Penalty Methods for the Solution of Generalized Nash Equilibrium Problems

Francisco Facchinei
La Sapienza University of Rome

joint work with Christian Kanzow

Zürich, 6 April 2009
Outline

- Definition of the problem
- Review of algorithms
- The new penalty method
PART I

Definition of the problem
Optimization Problem

\[ \min_x \theta(x) \quad x \in X \]

- One “optimizer”
- A single function to be optimized
Nash Equilibrium Problems (NEP)

- **N** players
- Each player controls \(x^\nu \in \mathbb{R}^{n_\nu}\)
- Set \(n := \sum_{\nu=1}^{N} n_\nu\), \(x := \begin{pmatrix} x^1 \\ \vdots \\ x^N \end{pmatrix} \in \mathbb{R}^n\), \(x^{-\nu} := \begin{pmatrix} x^1 \\ \vdots \\ x^{\nu-1} \\ x^{\nu+1} \\ \vdots \\ x^N \end{pmatrix}\)
- \(x = (x^\nu, x^{-\nu})\)
\[
\min_{x^1} \quad \theta_1(x^1, x^{-1}) \quad \min_{x^\nu} \quad \theta_\nu(x^\nu, x^{-\nu}) \quad \min_{x^N} \quad \theta_N(x^N, x^{-N})
\]

\[
x^1 \in X_1 \quad \cdots \quad x^\nu \in X_\nu \quad \cdots \quad x^N \in X_N
\]

- Several “optimizers (or players)”
- Every player minimizes a different obj. f.
- The obj. f. depend on the variables of the other players
- The feasible sets are independent of the choices of the other players
$S_\nu(x^{-\nu})$: Optimal solution set of player $\nu$ for a given $x^{-\nu}$ of the other players

$x$ is a Nash equilibrium if

$$x^\nu \in S_\nu(x^{-\nu})$$

for all players $\nu$

No player can improve by unilaterally deviating from the current situation
Generalized Nash Equilibrium Problem (GNEP)

\[
\begin{align*}
\min_{x^1} & \quad \theta_1(x^1, x^{-1}) \\
\min_{x^\nu} & \quad \theta_\nu(x^\nu, x^{-\nu}) \\
\min_{x^N} & \quad \theta_N(x^N, x^{-N})
\end{align*}
\]

\[
\begin{align*}
x^1 & \in X_1(x^{-1}) \\
x^\nu & \in X_\nu(x^{-\nu}) \\
x^N & \in X^N(x^{-N})
\end{align*}
\]

- Several “optimizers (or players)”
- Every optimizer minimizes a different obj. f.
- The obj. f. depend on the variables of the other players
- Also the feasible sets depend on the other players’ variables
$S_\nu(x^{-\nu})$: Optimal solution set of player $i$ for a given $x^{-\nu}$ of the other players

$x$ is a Nash equilibrium if

$x^\nu \in S_\nu(x^{-\nu})$ for all players $\nu$
Short History

- Prehistory:
  - Cournot (1838)
  - von Neumann (1928)
  - von Neumann and Morgenstern (1944)

- Nash (1950/1) $\Rightarrow$ Nash Equilibrium Problem
- Debreu (1952)
- Arrow and Debreu (1954) $\Rightarrow$ Generalized Nash Eq.

- 1954–beginning of the 1990s: The GNEP is studied mainly within the economic domain

- From the 1990s, modern engineering applications:
  - Robinson (1993, effectiveness in combat models)
  - Scotti (1995, structural design)
  - liberalized markets, telecommunications, web protocols, pollution analysis....
The economic equilibrium model: how are commodities produced and exchanged among individuals?

Walras was probably the first author to tackle this issue in a modern mathematical perspective.

Arrow and Debreu considered a general economic system along with a corresponding definition of equilibrium. They then showed that the equilibria of their model are those of a suitably defined GNEP; on this basis, they were able to prove important results on the existence of economic equilibria.

