

A new smoothing method for nonlinear complementarity problems involving \mathcal{P}_0 -function

El Hassene Osmani ^{1,2,*}, Mounir Haddou ², Naceurdine Bensalem ¹, Lina Abdallah ³

¹University Ferhat Abbas of Setif 1, Faculty of Sciences, Laboratory of Fundamental and Numerical Mathematics, Setif 19000, Algeria

²University of Rennes, INSA Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France

³Lebanese university, Tripoli, Lebanon

Abstract In this paper, we present a family of smoothing methods to solve nonlinear complementarity problems (NCPs) involving \mathcal{P}_0 -function. Several regularization or approximation techniques like Fisher-Burmeister's method, interior-point methods (IPMs) approaches, or smoothing methods already exist. All the corresponding methods solve a sequence of nonlinear systems of equations and depend on parameters that are difficult to drive to zero. The main novelty of our approach is to consider the smoothing parameters as variables that converge by themselves to zero. We do not need any complicated updating strategy, and then obtain nonparametric algorithms. We prove some global and local convergence results and present several numerical experiments, comparisons, and applications that show the efficiency of our approach.

Keywords Nonlinear complementarity problems, Nonparametric algorithm, Newton's method, Smoothing functions, \mathcal{P}_0 -function, Interior-point methods

AMS 2010 subject classifications 47H05, 90C33

DOI: 10.19139/soic-2310-5070-1493

1. Introduction

The nonlinear complementarity problem (NCP) consists in finding $x \in \mathbb{R}^n$ satisfying

$$x \geq 0 \quad F(x) \geq 0 \quad x^T F(x) = 0, \quad (1)$$

where $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$. When F is linear, problem (1) reduces to a linear complementarity problem (LCP).

NCPs arise in many practical applications, for example, the Karush-Kuhn-Tucker (KKT) systems of mathematical programming problem, economic equilibria, and engineering design problems, can be formulated as NCPs (see, for instance, [11, 20, 29]).

Different concepts have been developed to study and solve these problems: reformulation as a system of nonlinear equations or a minimization problem (see [12, 15, 23, 27, 30, 31, 32]). Recently, there have been strong interests in equation reformulation methods for solving the NCPs. One of the most effective methods is to transform the NCP into semi-smooth equation (NCP functions) and solve using semi-smooth Newton methods. The most well-known NCP functions are the Fisher-Burmeister's function introduced by Fisher Burmeister in [13] and the min function studied by Kanzow, Yamashita and Fukushima [22]. Another well-known class of algorithms corresponds the smoothing methods. The main idea of smoothing approaches is to approximate or regularize the NCP to obtain smooth equations depending on some parameter (see, for example, [7, 24, 25, 33, 39]).

*Correspondence to: El Hassene Osmani (Email: el-hassene.osmani@insa-rennes.fr). University of Rennes, INSA Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France.

In this paper, we present, smoothing approximation scheme to solve (1). We replace

$$0 \leq x \perp F(x) \geq 0, \quad (2)$$

by a sequence of smoothed systems of the form

$$G_r(x, F(x)) := (G_r(x_i, F_i(x)))_{i=1, \dots, n} := \left(r\psi^{-1} \left[\psi \left(\frac{x_i}{r} \right) + \psi \left(\frac{F_i(x)}{r} \right) \right] \right)_{i=1, \dots, n} = 0. \quad (3)$$

All the functions and parameters involved in (3) will be explicit later. Depending on the context, several functions (θ, ψ, G_r, \dots) we apply on reals or vectors. When applied to vectors, we consider that they apply component-wise. The novelty of our approach is that we do not need any complicated strategy to update the regularization parameter r since we will consider it as a new variable. To solve the smoothed equations system we will use the standard Newton-like method. Without requiring a strict complementarity assumption at the solution of equation (1), we prove that the proposed algorithm is well defined, globally and superlinearly convergent. At the end of the paper, we present numerical results to prove the effectiveness of the algorithm.

This paper is organized as follows: some definitions are introduced in section 2. We present our approximation and formulation in section 3. In section 4, we discuss our approach and scheme to solve (1). The convergence properties of the algorithm are given in section 5. The section 6 is devoted to the numerical results with a comparison of our method with other approaches. Finally, we conclude our paper.

2. Preliminaries and problem setting

Consider the NCP, which is to find a solution of the system:

$$x \geq 0, \quad F(x) \geq 0 \quad \text{and} \quad x^T F(x) = 0 \quad \text{or} \quad 0 \leq x \perp F(x) \geq 0, \quad (4)$$

where $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a continuous function satisfying some additional assumptions to be precised later. From (4), we obtain the equivalent formulation for component-wise products

$$x \geq 0, \quad F(x) \geq 0 \quad \text{and} \quad x_i F_i(x) = 0, \quad i = 1, 2, \dots, n.$$

Or equivalently

$$x.F(x) = 0, \quad x \geq 0, \quad F(x) \geq 0,$$

where “.” stands for the Hadamard product. It provides an explanation for the term “complementarity”, namely, for all $i = 1, 2, \dots, n$, x_i and $F_i(x)$ are complementary in the sense that if one of them is positive then the other term must be zero.

A particular and important class of NCP is the LCP class defined below.

Definition 2.1

When F is affine function:

$$F(x) = Mx + q, \quad x \in \mathbb{R}^n, \quad q \in \mathbb{R}^n, \quad M \in \mathbb{R}^{n \times n}.$$

The corresponding NCP is called an LCP. So an problem is to find $x \in \mathbb{R}^n$ such that

$$x \geq 0, \quad Mx + q \geq 0 \quad \text{and} \quad x^T (Mx + q) = 0.$$

To solve NCP, there are essentially three different classes of methods: equation-based methods (smoothing), merit functions, and projection-type methods. Our goal in this paper is to present new and very simple smoothing and approximation schemes to solve NCP and to produce efficient numerical methods. In our approach, we do not need any complicated strategy to update the smoothing parameter since we will consider it as a new variable.

First, let us introduce the usual assumptions on F and the ones that will be used in this paper. A well-known and

studied situation corresponds to monotone functions F and several methods and algorithms have been developed in this case.

Almost all the solution methods consider at least the following important and standard condition on the mapping F (monotonicity): We recall that F is said to be monotone if $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfies for any $(x, y) \in \mathbb{R}^n$,

$$(x - y)^T (F(x) - F(y)) \geq 0.$$

In this work, we will consider a weaker assumption on F :

$$F \text{ is a } \mathcal{P}_0\text{-function,} \tag{H_0}$$

to prove the convergence of our approach. We recall the following definitions of \mathcal{P}_0 and \mathcal{P} functions. We say that $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a \mathcal{P}_0 -function (respectively \mathcal{P} -function) if $\forall x, y \in \mathbb{R}^n$ with $x \neq y$, there exists an index $i_0 \in \{1, 2, \dots, n\}$ such that

$$(x_{i_0} - y_{i_0})[F_{i_0}(x) - F_{i_0}(y)] \geq 0,$$

$$\text{(respectively } (x_{i_0} - y_{i_0})[F_{i_0}(x) - F_{i_0}(y)] > 0\text{)}.$$

It is important to notice that the index i_0 can depend on x and y .

