence properties. The definitions of the function ψ, the set Z_I and the vector a can be found in Subsection 3.1.

Algorithm 3: Hybrid Algorithm for GNEPs

`(S.0)`: Choose $z^0 := \hat{z} \in Z_I$ and parameters $\beta, \eta, \theta \in (0,1)$, $\zeta > m$, $0 < \tau_{\min} \leq \tau_{\max}$, $\tau_0 \in [\tau_{\min}, \tau_{\max}]$. Set $k := 0$.

`(S.1)`: If $H(z^k) = 0$ then STOP.

If $\|H(z^k)\| \leq \tau_k$ set $\sigma_k := 0$ and go to `(S.4)`, else go to `(S.2)`.

`(S.2)`: Choose $\sigma_k \in [0,1)$ and compute a solution d^k of the linear system

$$JH(z^k)d = -H(z^k) + \sigma_k \frac{a^\top H(z^k)}{\|a\|^2}\, a.$$

`(S.3)`: Compute a stepsize $t_k := \max\left\{\beta^i \mid i = 0, 1, 2, \dots\right\}$ such that

$$z^k + t_k d^k \in Z_I \quad \text{and}$$
$$\psi(z^k + t_k d^k) \leq \psi(z^k) + \eta t_k \nabla\psi(z^k)^\top d^k.$$

Set $z^{k+1} := z^k + t_k d^k$, $\hat{z} := z^{k+1}$, $\tau_{k+1} := \tau_k$, $k := k+1$, go to `(S.1)`.

`(S.4)`: Compute a solution $(\tilde{z}^{k+1}, \gamma^{k+1})$ of the linear program

$$\begin{aligned} \min_{z,\gamma} \gamma \qquad \text{s.t.} \qquad & z \in \Omega, \\ & \|H(z^k) + JH(z^k)(z - z^k)\|_\infty \leq \gamma \|H(z^k)\|_\infty^2, \\ & \|z - z^k\|_\infty \leq \gamma \|H(z^k)\|_\infty. \end{aligned}$$

If $\|H(\tilde{z}^{k+1})\| \leq \theta \|H(z^k)\|$ then
set $z^{k+1} := \tilde{z}^{k+1}$, $\tau_{k+1} := \tau_k$, $k := k+1$, go to `(S.1)`,
else set $z^{k+1} := \hat{z}$, $k := k+1$, choose $\tau_{k+1} \in [\tau_{\min}, \tau_{\max}]$, go to `(S.2)`.

As long as the iterates are not close to a solution, i.e., if $\|H(z^k)\| \geq \tau_k$, Algorithm 3 coincides with Algorithm 1. Therefore, in each iteration a linear system of equations is solved in (S.2) and a linesearch procedure is performed in (S.3). When supposedly near

a solution the algorithm switches to an LP-Newton step in (S.4), a solution of a linear program is computed. If the LP-Newton steps do not yield sufficiently fast convergence we assume that we actually were still too far from a solution. In this case, the algorithm switches back to the interior point method from the last iterate $\hat{z}$ that was used before the switch to the LP-Newton. Note that we can not easily use an iterate z computed by the LP-Newton part to continue with the potential reduction part, since $z \in Z_I$ is not guaranteed.

By choosing $\tau_{k+1} \in [\tau_{\min}, \tau_{\max}]$ in Step (S.4) one can somehow control the neighborhood in which Algorithm 3 switches to the LP-Newton subproblem. A practical rule for adjusting τ_{k+1} is given in Subsection 3.4.

The following theorem shows that the new Hybrid Algorithm for GNEPs enjoys both global and local quadratic convergence properties.

Theorem 4. *Assume that $JH(z)$ is nonsingular for all $z \in Z_I$ and that $\limsup_{k\to\infty} \sigma_k < 1$. Then, Algorithm 3 is well defined. Moreover, for any infinite sequence $\{z^k\}$ generated by the algorithm the following assertions hold:*

(a) The sequence $\{H(z^k)\}$ is bounded.

(b) Any accumulation point of $\{z^k\}$ is a solution of (3).

(c) If $\{z^k\}$ has an accumulation point z^* where Assumption 1 is satisfied then the whole sequence $\{z^k\}$ converges to z^* with a Q-quadratic rate.