- $l$ commodities
- $s$ production units that control $y^j \in \mathbb{R}^l$
- $t$ consumption units that control $x^i \in \mathbb{R}^l$
- One (fictitious) "market" player that controls the prices $p \in \mathbb{R}^l$
The production-player's problem:

$$\max_{y^j} \quad p^T y^j$$

$$s.t. \quad y^j \in Y^j$$

The consumption-player’s problem:

$$\max_{x^i} \quad u_i(x^i)$$

$$s.t. \quad x^i \in X_i$$

$$p^T x^i \leq p^T \xi^i + \max \left( 0, \sum_{j=1}^{s} \alpha_{ij}(p^T y^j) \right)$$

The market-players problem:

$$\max_{p} \quad p^T \left( \sum_{i=1}^{t} x^i - \sum_{j=1}^{s} y^j - \sum_{i=1}^{t} \xi^i \right)$$

$$s.t. \quad p \geq 0$$

$$\sum_{h=1}^{l} p_h = 1$$
PART II

Algorithms
A general scheme

Given a GNEP

Transform it into another (better understood) problem X
Solve problem X
Two warnings

- In order to solve problem X one must impose some conditions; when these conditions are “brought back” to the original GNEP setting they are often very strong or not very often verified in practice (or both)

- There is too little practical experience in the solution of GNEPs to be able to make a serious comparison of the practical behavior of various methods
Conversion of a NEP to a VI

- $\theta_{\nu}(x^\nu, x^{-\nu})$ convex in $x^\nu$ for every $x^{-\nu}$
- $\theta_{\nu}$ continuously differentiable
- $X_{\nu}$ closed and convex

Solve the NEP $\iff$ Solve $VI(K, F)$

$$K := \prod_{\nu=1}^{N}X_{\nu}, \quad F(x) := \begin{pmatrix} \nabla_{x^1} \theta_1(x) \\ \vdots \\ \nabla_{x^N} \theta_N(x) \end{pmatrix}$$
Conversion of a \textbf{GNEP} to a \textbf{QVI}

- $\theta_{\nu}(x^\nu, x^{-\nu})$ convex in $x^\nu$ for every $x^{-\nu}$
- $\theta_{\nu}$ continuously differentiable
- $X_{\nu}(x^{-\nu})$ closed and convex

\begin{align*}
K(x) &:= \prod_{\nu=1}^{N} X_{\nu}(x^{-\nu}), \\
F(x) &:= \begin{pmatrix}
\nabla_{x^1} \theta_1(x) \\
\vdots \\
\nabla_{x^N} \theta_N(x)
\end{pmatrix}
\end{align*}

Solving the $QVI(K, F)$ means finding $\bar{x} \in K(\bar{x})$ s.t. $F(\bar{x})^T(y - \bar{x}) \geq 0, \forall y \in K(\bar{x})$
Using the KKT conditions

Assume the feasible sets are explicitly given by a set of parametric inequalities

\[
\min_{x^\nu} \quad \theta_\nu(x^\nu, \bar{x}^{-\nu}) \\
x^\nu \in X_\nu(\bar{x}^{-\nu})
\]  
\[\implies\]

\[
\min_{x^\nu} \quad \theta_\nu(x^\nu, \bar{x}^{-\nu}) \\
g^\nu(x^\nu, \bar{x}^{-\nu}) \leq 0
\]

We can write down the KKT conditions for each player

\[
\nabla_{x^\nu} \theta_\nu(x^\nu, \bar{x}^{-\nu}) + \nabla_{x^\nu} g^\nu(x^\nu, \bar{x}^{-\nu})^T \lambda^\nu = 0, \\
0 \leq \lambda^\nu \perp -g^\nu(x^\nu, \bar{x}^{-\nu}) \geq 0
\]

- Concatenate the KKT systems of all players and try to solve
- Very recently some theoretical results based on an homothopy approach
Assume that (as usual) some constraints do not depend on $x^{-\nu}$,

$$
\min_{x^\nu} \quad \theta_{\nu}(x^\nu, x^{-\nu}) \\
g^\nu(x^\nu, x^{-\nu}) \leq 0 \quad \iff \quad \min_{x^\nu} \quad \theta_{\nu}(x^\nu, x^{-\nu}) + \rho P(g^\nu(x^\nu, x^{-\nu})) \\
h^\nu(x^\nu) \leq 0
$$

where

$$
P(g^\nu(x^\nu, x^{-\nu})) = \begin{cases} 
\|\max(0, g^\nu(x^\nu, x^{-\nu}))\|^2 & \text{sequential penalization} \\
\|\max(0, g^\nu(x^\nu, x^{-\nu}))\| & \text{exact penalization}
\end{cases}
$$

- Solve the resulting NEP
For an important class of GNEPs some other possibilities arise