A matrix is called a \mathcal{P}_0 -matrix (resp. \mathcal{P} -matrix) if all its principal minors are nonnegative (resp. positive). Note that F is a \mathcal{P}_0 -function if and only if $\nabla F(x)$ is a \mathcal{P}_0 -matrix for all $x \in \mathbb{R}^n$. If $\nabla F(x)$ is a \mathcal{P} -matrix for all $x \in \mathbb{R}^n$, then F is a \mathcal{P} -function. However, the converse is not necessarily true.

3. Smoothing approximation functions

In this section, we present a new smoothing function for NCP. A function $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ is said to be a NCP function if ϕ satisfies

$$\phi(a, b) = 0 \iff a \geq 0, b \geq 0, ab = 0. \tag{5}$$

An example of such function is $\phi_{\min} : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$\phi_{\min}(a, b) = \min\{a, b\}.$$

Then, ϕ_{\min} is a NCP function. Problem (1) is then equivalent to the following system of nonlinear equations:

$$H(x) = \begin{pmatrix} \phi_{\min}(x_1, F_1(x)) \\ \phi_{\min}(x_2, F_2(x)) \\ \vdots \\ \phi_{\min}(x_n, F_n(x)) \end{pmatrix} = 0. \tag{6}$$

This system is clearly nonsmooth; classical Newton-like methods can not be used to try to solve it. To overcome this difficulty, there exist several semi-smooth approaches. These techniques may present difficulties to converge. An efficient approach is to approximate (6) by a smooth one. The following subsection introduces some smoothing functions and establishes different properties that will be useful for our study.

3.1. θ -smoothing

In this section, we elaborate on how such a regularized function can be actually built up from the function (6). Our smoothing technique is based on the continuous approximation of a more elementary object, namely the step function. The step function is understood here to be the function $\mathfrak{S} : \mathbb{R}_+ \rightarrow \{0, 1\}$ defined as

$$\mathfrak{S}(t) = \begin{cases} 0 & \text{if } t = 0, \\ 1 & \text{if } t > 0. \end{cases} \tag{7}$$

As an indicator of positive arguments $t > 0$ over \mathbb{R}_+ , the step function \mathfrak{S} “discriminates” the argument $t = 0$ by assigning a zero value to it. The price to be paid for this sharp detection is the discontinuity of \mathfrak{S} at $t = 0$. We wish to have a regularization of \mathfrak{S} , that is, a family of functions

$$\{\tilde{\mathfrak{S}}(\cdot, r) : \mathbb{R}_+ \rightarrow [0, 1), r > 0\}, \tag{8}$$

such that

- $\tilde{\mathfrak{S}}(\cdot, r)$ is a smooth function of $t \geq 0$, for all $r > 0$;
- $\tilde{\mathfrak{S}}$ is continuous with respect to r , in some functional sense;
- $\lim_{r \downarrow 0} \tilde{\mathfrak{S}}(\cdot, r) = \mathfrak{S}(\cdot)$, in some functional sense.

To obtain such a family, we follow the methodology developed by Haddou and his coauthors [17, 2], the key ingredient of which is a smoothing function. This notion turned out to be a versatile tool in a wide variety of pure and applied mathematical problems [3, 18, 19, 34]. We begin with a “father” function, from which all other regularized functions will be generated.

Definition 3.1

(θ -smoothing function). A function $\theta : \mathbb{R} \rightarrow [0, 1)$ is said to be a θ -smoothing function if it is continuous, nondecreasing, concave, and

$$\begin{aligned} \theta(0) &= 0, \\ \lim_{t \rightarrow +\infty} \theta(t) &= 1. \end{aligned} \tag{9}$$

The two most common examples of smoothing functions are:

1. the rational function $\theta^1 : \mathbb{R} \rightarrow (-\infty, 1)$ defined by

$$\theta^1(t) = \frac{t}{t+1} \quad \text{for } t \geq 0 \quad \text{and} \quad \theta^1(t) = t \quad \text{for } t \leq 0. \tag{10}$$

2. the exponential function $\theta^2 : \mathbb{R} \rightarrow (-\infty, 1)$ defined by

$$\theta^2(t) = 1 - \exp(-t). \tag{11}$$

A more general “recipe” to build such function is to consider nonincreasing probability density functions $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and then take the corresponding cumulative distribution function on \mathbb{R}_+ i.e.,

$$\theta(t) = \int_0^t f(y)dy, \quad t \geq 0, \tag{12}$$

we complete the definition of θ on \mathbb{R}_- by $\theta(t) = t$ to get a continuous, nondecreasing function. The nonincreasing assumption on f gives the concavity of θ . Once a favorite θ -smoothing has been selected, the next step is to dilate or compress it in order to produce a family of regularized functions for the step function \mathfrak{S} .

Definition 3.2

(θ -smoothing family). Let θ be a θ -smoothing function. The family of functions

$$\left\{ \theta_r(t) := \theta\left(\frac{t}{r}\right), \quad r > 0 \right\}, \tag{13}$$

is said to be the θ -smoothing family associated with θ .

Obviously, θ_r is a smooth function of $t \geq 0$ for all $r > 0$. It is also continuous with respect to r at each fixed $r \geq 0$. From the defining properties (9), it can be readily shown that

$$\lim_{r \rightarrow 0} \theta_r(t) = \mathfrak{S}(t), \quad \forall t \geq 0. \tag{14}$$

In other words, \mathfrak{S} is the limit of θ_r in the sense of pointwise convergence. Thus, $\{\mathfrak{S}(\cdot, r) = \theta_r, r > 0\}$ is a good family of regularized functions in the sense of (8). Associated with the two examples (10)-(11) are:

1. the rational family $\theta_r^1 : \mathbb{R} \rightarrow (-\infty, 1)$ defined by

$$\theta_r^1(t) = \frac{t}{t+r} \quad \text{for } t \geq 0 \quad \text{and} \quad \frac{t}{r} \quad \text{for } t \leq 0. \tag{15}$$

2. the exponential family $\theta_r^2 : \mathbb{R} \rightarrow (-\infty, 1)$ defined by

$$\theta_r^2(t) = 1 - \exp(-t/r). \tag{16}$$

Figure 1 display the two families (15)-(16) for a few values of the parameter r . We can see that the smaller r is, the steeper is the slope at $t = 0$ and the closer to \mathfrak{S} the function is.

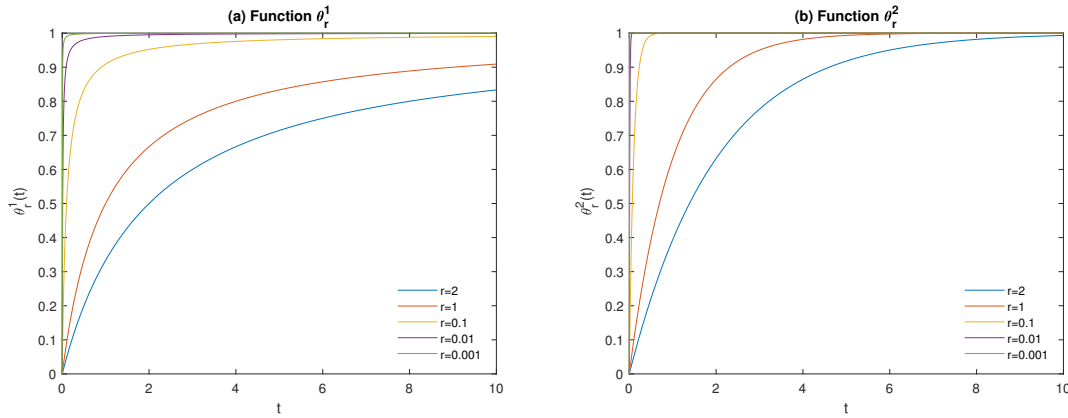


Figure 1. Function θ_r for a few values of r .