Proof. Algorithm 3 guarantees that the potential reduction part ((S.2) and (S.3)) always starts with an iterate $z^k \in Z_I$ since, on the one hand, the potential reduction itself generates only those points. On the other hand, whenever the algorithm switches from the LP-Newton step (S.4) back to (S.2) then (S.2) uses the iterate (saved in $\hat{z}$) that was produced by the very last run of the potential reduction part. Therefore, the nonsigularity of $JH(z)$ assumed for all $z \in Z_I$ guarantees that both the steps (S.2) and (S.3) are well defined, for details see [5]. As we mentioned in Subsection 3.2, any LP-Newton step (S.4) is well defined if $H(z^k) \neq 0$. This is true since any iterate generated in (S.3) belongs to Z_I and, thus, satisfies this. Moreover, the algorithm stops in (S1) if $H(z^k) = 0$ is encountered.

For any sequence $\{z^k\}$ generated by Algorithm 3 exactly one of the following two cases can occur:

(i) Whenever Algorithm 3 switches to (S.4), LP-Newton steps are used only a finite number of times before the algorithm gets back to (S.2) and (S.3).

(ii) For a certain iterate Algorithm 3 switches to (S.4) and then uses LP-Newton steps only.

As we already pointed out above, whenever Algorithm 3 gets back from (S.4) to the potential reduction part in (S.2) and (S.3), it employs the iterate $\hat{z}$ which was generated when step (S.3) was run the last time. Therefore, if case (i) occurs, Algorithm 3 is equivalent to the potential reduction algorithm (Algorithm 1) and assertions (a) and (b) follow from Theorem 2.

In case (ii) we have a $k_0 \in \mathbb{N}$ such that $\|H(z^{k+1})\| \leq \theta \|H(z^k)\|$ for all $k \geq k_0$, and since $\theta \in (0, 1)$, this also implies (a) and (b).

To prove that assertion (c) is valid let z^* denote an accumulation point of $\{z^k\}$ satisfying Assumption 1. Then, Theorem 1 and the local Lipschitz continuity of JH imply, by

[8, Proposition 2], that Assumptions 1 – 4 from [8] are satisfied. Hence, due to Lemma 1 and Lemma 2 from [8], we can conclude that there are $C_1 > 0$, $C_2 > 0$, and $\delta_* \in (0, \delta)$, with δ from Theorem 1, such that

$$\|z(z^k) - z^k\| \leq C_1 \operatorname{dist}[z^k, S_\Omega] \tag{34}$$

and

$$\operatorname{dist}[z(z^k), S_\Omega] \leq C_2 \operatorname{dist}[z^k, S_\Omega]^2 \leq \frac{1}{2} \operatorname{dist}[z^k, S_\Omega] \tag{35}$$

for all $z^k \in \mathcal{B}(z^*, \delta_*) \cap \Omega$ and any solution $z(z^k)$ of the LP-Newton subproblem (33). Since $\delta_* < \delta$, H is Lipschitz continuous on $\mathcal{B}(z^*, \delta_*)$ with modulus $L > 0$, see (6). With $\ell > 0$ from Theorem 1, let us define

$$\epsilon := \min\left\{\frac{\delta_*}{C_1 + 1}, \frac{\theta}{C_2 L \ell}\right\} \quad \text{and} \quad \rho := \frac{\epsilon}{1 + 2C_1}.$$

Since z^* is an accumulation point of $\{z^k\}$ and $\tau_k \geq \tau_{\min} > 0$ holds for all $k \in \mathbb{N}$, there is $K \in \mathbb{N}$ so that

$$z^K \in \mathcal{B}(z^*, \rho) \cap \Omega \quad \text{and} \quad \|H(z^K)\| \leq \tau_K.$$

The latter means that Step (S.4) is run for this particular K. We now show by induction that the following assertions are valid for all $k \geq K$:

$$z^k \in \mathcal{B}(z^*, \epsilon) \cap \Omega \tag{36}$$

and

$$z^{k+1} = \tilde{z}^{k+1} \text{ is generated in (S.4)}, \tag{37}$$

where $\tilde{z}^{k+1}$ denotes the corresponding solution of the LP-Newton subproblem (33) in (S.4). For $k = K$ assertion (36) is clear since $\rho < \epsilon$. Thus, using (34), we have

$$\|\tilde{z}^{K+1} - z^*\| \leq \|\tilde{z}^{K+1} - z^K\| + \|z^K - z^*\| \leq (C_1 + 1)\epsilon \leq \delta_*$$