Let $K \subset \mathbb{R}^n$ be closed and convex

GNEP is **jointly convex** if $X_\nu(x^-) := \{x^\nu : (x^\nu, x^-) \in K\}$ or, equivalently, if

$$X_\nu(x^-) = \{x^\nu : g(x^\nu, x^-) \leq 0\}$$

with $g(\cdot, \cdot)$ convex

Solve the GNEP. Solve $VI(K, F)$

Jointly Convex

The solutions of the VI are called **normalized equilibria**
An example: internet switching problem

\[ x^\nu = \begin{align*} 
\text{packets sent to buffer} \\
\text{by player } \nu 
\end{align*} \]

maximize \[ x^\nu \left( 1 - \frac{\sum_{\nu=1}^{N} x^\nu}{B} \right) \]

subject to \[ x^\nu \geq 0, \]
\[ \sum_{\nu=1}^{N} x^\nu \leq B \]
\[
\min_x \ (x - 1)^2 \quad \quad \min_y \ (y - \frac{1}{2})^2
\]
\[
x + y \leq 1 \quad \quad x + y \leq 1.
\]

Solutions: \((\alpha, 1 - \alpha)\) for every \(\alpha \in [1/2, 1]\)

This is a jointly convex GNEP

\[
K = \{(x, y) : x + y \leq 1\}, \quad F = \begin{pmatrix} 2x - 2 \\ 2y - 1 \end{pmatrix}
\]

VI \((K, F)\) is strongly monotone and its unique solution is \((3/4, 1/4)\); this is the normalized equilibrium of the GNEP.
Having reduced the (G)NEP to a VI we must pay attention.

The assumptions needed for the convergence of VI algorithms are often very strong (if not meaningless) when “translated” back into the (G)NEP environment. For example, a condition often invoked for the convergence of VI algorithms is that $F$ be monotone. These means assuming that

$$JF(x) = \begin{pmatrix}
\frac{\partial^2 \theta_1(x)}{\partial x^1 \partial x^1} & \cdots & \frac{\partial^2 \theta_1(x)}{\partial x^1 \partial x_\nu} & \cdots & \frac{\partial^2 \theta_1(x)}{\partial x^1 \partial x_N} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
\frac{\partial \theta_\nu(x)}{\partial x_\nu \partial x^1} & \cdots & \frac{\partial^2 \theta_\nu(x)}{\partial x_\nu \partial x_\nu} & \cdots & \frac{\partial^2 \theta_\nu(x)}{\partial x_\nu \partial x_N} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
\frac{\partial \theta_N(x)}{\partial x_N \partial x^1} & \cdots & \frac{\partial^2 \theta_N(x)}{\partial x_N \partial x_\nu} & \cdots & \frac{\partial^2 \theta_N(x)}{\partial x_N \partial x_N}
\end{pmatrix}
$$

be positive semidefinite. Is this sensible?
The mapping

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

is called the *Nikaido-Isoda function*. Each summand gives the improvement that the $\nu$-th player gets when changing (unilaterally) from $x^{\nu}$ to $y^{\nu}$.

Define the *nonnegative* function

$$V(x) := \sup_{y \in K} \Psi(x, y) \geq 0$$

$x \in K$ is a normalized equilibrium $\iff V(x) = 0$
Solve the problem \( \min_{x \in K} V(x) \) in order to get a normalized equilibrium.

Note that in general \( V \) is nondifferentiable and difficult to compute.

Specialized methods can be developed, the most important one being the relaxation method. Also regularized (differentiable) versions exist.

Relaxation method Pros:

- It is probably the method that has been used most in practice
- Conceptually does not require the differentiability of the \( \theta \).

Relaxation method Cons:

- Assumptions are still strong
- Computationally very intensive
<table>
<thead>
<tr>
<th></th>
<th>Ott.</th>
<th>Nash</th>
<th>Gen. Nash</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Alg.</td>
<td>Yes</td>
<td>Conversion to VI</td>
<td>“Jointly Convex” <img src="Image" alt="General" /></td>
</tr>
<tr>
<td>Newton</td>
<td>Yes</td>
<td>Conversion to VI</td>
<td>Yes, but <img src="Image" alt="Penalties &amp; this talk" /></td>
</tr>
<tr>
<td>Decompos.</td>
<td>Yes</td>
<td>Conversion to VI</td>
<td><img src="Image" alt="Via KKT" /></td>
</tr>
</tbody>
</table>