3.1.1. θ -smoothing of a complementarity condition A θ -smoothing function paves the way for a smooth approximation of a complementarity condition.

Let $(x, z) \in \mathbb{R}^2$ be two scalars such that

$$0 \leq x \perp z \geq 0, \tag{17}$$

that is,

$$x \geq 0, \quad z \geq 0, \quad xz = 0.$$

In the (x, z) -plane, the set of points obeying (17) is the union of the two semi-axes $\{x \geq 0, z = 0\}$ and $\{x = 0, z \geq 0\}$. Visually, the nonsmoothness of (17) is manifested by the “kink” at the corner $(x, z) = (0, 0)$. It is also clear that the corresponding set is nonconvex. We consider two possible smooth approximations of (17), depending how it is rewritten in terms of the step function \mathfrak{S} .

Lemma 3.1

[17] Assuming $x \geq 0$ and $z \geq 0$, we have the equivalence

$$xz = 0 \iff \mathfrak{S}(x) + \mathfrak{S}(z) \leq 1. \tag{18}$$

The equivalence (18) suggests us to impose

$$x \geq 0, \quad z \geq 0, \quad \theta_r(x) + \theta_r(z) = 1, \tag{19}$$

for $r > 0$, as a smooth approximation of (17). Replacing \mathfrak{S} by θ_r in (18) is logical. Replacing “ \leq ” by “ $=$ ” in (18) and the (19) seems to be a bold move, but this is motivated by the fact that we want an equality to be mounted into the system of equations. Some times an additional assumption (strict complementarity $x + z > 0$) is made to get such equations.

3.2. A new smoothing function using θ -function

Our aim is to propose a large class of θ -functions for which the problems

$$x^{(r)} \geq 0, \quad F(x^{(r)}) \geq 0 \quad \text{and} \quad \theta_r(x^{(r)}) + \theta_r(F(x^{(r)})) = 1, \quad (20)$$

are well-posed and any limit point of $(x^{(r)})$ when r goes to 0, is a solution of NCP. In the multidimensional case, the equation just above has to be interpreted as a system of n equations,

$$\theta_r(x_i^{(r)}) + \theta_r(F_i(x^{(r)})) = 1, \quad i = 1, \dots, n.$$

Note that the relation (20) is symmetric in x and $F(x)$. Thus, our problem can be seen as a fixed point problem for the function $F_{r,\theta}(x)$ defined just below. Indeed, the equation (20) is equivalent to

$$x = \theta_r^{-1}(1 - \theta_r(F(x))) = r\theta^{-1}(1 - \theta(F(x)/r)) =: F_{r,\theta}(x).$$

By symmetry of the equation (20), we also have the relations:

$$F(x) = \theta_r^{-1}(1 - \theta_r(x)) = r\theta^{-1}(1 - \theta(x/r)),$$

but we shall not go that direction. As in [18] we propose another way to approximate a solution of NCP as follows. Let $\psi_r(t) = 1 - \theta_r(t)$, the relation (20) is equivalent to the three following equalities

$$\psi_r(x) + \psi_r(F(x)) = 1 = \psi_r(0),$$

$$\psi_r^{-1}[\psi_r(x) + \psi_r(F(x))] = 0 \quad \text{and}$$

$$r\psi^{-1}\left[\psi\left(\frac{x}{r}\right) + \psi\left(\frac{F(x)}{r}\right)\right] = 0.$$

(with $\psi = 1 - \theta$). For the sequel, we set for any $x, y \in \mathbb{R}^n$ and any $r > 0$

$$G_r(x, y) := (G_r(x_i, y_i))_{i=1, \dots, n} := \left(r\psi^{-1}\left[\psi\left(\frac{x_i}{r}\right) + \psi\left(\frac{y_i}{r}\right)\right] \right)_{i=1, \dots, n}. \quad (21)$$

where $\psi : \mathbb{R} \rightarrow]0, +\infty[$.

First, we characterize the solutions (x, y) of $G_r(x, y) = 0$ when ψ satisfies some conditions independent of F .

Let $0 < a < 1$. We say that ψ satisfies condition (H_a) if there exists $s_a > 0$ such that

$$\psi(s) \leq \frac{1}{2}\psi(as) \quad \forall s \geq s_a \quad \text{or equivalently} \quad \frac{1}{2} + \frac{1}{2}\theta(as) \leq \theta(s) \quad \forall s \geq s_a. \quad (H_a)$$

The condition (H_a) imposes that the decay of $\psi(s)$ is under some uniform control for large s or in terms of θ that $\theta(s)$ should grow enough quickly with some uniformity for large s . Since ψ and θ are monotone, it is interesting to take a as large as possible in the condition (H_a) since $(H_a) \implies (H_b)$ for $b < a$.

Note that we can never take $a = 1$ because $\theta \leq 1$ unless θ is constant and equal to one for large s . But in some cases, a can be chosen as close to 1, see for instance θ^2 .

One can obtain by simple calculations that:

1. For θ^1 , we have

$$\psi^1(t) = \begin{cases} \frac{1}{t+1} & \text{if } t \geq 0, \\ 1-t & \text{if } t < 0, \end{cases}$$

and the condition (H_a) is only satisfied for $0 < a < 1/2$ with $s_a \geq \frac{1}{1-2a}$.

- For θ^2 , we have $\psi(t) = e^{-t}$ and the condition (H_a) is satisfied for any $0 < a < 1$ with $s_a = \frac{\ln 2}{1 - a}$.

From now on, all the results use the function ψ . Obviously, everything can be easily transposed on θ . We summarize here some important facts proved in [18]. The following lemma compare the function G_r defined in (21) to the min function and will be useful for the rest of our analysis.

Lemma 3.2

If $\psi : \mathbb{R} \rightarrow]0, +\infty[$ is an invertible non-increasing function, then for any $(s, t) \in \mathbb{R}^2$ and any $r > 0$

$$G_r(s, t) \leq \min(s, t).$$

The next theorem shows how the condition (H_a) gives information about the behavior of G_r .

Theorem 3.1

Let $\psi : \mathbb{R} \rightarrow]0, +\infty[$ be an invertible non-increasing function such that

$$\lim_{t \rightarrow -\infty} \psi(t) = +\infty, \psi(0) = 1, \text{ and } \lim_{t \rightarrow +\infty} \psi(t) = 0.$$

If ψ satisfies the condition (H_a) for some $a \in]0, 1[$, then for all $s, t \in \mathbb{R}$,

$$\lim_{r \searrow 0} G_r(s, t) = 0 \iff \min(s, t) = 0.$$

For both θ_r^1 and θ_r^2 examples, the assertion of Theorem 3.1 is satisfied. Indeed direct computations lead to

- For $s > 0$ and $t > 0$ such that $\frac{1}{s} + \frac{1}{t} \leq \frac{1}{r}$, we have the following explicit expression

$$G_r^1(s, t) = \frac{st - r^2}{s + t + 2r}. \tag{22}$$

Note that the denominator is not zero when s, t are non-negative even when $s = t = 0$. In addition, when $\min(s, t) > 0$ we have $\lim_{r \searrow 0} G_r^1(s, t) = \frac{st}{s + t} < \min(s, t)$.