The local Lipschitz continuity of H on $\mathcal{B}(z^*, \delta_*)$ now yields

$$\|H(\tilde{z}^{K+1})\| \leq L \operatorname{dist}[\tilde{z}^{K+1}, S_\Omega].$$

Therefore, by the first inequality in (35), (36) for $k = K$, Theorem 1, and the definition of ϵ, we obtain

$$\|H(\tilde{z}^{K+1})\| \leq C_2 L \operatorname{dist}[z^K, S_\Omega]^2 \leq C_2 L \ell \epsilon \|H(z^K)\| \leq \theta \|H(z^K)\|,$$

so that the test in (S.4) is satisfied and $z^{K+1} = \tilde{z}^{K+1}$ is generated in (S.4), i.e., (37) is true for $k = K$.

Suppose now that (36) and (37) hold for $k = K, \ldots, \kappa$ with some $\kappa \geq K$. In particular, all iterates $z^{K+1}, \ldots, z^{\kappa+1}$ were generated by the LP-Newton step (S.4). Using the triangle inequality, we obtain

$$\|z^{\kappa+1} - z^*\| \leq \underbrace{\|z^K - z^*\|}_{\leq \rho} + \sum_{j=K}^{\kappa} \|z^{j+1} - z^j\|. \tag{38}$$

Taking into account (36) for $k = K, \ldots, \kappa$, (34), and (35), it follows for all $j = K, \ldots, \kappa$

$$\|z^{j+1} - z^j\| \leq C_1 \operatorname{dist}[z^j, S_\Omega] \leq C_1 \left(\frac{1}{2}\right)^{j-K} \operatorname{dist}[z^K, S_\Omega].$$

This, (38), and the definition of ρ imply

$$\|z^{\kappa+1} - z^*\| \le \rho + C_1 \underbrace{\operatorname{dist}[z^K, S_\Omega]}_{\le \rho} \underbrace{\sum_{j=0}^{\kappa-K} \left(\frac{1}{2}\right)^j}_{\le 2} \le (1 + 2C_1)\rho \le \epsilon.$$

Thus, $z^{\kappa+1} \in \mathcal{B}(z^*, \epsilon) \cap \Omega$ holds. Moreover, (37) for $k = \kappa + 1$ can be shown as for $k = K$. Therefore, (36) and (37) are valid for all $k \ge K$. Thus, for $k \ge K$, Algorithm 3 runs LP-Newton steps only. Hence, we are in case (ii) and the sequence $\{z^k\}$ converges by Theorem 3 Q-quadratically to a solution of (3). Obviously, z^* must be the limit. This shows assertion (c). □

3.4 Numerical results

Now, we briefly describe results of a preliminary implementation of the hybrid algorithm and compare it to the potential reduction algorithm. The algorithms were implemented in Matlab® with the parameters

$$\beta = 0.5,\ \eta = 10^{-2},\ \theta = 0.9,\ \zeta = 2m,\ \tau_{\min} = 10^{-11},\ \tau_{\max} = \tau_0 = 10^{-3}.$$

For updating σ_k in (S.2) and τ_k in the else-branch of (S.4) we used

$$\sigma_k = 0.1 \quad \text{and} \quad \tau_{k+1} = \max\{\tau_{\min}, \theta\|H(z^k)\|\}.$$

If the matrix $JH(z^k)$ is ill conditioned, i.e., if its estimated condition number is larger than 10^{16}, we try to avoid running into numerical trouble by solving the linear system in (S.2) with the pertubed matrix $JH(z^k) + 10^j I$, where the smallest $j \in \{-2, -1, 0, \ldots\}$ is used so that the estimated condition number of $JH(z^k) + 10^j I$ is not larger than 10^{16}. By this strategy we avoid that the algorithm fails to compute a search direction. Furthermore, if the computed direction d^k from Step (S.2) does not satisfy $\nabla\psi(z^k)^\top d^k \le -10^{-8}\|d^k\|^{2.1}$, that is if the direction d^k is almost orthogonal to $\nabla\psi(z^k)$, we use the anti-gradient direction $-\nabla\psi(z^k)$ as search direction d^k, which was already suggested in the numerical part of [5]. For reporting results with these modifications in Step (S.3) we refer to "Algorithm 1 (modified)" and "Algorithm 3 (modified)". For the solution of the linear program in (S.4) the "lp-minos" solver from the Optimization Toolbox of Tomlab® was applied.