In these cases caution must be exerted

[Facchini et al. (soon)](Image) Specific applications
PART III

The New Penalty Approach
Penalization again

\[
\begin{align*}
\min_{x^\nu} & \quad \theta^\nu(x^\nu, x^{-\nu}) \\
g^\nu(x^\nu, x^{-\nu}) \leq 0 & \quad \Longrightarrow \quad \min_{x^\nu} \quad \theta^\nu(x^\nu, x^{-\nu}) + \rho^\nu \|g^\nu_+(x^\nu, x^{-\nu})\|_3 \\
& \quad x^\nu \in \mathbb{R}^{n^\nu}
\end{align*}
\]

We set \( P^\nu(x^\nu, x^{-\nu}, \rho^\nu) = \theta^\nu(x^\nu, x^{-\nu}) + \rho^\nu \|g^\nu_+(x^\nu, x^{-\nu})\|_3 \)

We assume throughout “sufficient” differentiability of all functions involved

Important: the 3-norm is used, so that the penalty term is continuously differentiable at all infeasible points
Two main issues:

- How to choose and/or update the penalty parameter $\rho$
- How to solve the (unconstrained) penalized problems

We decouple completely these two issues.
In order to deal with the updating of the penalty parameter $\rho$ we assume temporarily that an iterative algorithm $\mathcal{A}$ is available that, given a point $x^k$, generates a new point $x^{k+1} := \mathcal{A}[x^k]$. We make the following absolutely natural blanket assumption on $\mathcal{A}$.

For every $x^0$, the sequence $\{x^k\}$ obtained by $x^{k+1} = \mathcal{A}[x^k]$ is such that every limit point (if any) is a solution of the unconstrained penalized problem.

All results on the updating below hold whatever the algorithm $\mathcal{A}$
How to update $\rho$ II

If, for any value of the penalty parameters $\rho_\nu$, we find a solution of the penalized problem that is also feasible for the original constrained problem, then this is a solution of original constrained problem. We try to force feasibility by increasing the penalty parameters.

If a solution $\bar{x}$ of the penalty problem is not feasible for player $\nu$, then $P_\nu(x, \rho_\nu)$ is continuously differentiable at $\bar{x}$ so that

$$\| \nabla_{x_\nu} \theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \| = \rho_\nu \| \nabla_{x_\nu} \| g^\nu_+(\bar{x}^\nu, \bar{x}^{-\nu}) \|_3 \| .$$

The idea of the updating scheme is to detect when this “dangerous” situation occurs (see the test at Step 2), and to increase the value of the penalty parameter in this case.
Updating Scheme

**Data:** \( x^0 \in \mathbb{R}^n \) and \( \rho_\nu > 0 \) for all \( \nu = 1, \ldots, N \). Set \( k := 0 \).

**Step 1:** If \( x^k \) is a solution of the GNEP: STOP.

**Step 2:** Let \( I^k := \{ \nu \mid (x^k)\nu \notin X_\nu((x^k)_{-\nu}) \} \) (violated constraints)

For every \( \nu \in I^k \), if

\[
\| \nabla x^\nu \theta_\nu((x^k)\nu, (x^k)_{-\nu}) \| > 0.1 \left[ \rho_\nu \left\| \nabla x^\nu \left\| g^\nu + ((x^k)^\nu, (x^k)_{-\nu}) \right\|_3 \right\| \right],
\]

then double the penalty parameters \( \rho_\nu \).

**Step 3:** Compute \( x^{k+1} = A[x^k] \), set \( k \leftarrow k + 1 \), and go to Step 1.
Let \( \{x^k\} \) be the sequence generated by Updating Scheme. If the penalty parameters are updated a finite number of times only, then every limit point \( \bar{x} \) of this sequence is a solution of the GNEP.

If instead some penalty parameters grow to infinity and the sequence \( \{x^k\} \) is bounded, then, for each \( \nu \) such that \( \rho_\nu \to \infty \), there is a limit point \( \bar{x} \) for which one of the following assertions is true:

(a) \( \bar{x}^\nu \) is a global minimizer of the constraint violation \( \|g_+^\nu(\cdot, \bar{x}^{-\nu})\|_3 \) with \( \|g_+^\nu(\bar{x}^\nu, \bar{x}^{-\nu})\|_3 > 0 \);

(b) \( \bar{x}^\nu \) is Fritz John point for the player’s problem, but not a solution of it

(c) \( \bar{x}^\nu \) is an optimal solution for the player’s problem.