- For any $s, t \in \mathbb{R}$, we have the following explicit expression

$$G_r^2(s, t) = -r \log(e^{-s/r} + e^{-t/r}). \tag{23}$$

Assume $s = \min(s, t)$. Then we have $s - r \log(2) \leq G_r^2(s, t)$ because

$$e^{-s/r} + e^{-t/r} \leq 2e^{-s/r}.$$

Thus, $\min(s, t) - r \log(2) \leq G_r^2(s, t) \leq \min(s, t)$. Passing to the limit as r goes to 0, we conclude that $\lim_{r \searrow 0} G_r^2(s, t) = \min(s, t)$.

Figure 2 illustrate the behaviour of $G_r^2(x, -x)$.

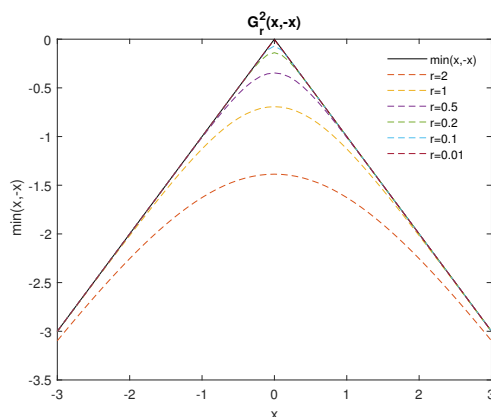


Figure 2. Comparison of $G_r^2(x, -x)$ and $\min(x, -x)$.

Now, we focus on the case where ψ satisfies (H_a) for all $a \in]0, 1[$ and state a stronger result (see [18]).

Theorem 3.2

Let $\psi : \mathbb{R} \rightarrow]0, +\infty[$ be an invertible non-increasing function such that

$$\lim_{t \rightarrow -\infty} \psi(t) = +\infty, \psi(0) = 1, \text{ and } \lim_{t \rightarrow +\infty} \psi(t) = 0.$$

If ψ satisfies (H_a) for all $a \in]0, 1[$, then for any $s, t > 0$,

$$\lim_{r \searrow 0} G_r(s, t) = \min(s, t).$$

3.3. An approximate formulation

In this section, we present two new reformulations of the complementarity problem (4) by corresponding to two approximation schemes.

3.3.1. *Approximation of NCP using θ_r^1 -function and θ_r^2 -function* Using θ_r -function, we regularize each complementarity constraint by considering

$$x_i z_i = 0, x_i \geq 0, z_i \geq 0 \text{ by } G_r(x_i, z_i) := 0, \quad i = 1, \dots, n.$$

This approximation yields the following formulation

$$(\tilde{P}_\theta) \quad \begin{cases} F(x) = z, \\ x \geq 0, z \geq 0, \quad r \searrow 0, \\ G_r(x, z) = 0. \end{cases} \tag{24}$$

We consider the family $\{H_\theta^r(\cdot), r > 0\}$, where

$$H_\theta^r(x, z) = \begin{bmatrix} F(x) - z \\ G_r(x, z) \end{bmatrix}, \tag{25}$$

is a regularized function of H defined in (6).

Lemma 3.3

Let $H_{\theta_1}^r(x, z)$ and $H_{\theta_2}^r(x, z)$ be defined by (25) respectively for θ_1^1 and θ_2^2 . Then, for any $(x, z) \in \mathbb{R}_+^{2n}$ the Jacobian

matrix of $H_{\theta_1}^r(x, z)$ is

$$\nabla H_{\theta_1}^r(x, z) = \begin{pmatrix} \nabla F(x) & -I \\ D_a(x, z) & D_b(x, z) \end{pmatrix},$$

where $D_a(x, z) = \text{diag}\{a_1(x, z), \dots, a_n(x, z)\}$ and $D_b(x, z) = \text{diag}\{b_1(x, z), \dots, b_n(x, z)\}$ are two diagonal matrices, given by

$$a_i(x, z) = \left(\frac{z_i + r}{x_i + z_i + 2r}\right)^2, \quad b_i(x, z) = \left(\frac{x_i + r}{x_i + z_i + 2r}\right)^2, \quad i = 1, \dots, n.$$

and, for any $(x, z) \in \mathbb{R}_+^{2n}$ the Jacobian matrix of $H_{\theta_2}^r(x, z)$ is

$$\nabla H_{\theta_2}^r(x, z) = \begin{pmatrix} \nabla F(x) & -I \\ Q_k(x, z) & Q_l(x, z) \end{pmatrix},$$

where $Q_k(x, z) = \text{diag}\{k_1(x, z), \dots, k_n(x, z)\}$ and $Q_l(x, z) = \text{diag}\{l_1(x, z), \dots, l_n(x, z)\}$ are two diagonal matrices, given by

$$k_i(x, z) = \frac{e^{-x_i/r}}{e^{-x_i/r} + e^{-z_i/r}}, \quad l_i(x, z) = \frac{e^{-z_i/r}}{e^{-x_i/r} + e^{-z_i/r}}, \quad i = 1, \dots, n.$$

In order to study the nonsingularity of the Jacobian matrix of $H_{\theta_1}^r(x, z)$ (resp. $H_{\theta_2}^r(x, z)$), we state first a basic but essential lemma.

Lemma 3.4

Let $M \in \mathbb{R}^{n \times n}$ be a \mathcal{P}_0 -matrix. Then any matrix in the following form is nonsingular:

$$N_s + N_t M,$$

where $N_s \in \mathbb{R}^{n \times n}$ is a positive (negative) diagonal matrix, and $N_t \in \mathbb{R}^{n \times n}$ is a nonnegative (non-positive) diagonal matrix.

Proof

Let $N_s = \text{diag}(s_1, s_2, \dots, s_n)$ and $N_t = \text{diag}(t_1, t_2, \dots, t_n)$. If N_s is positive, and N_t is nonnegative, then $s_i > 0$ and $t_i \geq 0$ for all $i = 1, 2, \dots, n$.

Let $v \in \mathbb{R}^n$ be a vector such that $(N_s + N_t M)v = 0$. Then, we have $v_i = -\frac{t_i}{s_i}(Mv)_i$.

It yields $v_i^2 = -\frac{t_i}{s_i}v_i(Mv)_i$. If $t_i = 0$, then $v_i = 0, \forall i = 1, \dots, n$.

If $v_i \neq 0$, we have $\frac{t_i}{s_i} > 0$. Owing to $v_i^2 \geq 0$, we have $v_i(Mv)_i \leq 0$. If $v_i(Mv)_i = 0$, then $v_i = 0$. Otherwise, $v_i(Mv)_i < 0$ contradicts the property of M . Based on the above discussion, it is concluded that $v = 0$, then $N_s + N_t M$ is a nonsingular matrix. □

By Lemma 3.4, we can obtain a property of $H_{\theta_1}^r$ and $H_{\theta_2}^r$ if F is a \mathcal{P}_0 -matrix.