The algorithms were tested on a library of GNEPs proposed in [5]; these test problems are discussed in detail in [3]. In order to compare our results the algorithms use the same stopping criterion

$$V(x, \lambda) := \frac{1}{\sqrt{n+m}} \left\| \begin{pmatrix} F(x, \lambda) \\ \min\{-g(x), \lambda\} \end{pmatrix} \right\| \le \varepsilon$$

with accuracy $\varepsilon \in \{10^{-4}, 10^{-6}, 10^{-8}, 10^{-10}\}$. A failure is deteced if the maximum iteration limit of 100 is exceeded (using a larger iteration limit essentially does not change the results described below). Also if the stepsize t_k in (S.3) becomes smaller than $t_{\min} := 10^{-8}$ the algorithms stop with a failure. Depending on the accuracy ε, Table 1 reports the number of failures. A first observation is that the modifications of Step (S.3) lead to an improvement in the robustness of the potential reduction method for small accuracy, see the results in Table 1 for Algorithm 1 and for Algorithm 1 (modified). Therefore, we decided to use the above modifications also for Step (S.3) of Algorithm 3 which leads to Algorithm 3 (modified). Moreover, Table 1 shows that Algorithm 3 (modified) is

Table 1: Failures on 57 test runs depending on the accuracy ε.

ε	10^{-4}	10^{-6}	10^{-8}	10^{-10}
Algorithm 1	6	8	10	11
Algorithm 1 (modified)	1	4	9	14
Algorithm 3 (modified)	1	2	2	2

significantly more robust on the test library then the pure potential reduction algorithms. The new hybrid algorithm is able to calculate in all but two cases a solution with a quite high precision. This, coupled with the better theoretical convergence properties makes the hybrid algorithm look very promising and a good candidate for developing a general purpose solution algorithm for GNEPs. In most of the test runs the hybrid algorithm switches only once or twice to the LP-Newton step (S.4), then stays there, and converges in a few iterations. To illustrate this we have added Table 2 which shows more detailed results on problem Tr3 ($N = 7$, $n = 80$, $m = 304$). The potential reduction method (Algorithm 1 (modified)) only reaches the accuracy bound of $\varepsilon = 10^{-6}$ and fails for $k = 63$, whereas the hybrid method (Algorithm 3 (modified)) shows no failure and reaches $V(x^k, \lambda^k) \leq 10^{-10}$. Moreover, neither the regularization strategy nor the use of an anti-gradient direction $-\nabla\psi(z^k)$ improves the situation for the potential reduction method.

Table 2: Detailed numerical results for problem Tr3.

	Algorithm 3 (modified)				Algorithm 1 (modified)	
k	LS	LP	stepsize t_k	$V(x^k, \lambda^k)$	stepsize t_k	$V(x^k, \lambda^k)$
0	0	0	0.000000	64.555823931987		
1	1	0	0.015625	63.533509705715		
2	2	0	0.015625	62.239660552331		
3	3	0	0.031250	59.563728071708		
4	4	0	0.062500	54.773243024867		
5	5	0	0.125000	47.399007958429		
10	10	0	0.500000	18.025445688612		
20	20	0	0.125000	0.279047403143		
30	30	0	0.250000	0.033002833980		
40	40	0	0.062500	0.007860354749		
50	50	0	0.250000	0.002304240403		
55	55	0	0.250000	0.000653997253		
56	56	0	0.500000	0.000563301075		
57	57	0	0.500000	0.000510204495		
58	58	0	0.500000	0.000286881425		
59	59	0	1.000000	0.000043808966		
60	59	1	1.000000	0.000043808966	1.000000	0.000006142406
61	60	1	1.000000	0.000006142406	1.000000	0.000000877795
62	60	2	1.000000	0.000002017306	1.000000	0.000000202210
63	60	3	1.000000	0.000000282627	failure since $t_k < t_{\min}$	
64	60	4	1.000000	0.000000000067		

Table 2 is built as follows. Its first column contains the counter k for the iteration number. The next four columns are for Algorithm 3 (modified), reporting the number

of linear systems solved (LS), the number of linear programs solved (LP), the stepsizes computed in Step (S.3), and the values $V(x^k, \lambda^k)$. The last two columns report the latter values for Algorithm 1 (modified). Note that the first iterations (potential reduction steps) are the same for both methods and therefore are only reported once in the column belonging to Algorithm 3 (modified).