Our next aim is giving conditions ensuring that (a), (b) and (c) cannot occur, so that the \( \rho_\nu \) remain finite and every limit point is a solution of the GNEP.
How to update $\rho$ V

We need some constraint qualifications

- $\partial_{x^\nu}^* \|g^\nu_+(x^\nu, x^{-\nu})\|_3 := \{ \xi \in \mathbb{R}^{n\nu} \mid \exists \{y^k\} with (y^k)^\nu not feasible for player \nu such that \{y^k\} \to x and \nabla_{x^\nu} \|g^\nu_+((y^k)^\nu, (y^k)^{-\nu})\|_3 \to \xi \}$

We say that the GNEP satisfies the constraint qualification $CQ_3$ at a point $\bar{x}$ if

$$0 \notin \partial_{x^\nu}^* \|g^\nu_+(\bar{x}^\nu, \bar{x}^{-\nu})\|_3, \quad \forall \nu = 1, \ldots, N$$

- We say that the GNEP satisfies the $EMFCQ$ at a point $\bar{x}$ if, for every player $\nu = 1, \ldots, N$, there exists a vector $d^\nu$ such that

$$\nabla_{x^\nu} g^\nu_i(\bar{x}^\nu, \bar{x}^{-\nu})^T d^\nu < 0 \quad \forall i \in I^\nu_+(\bar{x}),$$

where $I^\nu_+(\bar{x}) := \{ i \in \{1, \ldots, m^\nu \} \mid g^\nu_i(\bar{x}^\nu, \bar{x}^{-\nu}) \geq 0 \}$
Assume that the sequence \( \{x^k\} \) generated by Updating Scheme is bounded. Consider the following assertions:

(a) The EMFCQ holds at every limit point \( \bar{x} \) of \( \{x^k\} \);
(b) The CQ\(_3\) condition holds at every limit point \( \bar{x} \) of \( \{x^k\} \);
(c) The penalty parameters are updated a finite number of times only

Then the following implications hold:

\[(a) \implies (b) \implies (c).\]
Our algorithm $\mathcal{A}$ is based on smoothing. Recall that the objectives functions of the penalized game are

\[
P_{\nu}(x, \rho_{\nu}) = P_{\nu}(x^\nu, x^{-\nu}, \rho_{\nu})
\]
\[
= \theta_{\nu}(x^\nu, x^{-\nu}) + \rho_{\nu}\|g_+^{\nu}(x^\nu, x^{-\nu})\|_3
\]
\[
= \theta_{\nu}(x^\nu, x^{-\nu}) + \rho_{\nu}\left(\sum_{i=1}^{m_{\nu}} \max\{0, g_i^{\nu}(x^\nu, x^{-\nu})\}^3\right)^{1/3}.
\]

We approximate these functions by the smooth mappings

\[
\tilde{P}_{\nu}(x, \rho_{\nu}, \varepsilon) := \tilde{P}_{\nu}(x^\nu, x^{-\nu}, \rho_{\nu}, \varepsilon)
\]
\[
:= \theta_{\nu}(x^\nu, x^{-\nu}) + \rho_{\nu}\left(\sum_{i=1}^{m_{\nu}} \max\{0, g_i^{\nu}(x^\nu, x^{-\nu})\}^3 + \varepsilon\right)^{1/3}
\]
\[
+ \frac{\varepsilon}{2}\|x^\nu\|^2,
\]

where $\varepsilon > 0$.
We approximate the nonsmooth penalized game by the unconstrained smooth (actually $C^2$) game where each player’s problem is $\tilde{P}_\nu$ instead of $P_\nu$. The solutions of this smoothed game are the solutions of the following system of equations:

\[
F_\varepsilon(x) := \left( \begin{array}{c}
\nabla_{x^1} \tilde{P}_1(x^1, x^{-1}, \rho_1, \varepsilon) \\
\vdots \\
\nabla_{x^N} \tilde{P}_N(x^N, x^{-N}, \rho_N, \varepsilon)
\end{array} \right) = 0. \tag{1}
\]

Let $\{\varepsilon_k\}$ and $\{\eta_k\}$ be two sequences of positive numbers converging to 0 and, for every $k$, let $x(\varepsilon_k)$ be a point such that

\[
\|F_{\varepsilon_k}(x(\varepsilon_k))\| \leq \eta_k.
\]

Then every limit point of the sequence $x(\varepsilon_k)$ is a solution of the nonsmooth penalized game.
On this basis we can, roughly speaking, use any standard method for the solution of a smooth system of equations as algorithm $\mathcal{A}$.