Theorem 3.3

Let F be a \mathcal{P}_0 -function. Then, for any $r > 0$, and any $(x, z) \in \mathbb{R}_+^{2n}$ the Jacobian matrix $\nabla H_{\theta_1}^r(x, z)$ (resp. $\nabla H_{\theta_2}^r(x, z)$) is nonsingular.

Proof

For all $r > 0$ and from Lemma 3.3, it follows that the diagonal matrix $D_a(x, z)$ (resp. $Q_k(x, z)$) is non-negative, and $D_b(x, z)$ (resp. $Q_l(x, z)$) is non-negative diagonal matrix.

Since F is a \mathcal{P}_0 -function, the Jacobian matrix $\nabla F(x)$ is a \mathcal{P}_0 -matrix.

We have

$$\det(\nabla H_{\theta_1}^r(x, z)) = \det(D_a(x, z) + \nabla F(x)D_b(x, z)),$$

and

$$\det(\nabla H_{\theta_2}^r(x, z)) = \det(Q_k(x, z) + \nabla F(x)Q_l(x, z)).$$

Since $\nabla F(x)$ is a \mathcal{P}_0 -matrix and from Lemma 3.4, it follows that $D_a(x, z) + \nabla F(x)D_b(x, z)$ (resp. $Q_k(x, z) + \nabla F(x)Q_l(x, z)$) is nonsingular. Hence $\nabla H_{\theta_1}^r(x, z)$ (resp. $\nabla H_{\theta_2}^r$) is nonsingular. \square

4. New approach for solving nonlinear complementarity problems

In this section, we present the idea of our algorithms, we take inspiration from the well-known interior-point methods (IPMs) usually used in nonlinear programming. Even though we don't have any objective function to minimize, the regularization idea behind IPM can be used to tackle NCP.

One can replace the original nonsmooth problems NCPs by a sequence of regularized problems

$$H_r(\mathbf{X}) = 0, \quad (26)$$

where

$$\mathbf{X} = \begin{bmatrix} x \\ z \end{bmatrix} \in \mathbb{R}_+^{2n}, \quad H_r(\mathbf{X}) = \begin{bmatrix} F(x) - z \\ x \cdot z - r \mathbf{e} \end{bmatrix}, \quad (27)$$

and $r \geq 0$ is the smoothing parameter, $\mathbf{e} \in \mathbb{R}^n$ is the vector whose components are all equal to 1. The Jacobian matrix of H_r with respect to \mathbf{X} , does not depend on r and can be denoted by

$$\nabla H_r(\mathbf{X}) = \begin{pmatrix} \nabla F(x) & -I \\ Z & X \end{pmatrix}, \quad (28)$$

where Z and (resp. X) the diagonal matrix of z (resp. x).

The main difficulty in this approach, is to drive r to 0. In [36] the authors propose a new technique where r is considered as a new variable.

4.1. When the parameter becomes a variable

In the system (26), the status of the parameter r is very distinct from that of the variable \mathbf{X} . While \mathbf{X} is computed "automatically" by a Newton iteration, r has to be updated "manually" in an ad-hoc manner.

Our goal is to find a strategy that decreases r during iterations and ensures the nonnegative of variables. However, we must adjust the strategy when the model or its parameters are changed. To avoid this trouble, we consider r as an unknown of the system instead of a parameter as in [36].

We feel that it would be judicious to incorporate the parameter r into the variables. Let us, therefore, consider the enlarged vector of unknowns

$$\mathbb{X} = \begin{bmatrix} \mathbf{X} \\ r \end{bmatrix} \in \mathbb{R}_+^{2n} \times \mathbb{R}_+, \quad (29)$$

and then consider a system of $2n + 1$ equations

$$\mathbb{H}_\theta(\mathbb{X}) = 0, \quad (30)$$

to be on \mathbb{X} . To this end, let us remind ourselves that our ultimate goal is to solve $H_{\theta_1}^0(\mathbf{X})$ (we restrict our choice of θ -function to $\theta_r(t) = \theta_r^1(t)$), together with the inequalities $x \geq 0$, $z \geq 0$.

Thus, it is really natural to first consider

$$\mathbb{H}_\theta(\mathbb{X}) = \begin{bmatrix} H_{\theta_1}^r(\mathbf{X}) \\ r \end{bmatrix}. \quad (31)$$

This construction turns out to be too naive. Indeed, if we start from some r^0 and solve the smooth system (31) by the smooth Newton method since the last equation is linear, we end up with $r^1 = 0$ at the first iteration. Once the boundary of the interior region is reached, we are "stuck" there.

To prevent r from rushing to zero in just one iteration, we could set

$$\mathbb{H}_\theta(\mathbb{X}) = \begin{bmatrix} H_{\theta_1}^r(\mathbf{X}) \\ r^2 \end{bmatrix}. \tag{32}$$

At this stage, system (32) is not yet fully adequate. Indeed, the last equation is decoupled from the others. Everything happens as if r follows a prefixed sequence, generated by the Newton iterates of the scalar equation $r^2 = 0$, regardless of \mathbf{X} . It is desirable to a couple r and \mathbf{X} in a tighter way. In this respect, we advocate

$$\mathbb{H}_\theta(\mathbb{X}) = \begin{bmatrix} H_{\theta_1}^r(\mathbf{X}) \\ \frac{1}{2}\|x^-\|^2 + \frac{1}{2}\|z^-\|^2 + r^2 \end{bmatrix}, \tag{33}$$

where x^- is the vector of components $x_i^- = \min(x_i, 0)$ and similarly for z^- and

$$\|x^-\|^2 = \sum_{i=1}^n \min^2(x_i, 0), \quad \|z^-\|^2 = \sum_{i=1}^n \min^2(z_i, 0).$$

This choice has the benefit of taking into account the nonnegativity condition $x \geq 0$ and $z \geq 0$.

Indeed, the last equation of (33) implies that, as long as $r \geq 0$, we are ascertained that $x^- = z^- = 0$. This amounts to saying that $x \geq 0$ and $z \geq 0$. Should a component of x or z become negative during the iteration, this equation would contribute to “penalize” it.

Since r is now considered a variable and the scalar function $t \mapsto \frac{1}{2}|\min(t, 0)|^2$ is differentiable and its derivative is equal to $\min(t, 0)$. From this observation, the Jacobian matrix of \mathbb{H}_θ is:

$$\nabla_{\mathbb{X}}\mathbb{H}_\theta(\mathbb{X}) = \begin{pmatrix} \nabla_x H_{\theta_1}^r & \nabla_z H_{\theta_1}^r & \partial_r H_{\theta_1}^r \\ (x^-)^T & (z^-)^T & 2r \end{pmatrix}, \tag{34}$$

where

$$\begin{aligned} \nabla_x H_{\theta_1}^r &= \begin{pmatrix} \nabla_x F(x) \\ D_a(x, z) \end{pmatrix}_{2n \times n}, & \nabla_z H_{\theta_1}^r &= \begin{pmatrix} -I \\ D_b(x, z) \end{pmatrix}_{2n \times n}, \\ \partial_r H_{\theta_1}^r &= \begin{bmatrix} 0_{n \times 1} \\ \text{diag} \left(\left(\frac{-2r}{x_i + z_i + 2r} + \frac{2(r^2 - x_i z_i)}{(x_i + z_i + 2r)^2} \right)_{1 \leq i \leq n} \right) \mathbf{e} \end{bmatrix}_{2n \times 1}, \end{aligned}$$