References

[1] Altangerel, L., Battur, G.: Perturbation approach to generalized Nash equilibrium problems with shared constraints. Optim. Lett. **6**, 1379–1391 (2012)

[2] Dorsch, D., Jongen, H.T., Shikhman, V.: On intrinsic complexity of Nash equilibrium problems and bilevel optimization. J. Optim. Appl., published online (2012) 10.1007/s10957-012-0210-7

[3] Dreves, A.: Globally Convergent Algorithms for the Solution of Generalized Nash Equilibrium Problems. Dissertation, University of Würzburg (2012), available online at `http://opus.bibliothek.uni-wuerzburg.de/volltexte/2012/6982`

[4] Dreves, A., von Heusinger, A.,: Kanzow, C., Fukushima, M.: A globalized Newton method for the computation of normalized Nash equilibria. J. Global Optim. published online (2011), DOI 10.1007/s10898-011-9824-9

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

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

[7] Fabian, M.J., Henrion, R., Kruger, A.Y., Outrata, J.V.: Error Bounds: Necessary and Sufficient Conditions. Set-Valued Var. Anal. **18**, 121–149 (2010)

[8] Facchinei, F., Fischer, A., Herrich. M.: An LP-Newton method: Nonsmooth equations, KKT systems, and nonisolated solutions. Technical Report MATH-NM-5-2011, Institute of Numerical Mathematics, Technische Universität Dresden (2011), available online at `http://www.optimization-online.org/DB_HTML/2011/09/3177.html`

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

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

[11] Facchinei, F., Kanzow, C.: Generalized Nash equilibrium problems. Ann. Oper. Res. **1755**, 177–211 (2010).

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

[13] Facchinei, F., Kanzow, C., Sagratella, S.: Solving quasi-variational inequalities via their KKT-conditions. Preprint 305, Institute of Mathematics, University of Würzburg (2012), available online at
http://www.mathematik.uni-wuerzburg.de/∼kanzow/paper/QVI_12_01_2012P.pdf

[14] Facchinei, F., Pang, J.-S.: Exact penalty function for generalized Nash games. In: Di Pillo, G., Roma, M. (eds.) Large-scale Nonlinear Optimization, Series Nonconvex Optimization and Its Applications **83** pp. 115–126. Springer, New York (2006)

[15] Facchinei, F., Pang, J.-S.: Nash equilibria: The variational approach. In: Eldar, Y. Palomar, D. (eds.) Convex Optimization in Signal Processing and Communications, pp. 443-493. Cambridge University Press, Cambridge (2009)

[16] Fukushima, M.: Restricted generalized Nash equilibria and controlled penalty algorithm. Comput. Manag. Sci. **8**, 201–218 (2011)

[17] Han, D., Zhang, H., Qian, G., Xu, L.: An improved two-step method for solving generalized Nash equilibrium problems. European J. Oper. Res. **216**, 613–623 (2012)

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

[19] Kanzow, C., Yamashita, N., Fukushima, M.: Levenberg-Marquardt methods with strong local convergence properties for solving equations with convex constraints. J. Comput. Appl. Math. **172**, 375–397 (2004)

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

[21] Monteiro, R.D.C., Pang, J.-S.: A potential reduction Newton method for constrained equations. SIAM J. Optim. **9**, 729–754 (1999)

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

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

[24] Pang, J.-S., Scutari, G.: Nonconvex games with side constraints. SIAM J. Optim. **21**, 1491–1522 (2012)

[25] Schiro, D.A., Pang, J.-S., Shanbhag U.V.: On the solution of affine generalized Nash equilibrium problems with shared constraints by Lemkes method. Math. Program., published online (2012) DOI 10.1007/s10107-012-0558-3

[26] Scutari, G., Facchinei, F., Pang, J.-S., Palomar, D.: Real and complex monotone communication games. Submitted to IEEE Transactions on Information Theory