The one theoretical problem we are left to deal with is “Can we ensure that we are able to solve the smooth systems $F_\varepsilon(x) = 0$?”

We do not actually need to solve these systems at each iteration but eventually we want to get more and more accurate solutions. While the approach seems to work very well in practice, this theoretical issue is still under investigation.
The overall algorithm

Data: \( \mathbf{x}^0 \in \mathbb{R}^n, \rho_\nu > 0, \varepsilon_0 > 0, S > 0 \) and integer. Set \( k := 0 \).

**Step 1:** If \( \mathbf{x}^k \) is such that

1. \( \| \max \{ 0, g(\mathbf{x}^k) \} \| \leq 10^{-6} \);  
2. \( \varepsilon^k \leq 10^{-6} \);  
3. \( \| F_{\varepsilon_k}(\mathbf{x}) \| \leq 10^{-6} \);

then STOP.

**Step 2:** Let \( I^k := \{ \nu \mid (\mathbf{x}^k)_\nu \not\in X_\nu((\mathbf{x}^k)^{-\nu}) \} \). For every \( \nu \in I^k \), if

\[
\| \nabla_{x^\nu} \theta_\nu((\mathbf{x}^k)_\nu, (\mathbf{x}^k)^{-\nu}) \| > 0.1 \left[ \rho_\nu \left\| \nabla_{x^\nu} g^\nu_+((\mathbf{x}^k)_\nu, (\mathbf{x}^k)^{-\nu}) \right\|_3 \right],
\]

then double the penalty parameter \( \rho_\nu \).

**Step 3:** Perform at most \( S \) steps of an equation solver to the nonlinear system of equations \( F_{\varepsilon_k}(\mathbf{x}) = 0 \). Let \( \mathbf{x}^{k+1} \) be the final iterate of this equation solver. 

If \( \| F_{\varepsilon_k}(\mathbf{x}^{k+1}) \| \leq 1000 \varepsilon_k \), set \( \varepsilon_{k+1} = 0.1 \varepsilon_k \), otherwise let \( \varepsilon_{k+1} := \varepsilon_k \).

Set \( k \leftarrow k + 1 \), and go to Step 1.
### Numerical results: non jointly convex