If $\mathbb{H}_\theta(\mathbb{X}) = 0$ where $\mathbb{X} \in \mathbb{R}_+^{2n} \times \mathbb{R}_+$ we obtain $r = 0$ and $x^- = z^- = 0$. Hence in this case, $\nabla_{\mathbb{X}}\mathbb{H}_\theta(\mathbb{X})$ becomes singular, since $\det(\nabla_{\mathbb{X}}\mathbb{H}_\theta(\mathbb{X})) = 0$. To solve this issue, we add a small enough positive parameter ε in the last equation. We get

$$\frac{1}{2}\|x^-\|^2 + \frac{1}{2}\|z^-\|^2 + r^2 + \varepsilon r = 0. \tag{35}$$

Hence, we define the following systems

$$\mathbb{H}_\theta(\mathbb{X}) = \begin{bmatrix} H_{\theta_1}^r(\mathbf{X}) \\ \frac{1}{2}\|x^-\|^2 + \frac{1}{2}\|z^-\|^2 + r^2 + \varepsilon r \end{bmatrix}. \tag{36}$$

Lemma 4.1

Let $\mathbf{X} \in \bar{\Xi}$ (the closure of Ξ), where Ξ is the interior region defined in

$$\Xi = \{\mathbf{X} = (x, z) \in \mathbb{R}^{2n} \mid x > 0, z > 0\}. \tag{37}$$

Let $r \in \mathbb{R}$ and $\mathbb{X} = [\mathbf{X}; r]^T$. Then,

$$\det \nabla \mathbb{H}_\theta(\mathbb{X}) = (\varepsilon + 2r) \det \nabla H_{\theta_1}^r(\mathbf{X}).$$

If $r > -\frac{\varepsilon}{2}$, the two Jacobian matrices are singular or nonsingular at the same time.

Proof

Thanks to the assumption $\mathbf{X} \in \bar{\Xi}$, we have $x \geq 0$ and $z \geq 0$, so that $x^- = z^- = 0$. Expanding the determinant of $\nabla \mathbb{H}_\theta(\mathbb{X})$ with respect to the last row yields the desired result. \square

5. Convergence

In this section, we propose a generic algorithm to solve NCPs and prove some convergence results. From now on, the enlarged equation (36) is selected as the reference system in the design of our new algorithm. The idea is simply to apply the standard Newton method to the smooth system (36). To enforce a global convergence behavior, we also recommend using α line search like Armijo back-tracking technique. Now, we present our algorithm:

<p>Algorithm 1 Nonparametric method with Armijo line search</p> <ol style="list-style-type: none"> 1. Chose $\mathbb{X}^0 = (\mathbf{X}^0, r^0)$, $\mathbf{X}^0 \in \Xi$, $r^0 = \langle x^0, z^0 \rangle / n$, $\tau \in (0, 1/2)$, $\varrho \in (0, 1)$. Set $k = 0$. 2. If $\mathbb{H}_\theta(\mathbb{X}^k) = 0$, stop. 3. Find a direction $\mathbf{d}^k \in \mathbb{R}^{2n+1}$ such that $\mathbb{H}_\theta(\mathbb{X}^k) + \nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbb{X}^k) \mathbf{d}^k = 0.$ 4. Choose $\zeta^k = \varrho^{j_k} \in (0, 1)$, where $j_k \in \mathbb{N}$ is the smallest integer such that $\Theta(\mathbb{X}^k + \zeta^k \mathbf{d}^k) - \Theta(\mathbb{X}^k) \leq \tau \varrho^{j_k} \nabla \Theta(\mathbb{X}^k)^T \mathbf{d}^k.$ 5. Set $\mathbb{X}^{k+1} = \mathbb{X}^k + \zeta^k \mathbf{d}^k$ and $k \leftarrow k + 1$. Go to step 2.

The merit function used in the line search is:

$$\Theta(\mathbb{X}) = \frac{1}{2} \|\mathbb{H}_\theta(\mathbb{X})\|^2.$$

A detailed description of nonparametric method is given in Algorithm 1. A few comments are in order:

- The initial point $\mathbb{X}^0 = (\mathbf{X}^0, r^0)$ must be an interior point, namely, $\mathbf{X}^0 > 0$ and the initial parameter $r^0 = \langle x^0, z^0 \rangle / n$ has the correct order of magnitude.
- If $\mathbf{X}^k \in \Xi$, then $(x^k)^- = (z^k)^- = 0$ and

$$\mathbf{d}^k = \begin{bmatrix} d\mathbf{X}^k \\ dr^k \end{bmatrix} = - \begin{pmatrix} \nabla_x H_{\theta_1}^r & \nabla_z H_{\theta_1}^r & \partial_r H_{\theta_1}^r \\ 0^T & 0^T & \varepsilon + 2r^k \end{pmatrix}^{-1} \begin{bmatrix} H_{\theta_1}^r(\mathbf{X}^k) \\ \varepsilon r^k + (r^k)^2 \end{bmatrix},$$

provided that the Jacobian matrix is invertible. The increment for the parameter is then

$$dr^k = - \frac{\varepsilon r^k + (r^k)^2}{\varepsilon + 2r^k}.$$

- There is no need to truncate the Newton direction \mathbf{d}^k to preserve positivity for x^{k+1} and z^{k+1} , since nonnegativity is “guaranteed” at convergence. However, if we want all the iterates to be nonnegative, so we need to carry out additional damping after Step 4 (Armijo’s line search).

Proposition 5.1

Let F be a continuous differentiable \mathcal{P}_0 -function. Then, step 3 in Algorithm 1 is well-defined.

Proof

We know that for all $k \geq 0$, $r^k > 0$, $\mathbf{X}^k > 0$, and $\varepsilon > 0$,

$$\det \nabla \mathbb{H}_\theta(\mathbb{X}^k) = (\varepsilon + 2r^k) \det \nabla H_{\theta_1}^r(\mathbf{X}^k).$$

By Theorem 3.3, $\nabla H_{\theta_1}^{r^k}(\mathbf{X}^k)$ is nonsingular, so step 3 of Algorithm 1 is well-defined. \square

5.1. Global convergence analysis

Definition 5.1

(Regular zero). Let $\mathbb{X}^* \in \mathbb{R}^{2n+1}$ be a zero of \mathbb{H}_θ , that is, $\mathbb{H}_\theta(\mathbb{X}^*) = 0$. If the Jacobian matrix $\nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbb{X}^*)$ is nonsingular, \mathbb{X}^* is said to be a regular zero of \mathbb{H}_θ .

The main interest of Algorithm 1 lies in the prospect of global convergence, as envisioned by the theory that we are developing now. This global convergence theory, is primarily based on the regularity of zeros [Definition (5.1)]. We reproduce a concise result that can be found in the book of Bonnans [5], in view of its importance to our algorithm.

We will prove the global convergence of Algorithm 1. First, we show that every $\mathbf{d} \in \Delta$ is a descent direction of Θ at \mathbb{X} , where

$$\Delta(\mathbb{X}) = \{\mathbf{d} \in \mathbb{R}^n \mid \nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbb{X})\mathbf{d} = -\mathbb{H}_\theta(\mathbb{X})\}. \quad (38)$$

Lemma 5.1

(see [37]) If \mathbb{X} is not a solution of NCP, i.e. $\Theta(\mathbb{X}) > 0$, then every $\mathbf{d} \in \Delta(\mathbb{X})$ satisfies the descent condition for \mathbb{X} , i.e., $\nabla \Theta(\mathbb{X})^T \mathbf{d} < 0$.