<table>
<thead>
<tr>
<th>Ex.</th>
<th>N</th>
<th>n</th>
<th>St.p.</th>
<th>k</th>
<th>( i_{\text{total}} )</th>
<th>( \varepsilon_f )</th>
<th>( | F_{\varepsilon_f}(x_f) | )</th>
<th>( | g_+(x_f) | )</th>
<th>( \rho_{f_{\max}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>10</td>
<td>10</td>
<td>0.01</td>
<td>5</td>
<td>34</td>
<td>1e-7</td>
<td>3.6e-13</td>
<td>7.6e-4</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.1</td>
<td>5</td>
<td>33</td>
<td>1e-7</td>
<td>3.6e-13</td>
<td>7.6e-4</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>5</td>
<td>30</td>
<td>1e-7</td>
<td>3.6e-13</td>
<td>7.6e-4</td>
<td>10</td>
</tr>
<tr>
<td>A2</td>
<td>10</td>
<td>10</td>
<td>0.01</td>
<td>6</td>
<td>46</td>
<td>1e-8</td>
<td>2.4e-12</td>
<td>6.2e-4</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.1</td>
<td>6</td>
<td>44</td>
<td>1e-08</td>
<td>7.6e-13</td>
<td>9.1e-4</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>3</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>1e-5</td>
<td>1.7e-11</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>1e-5</td>
<td>2.3e-11</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>3</td>
<td>8</td>
<td>1e-5</td>
<td>1.5e-11</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>A4</td>
<td>3</td>
<td>7</td>
<td>0</td>
<td>16</td>
<td>247</td>
<td>1e-10</td>
<td>1.4e-10</td>
<td>5.6e-4</td>
<td>1e+5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>29</td>
<td>510</td>
<td>1e-08</td>
<td>1.4e-10</td>
<td>5.6e-4</td>
<td>1e+5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>22</td>
<td>351</td>
<td>1e-9</td>
<td>1.6e-10</td>
<td>3.8e-4</td>
<td>1e+4</td>
</tr>
<tr>
<td>A5</td>
<td>3</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>43</td>
<td>1e-5</td>
<td>6.2e-6</td>
<td>4.3e-4</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>43</td>
<td>1e-5</td>
<td>6.2e-6</td>
<td>4.3e-4</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>5</td>
<td>62</td>
<td>1e-5</td>
<td>8.9e-16</td>
<td>4.3e-4</td>
<td>100</td>
</tr>
<tr>
<td>A6</td>
<td>3</td>
<td>7</td>
<td>0</td>
<td>10</td>
<td>88</td>
<td>1e-10</td>
<td>5.3e-9</td>
<td>3.5e-4</td>
<td>1e+4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>9</td>
<td>121</td>
<td>1e-9</td>
<td>4.7e-10</td>
<td>6.1e-4</td>
<td>1e+4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>10</td>
<td>99</td>
<td>1e-10</td>
<td>5.4e-9</td>
<td>3.5e-4</td>
<td>1e+4</td>
</tr>
<tr>
<td>A7</td>
<td>4</td>
<td>20</td>
<td>0</td>
<td>12</td>
<td>148</td>
<td>1e-9</td>
<td>3.6-10</td>
<td>0.001</td>
<td>1e+4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>12</td>
<td>151</td>
<td>1e-9</td>
<td>3.6-10</td>
<td>0.001</td>
<td>1e+4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A8</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>81</td>
<td>1e-5</td>
<td>4.6e-13</td>
<td>6.3e-5</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>4</td>
<td>31</td>
<td>1e-6</td>
<td>9.4e-14</td>
<td>2.0e-4</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A9a</td>
<td>7</td>
<td>56</td>
<td>0</td>
<td>6</td>
<td>169</td>
<td>1e-8</td>
<td>2.8e-11</td>
<td>3.7e-3</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>22</td>
<td>1034</td>
<td>1e-10</td>
<td>5.9e-8</td>
<td>9.6e-4</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A9b</td>
<td>7</td>
<td>112</td>
<td>0</td>
<td>11</td>
<td>246</td>
<td>1e-9</td>
<td>6.4e-10</td>
<td>1.5e-3</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>8</td>
<td>203</td>
<td>1e-8</td>
<td>3.1e-11</td>
<td>8.4e-3</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>101</td>
<td>14637</td>
<td>1e-8</td>
<td>4.8e-9</td>
<td>0.01</td>
<td>1e+4</td>
</tr>
<tr>
<td>A10a</td>
<td>8</td>
<td>24</td>
<td>0</td>
<td>11</td>
<td>1282</td>
<td>1e-7</td>
<td>2.3e-11</td>
<td>0.03</td>
<td>1000</td>
</tr>
<tr>
<td>A10c</td>
<td>37</td>
<td>222</td>
<td>0</td>
<td>12</td>
<td>1392</td>
<td>1e-7</td>
<td>1.6e-10</td>
<td>0.04</td>
<td>1000</td>
</tr>
<tr>
<td>A10d</td>
<td>37</td>
<td>370</td>
<td>0</td>
<td>11</td>
<td>1282</td>
<td>1e-7</td>
<td>2.3e-11</td>
<td>0.03</td>
<td>1000</td>
</tr>
<tr>
<td>A10e</td>
<td>48</td>
<td>576</td>
<td>0</td>
<td>12</td>
<td>1392</td>
<td>1e-7</td>
<td>1.6e-10</td>
<td>0.04</td>
<td>1000</td>
</tr>
</tbody>
</table>
### Numerical results: jointly convex