Using the preceding results, we can prove the following global convergence theorem.

Theorem 5.1

Every limit point $\mathbb{X}^* = (\mathbf{X}^*, r^*)$ of a sequence $\{\mathbb{X}^k\}$ generated by Algorithm 1 corresponds to a solution of NCP.

Proof

Since the sequence $\{\Theta(\mathbb{X}^k)\}$ is nonnegative and decreases monotonically, it converges to some $\Theta^* \geq 0$. We assume $\Theta^* > 0$. Let \mathbb{X}^* be an accumulation point of $\{\mathbb{X}^k\}$ and $\{\mathbb{X}^k\}_{k \in K}$ be a subsequence converging to $\{\mathbb{X}^*\}$. Taking a further subsequence if necessary, we can assume without loss of generality that $\lim_{k \rightarrow \infty} d^k = d^*$, because Δ is uniformly compact near and closed at \mathbb{X}^* (see [37]). Furthermore, by the closedness of Δ , we have

$$d^* \in \Delta(\mathbb{X}^*). \quad (39)$$

Since $\Theta(\mathbb{X}^k + \zeta^k \mathbf{d}^k) - \Theta(\mathbb{X}^k) \leq \zeta^k \tau \nabla \Theta(\mathbb{X}^k)^T \mathbf{d}^k \leq 0$. It is obvious that $\{\zeta^k \tau \nabla \Theta(\mathbb{X}^k)^T \mathbf{d}^k\}$ converges to 0. In order to prove $\nabla \Theta(\mathbb{X}^k)^T \mathbf{d}^k \rightarrow 0$, we show that $\{\zeta^k\}$ is bounded away from 0. Now suppose that there exists a subsequence such that $\zeta^k \rightarrow 0$. By the line search rule, we have

$$\frac{\Theta(\mathbb{X}^k + \sigma^k \mathbf{d}^k) - \Theta(\mathbb{X}^k)}{\sigma^k} > \tau \nabla \Theta(\mathbb{X}^k)^T \mathbf{d}^k, \quad (40)$$

where $\sigma^k = \frac{\zeta^k}{\rho^{\beta k}}$. Since $\sigma^k \rightarrow 0$, taking the limit of both sides of (40) yields

$$\nabla \Theta(\mathbb{X}^*)^T \mathbf{d}^* > \tau \nabla \Theta(\mathbb{X}^*)^T \mathbf{d}^*. \quad (41)$$

Since $\Theta(\mathbb{X}^*) = \Theta^* > 0$ by assumption, it follows from (39) and Lemma 5.1 that $\nabla \Theta(\mathbb{X}^*)^T \mathbf{d}^* < 0$. Since $\tau < 1$, this contradicts (41). This implies that $\{\zeta^k\}$ is bounded away from 0, and hence, $\{\nabla \Theta(\mathbb{X}^k)^T \mathbf{d}^k\}$ converge to 0. That is,

$$\lim_{k \rightarrow \infty} \nabla \Theta(\mathbb{X}^k)^T \mathbf{d}^k = \nabla \Theta(\mathbb{X}^*)^T \mathbf{d}^* = 0. \quad (42)$$

It then follows from (39) and Lemma 5.1 that $\Theta(\mathbb{X}^*) = 0$. This is contradictory to $\Theta^* > 0$. Therefore, we must have $\Theta(\mathbb{X}^k) \rightarrow 0$, which implies that any accumulation point \mathbb{X}^* of $\{\mathbb{X}^k\}$ satisfies $\Theta(\mathbb{X}^*) = 0$ and hence is a solution of NCP. The proof is complete. \square

Below is a result about the Jacobian matrix of $\mathbb{H}_\theta(\mathbb{X})$, when r goes to 0.

Lemma 5.2

Suppose that $\mathbf{X}^* = (x^*, z^*)$ is a solution of NCP, then we have the following equality

$$\liminf_{r \rightarrow 0} \max \left((\nabla_x G_r^1(x^*, z^*))_{ii}, (\nabla_z G_r^1(x^*, z^*))_{ii} \right) \geq \frac{1}{4}, \quad \forall i = 1, \dots, n,$$

where

$$\nabla_x G_r^1(x^*, z^*) = \text{diag} \left(\left(\left(\frac{z_i^* + r}{x_i^* + z_i^* + 2r} \right)^2 \right)_{1 \leq i \leq n} \right), \quad \text{and} \quad \nabla_z G_r^1(x^*, z^*) = \text{diag} \left(\left(\left(\frac{x_i^* + r}{x_i^* + z_i^* + 2r} \right)^2 \right)_{1 \leq i \leq n} \right).$$

Proof

By considering the two possible situations $(x_i^* \geq z_i^*)$ and $(x_i^* < z_i^*)$, very simple calculation give:

1. $x_i^* \geq z_i^*$
 $\max \left(\left(\frac{z_i^* + r}{x_i^* + z_i^* + 2r} \right)^2, \left(\frac{x_i^* + r}{x_i^* + z_i^* + 2r} \right)^2 \right) = \left(\frac{x_i^* + r}{x_i^* + z_i^* + 2r} \right)^2 \geq \left(\frac{x_i^* + r}{2x_i^* + 2r} \right)^2 = \frac{1}{4}.$
2. $z_i^* > x_i^*$
 $\max \left(\left(\frac{z_i^* + r}{x_i^* + z_i^* + 2r} \right)^2, \left(\frac{x_i^* + r}{x_i^* + z_i^* + 2r} \right)^2 \right) = \left(\frac{z_i^* + r}{x_i^* + z_i^* + 2r} \right)^2 \geq \left(\frac{z_i^* + r}{2z_i^* + 2r} \right)^2 = \frac{1}{4}.$

Since $\nabla_x G$ and $\nabla_z G$ are bounded, the proof is complete by passing to the limit. □

Lemma 5.3

Let $\mathbf{X}^* = (x^*, z^*)$ be a solution of NCP satisfying the strict complementarity condition (i.e. $x_i^* + z_i^* > 0, \forall i \in \{1, \dots, n\}$). We have

$$\lim_{r \rightarrow 0} (\nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbf{X}^*, r)) = \begin{pmatrix} \nabla F(x^*) & -I_{n \times n} & 0_{n \times 1} \\ \phi(Z^*) & \phi(X^*) & 0_{n \times 1} \\ 0_{1 \times n} & 0_{1 \times n} & \varepsilon \end{pmatrix},$$

where

$$\phi(Z^*)_{ii} = \begin{cases} 1 & \text{if } z_i^* \neq 0 \text{ and } x_i^* = 0 \\ 0 & \text{if } z_i^* = 0 \text{ and } x_i^* \neq 0, \end{cases} \quad \text{and} \quad \phi(X^*)_{ii} = \begin{cases} 1 & \text{if } x_i^* \neq 0 \text{ and } z_i^* = 0 \\ 0 & \text{if } x_i^* = 0 \text{ and } z_i^* \neq 0. \end{cases}$$

Proof

by definition

$$\mathbb{H}_\theta(\mathbb{X}) = \begin{bmatrix} F(x) - z \\ G_r^1(x, z) \\ \frac{1}{2} \|x^-\|^2 + \frac{1}{2} \|z^-\|^2 + r^2 + \varepsilon r \end{bmatrix}.$$