<table>
<thead>
<tr>
<th>Ex.</th>
<th>$N$</th>
<th>$n$</th>
<th>St.p.</th>
<th>$k$</th>
<th>$i_{total}$</th>
<th>$\varepsilon_f$</th>
<th>$|F_{\varepsilon_f}(x_f)|$</th>
<th>$|g_+(x_f)|$</th>
<th>$\rho_{max}^f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>9</td>
<td>60</td>
<td>1e-11</td>
<td>9.9e-12</td>
<td>9.6e-5</td>
<td>10</td>
</tr>
<tr>
<td>B2</td>
<td>2</td>
<td>2</td>
<td>(2, 0)</td>
<td>3</td>
<td>5</td>
<td>1e-5</td>
<td>3.6e-12</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>B3</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>10</td>
<td>84</td>
<td>1e-11</td>
<td>1.7e-9</td>
<td>1.5e-4</td>
<td>10</td>
</tr>
<tr>
<td>B4</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>3</td>
<td>16</td>
<td>1e-5</td>
<td>2.1e-10</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>B5</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>3</td>
<td>9</td>
<td>1e-5</td>
<td>4.7e-10</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>B6</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>102</td>
<td>1e-10</td>
<td>1.3e-7</td>
<td>3.9e-4</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td>74</td>
<td>1e-10</td>
<td>1.7e-8</td>
<td>3.1e-4</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>101</td>
<td>1e-11</td>
<td>3.3e-8</td>
<td>2.9e-4</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>73</td>
<td>1e-11</td>
<td>2.8e-8</td>
<td>2.3e-4</td>
<td>10</td>
</tr>
<tr>
<td>B7</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>9</td>
<td>71</td>
<td>1e-10</td>
<td>1.5e-10</td>
<td>1.9e-4</td>
<td>100</td>
</tr>
<tr>
<td>B8</td>
<td>2</td>
<td>12</td>
<td>0</td>
<td>19</td>
<td>346</td>
<td>1e-7</td>
<td>1.1e-7</td>
<td>0.001</td>
<td>1e+4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>12</td>
<td>1e-7</td>
<td>4.1e-7</td>
<td>5.2e-4</td>
<td>10e+4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>11</td>
<td>1e-7</td>
<td>4.4e-8</td>
<td>5.7e-4</td>
<td>1e+4</td>
</tr>
</tbody>
</table>
### Numerical results for Problem B7

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\rho_k$</th>
<th>$(x_1^1)^k$</th>
<th>$(x_1^2)^k$</th>
<th>$(x_2^2)^k$</th>
<th>$| \max{0, g(x^k)} |$</th>
<th>$\varepsilon_k$</th>
<th>$| F_{\varepsilon_k}(x^k) |$</th>
<th>$i_{\text{total}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>$10^{-3}$</td>
<td>0.000</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1.9482194</td>
<td>14.5776644</td>
<td>4.4138139</td>
<td>31.6005704</td>
<td>$10^{-3}$</td>
<td>0.000</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>0.3147472</td>
<td>10.6578689</td>
<td>7.5978619</td>
<td>0.0652462</td>
<td>$10^{-4}$</td>
<td>3.607</td>
<td>36</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>0.0012125</td>
<td>11.0019638</td>
<td>7.9970862</td>
<td>0.0186015</td>
<td>$10^{-4}$</td>
<td>0.000</td>
<td>41</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>0.0003897</td>
<td>11.0010847</td>
<td>7.9988202</td>
<td>0.0086351</td>
<td>$10^{-5}$</td>
<td>0.000</td>
<td>47</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>0.0001635</td>
<td>11.0005208</td>
<td>7.9994696</td>
<td>0.0040083</td>
<td>$10^{-6}$</td>
<td>0.000</td>
<td>53</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>0.0000742</td>
<td>11.0002434</td>
<td>7.9997555</td>
<td>0.0018605</td>
<td>$10^{-7}$</td>
<td>0.000</td>
<td>59</td>
</tr>
<tr>
<td>7</td>
<td>100</td>
<td>0.0000342</td>
<td>11.0001132</td>
<td>7.9998867</td>
<td>0.0008636</td>
<td>$10^{-8}$</td>
<td>0.000</td>
<td>65</td>
</tr>
<tr>
<td>8</td>
<td>100</td>
<td>0.0000158</td>
<td>11.0000525</td>
<td>7.9999474</td>
<td>0.0004008</td>
<td>$10^{-9}$</td>
<td>0.000</td>
<td>71</td>
</tr>
<tr>
<td>9</td>
<td>100</td>
<td>0.0000073</td>
<td>11.0000244</td>
<td>7.9999756</td>
<td>0.0001860</td>
<td>$10^{-10}$</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>
The algorithm compares favourably to existing penalty methods (Fukushima Pang (2005), Facchinei Pang (2006), Fukushima (2008))

We have proposed, studied and tested a penalty method for the solution of GNEPs

The algorithms seems to perform very well in practice

We believe that what we report is by far the largest numerical testing in the literature (for general problems)

Some theoretical questions on the algorithm $\mathcal{A}$ require further study

Many variants are possible and we plan to explore them