The Jacobian matrix of \mathbb{H}_θ is:

$$\nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbb{X}) = \begin{pmatrix} \nabla_x F(x) & -I_{n \times n} & 0_{n \times 1} \\ \nabla_x G_r^1(x, z) & \nabla_z G_r^1(x, z) & \partial_r G_r^1(x, z) \\ (x^-)^T & (z^-)^T & 2r + \varepsilon \end{pmatrix}.$$

1. The derivative of $G_r^1(x, z)$ with respect to x is:

$$\nabla_x G_r^1(x^*, z^*) = \text{diag} \left(\left(\left(\frac{z_i^* + r}{x_i^* + z_i^* + 2r} \right)^2 \right)_{1 \leq i \leq n} \right),$$

when r goes to 0 the only two cases to consider are:

- $x_i^* \rightarrow 0$, and $z_i^* > 0 \forall i \in \{1, \dots, n\}$ then

$$\lim_{\substack{r \rightarrow 0 \\ x_i^* \rightarrow 0 \\ z_i^* \neq 0}} (\nabla_x G_r^1(x^*, z^*))_{ii} = \lim_{r \rightarrow 0} \left(\frac{z_i^* + r}{z_i^* + 2r} \right)^2 = 1.$$
- $x_i^* > 0$, and $z_i^* \rightarrow 0 \forall i \in \{1, \dots, n\}$ then

$$\lim_{\substack{r \rightarrow 0 \\ z_i^* \rightarrow 0 \\ x_i^* \neq 0}} (\nabla_x G_r^1(x^*, z^*))_{ii} = \lim_{r \rightarrow 0} \left(\frac{r}{x_i^* + 2r} \right)^2 = 0.$$

2. The derivative of $G_r^1(x, z)$ with respect to z is:

$$\nabla_z G_r^1(x^*, z^*) = \text{diag} \left(\left(\left(\frac{x_i^* + r}{x_i^* + z_i^* + 2r} \right)^2 \right)_{1 \leq i \leq n} \right),$$

as below, the only two cases to consider are:

- $x_i^* \rightarrow 0$, and $z_i^* > 0 \forall i \in \{1, \dots, n\}$ then

$$\lim_{\substack{r \rightarrow 0 \\ x_i^* \rightarrow 0 \\ z_i^* \neq 0}} (\nabla_z G_r^1(x^*, z^*))_{ii} = \lim_{r \rightarrow 0} \left(\frac{r}{z_i^* + 2r} \right)^2 = 0.$$
- $x_i^* > 0$, and $z_i^* \rightarrow 0 \forall i \in \{1, \dots, n\}$ then

$$\lim_{\substack{r \rightarrow 0 \\ z_i^* \rightarrow 0 \\ x_i^* \neq 0}} (\nabla_z G_r^1(x^*, z^*))_{ii} = \lim_{r \rightarrow 0} \left(\frac{x_i^* + r}{x_i^* + 2r} \right)^2 = 1.$$

3. The derivative of $G_r^1(x, z)$ with respect to r is:

$$\partial_r G_r^1(x^*, z^*) = \left(\frac{-2r}{x_i^* + z_i^* + 2r} + \frac{2(r^2 - x_i^* z_i^*)}{(x_i^* + z_i^* + 2r)^2} \right)_{i=1, \dots, n},$$

when r goes to 0 the only two cases to consider are:

- $x_i^* \rightarrow 0$, and $z_i^* > 0 \forall i \in \{1, \dots, n\}$ then

$$\lim_{\substack{r \rightarrow 0 \\ x_i^* \rightarrow 0 \\ z_i^* \neq 0}} (\partial_r G_r^1(x^*, z^*))_i = \lim_{r \rightarrow 0} \left(\frac{-2r}{z_i^* + 2r} + \frac{2r^2}{(z_i^* + 2r)^2} \right) = 0.$$
- $x_i^* > 0$, and $z_i^* \rightarrow 0 \forall i \in \{1, \dots, n\}$ then

$$\lim_{\substack{r \rightarrow 0 \\ z_i^* \rightarrow 0 \\ x_i^* \neq 0}} (\partial_r G_r^1(x^*, z^*))_i = \lim_{r \rightarrow 0} \left(\frac{-2r}{x_i^* + 2r} + \frac{2r^2}{(x_i^* + 2r)^2} \right) = 0.$$

□

Finally, since $\mathbf{X}^* = (x^*, z^*)$ is a solution of NCP, we have $x^* \geq 0$ and $z^* \geq 0$, so that $x^- = z^- = 0$. Hence

$$\lim_{r \rightarrow 0} (\nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbf{X}^*, r)) = \begin{pmatrix} \nabla F(x^*) & -I_{n \times n} & 0_{n \times 1} \\ \phi(Z^*) & \phi(X^*) & 0_{n \times 1} \\ 0_{1 \times n} & 0_{1 \times n} & \varepsilon \end{pmatrix}.$$

We present now, two situations where we can conclude about the nonsingularity of $\lim_{r \rightarrow 0} \nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbf{X}^*, r)$.

Lemma 5.4

Suppose that $\mathbf{X}^* = (x^*, z^*)$ is a solution of NCP, we have two possibilities when computing the determinant of \mathbb{H}_θ on (x^*, z^*) .

- If $\mathbf{X}^* = (x^*, z^*)$ satisfies the strict complementarity condition, then by Lemma 5.6, $\forall \mathbb{I} \subset \{1, \dots, n\}$ we have:

$$\lim_{r \rightarrow 0} \det(\nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbf{X}^*, r)) = \left| \begin{pmatrix} \left(\begin{array}{cc} \nabla F(x^*) & -I_{n \times n} \\ \phi(Z^*) & \phi(X^*) \end{array} \right) & 0 \\ 0 & 0 & \varepsilon \end{pmatrix} \right| = \varepsilon \left| \begin{pmatrix} \nabla F(x^*) & -I_{n \times n} \\ \phi(Z^*) & \phi(X^*) \end{pmatrix} \right| = \varepsilon |\nabla F(x^*)_{\mathbb{I}}|,$$

therefore the matrix $\lim_{r \rightarrow 0} \nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbf{X}^*, r)$ exists and is invertible if F is \mathcal{P} -function.

- If $\mathbf{X}^* = (x^*, z^*)$ does not satisfy the strict complementarity condition, then by Lemma 5.6, $\forall \mathbb{I} \subset \{1, \dots, n\}$ we have:

$$\begin{aligned} \lim_{r \rightarrow 0} \det(\nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbf{X}^*, r)) &= \left| \begin{pmatrix} \left(\begin{array}{cc} \nabla F(x^*) & -I_{n \times n} \\ \phi(Z^*) & \phi(X^*) \end{array} \right) & 0 \\ 0 & 0 & \varepsilon \end{pmatrix} \right| \\ &= \varepsilon \left| \begin{pmatrix} \nabla F(x^*) & -I_{n \times n} \\ \phi(Z^*) & \phi(X^*) \end{pmatrix} \right| = \varepsilon \left| \begin{pmatrix} \nabla F(x^*) - I_{n \times n} & -I_{n \times n} \\ \phi(Z^*) + \phi(X^*) & \phi(X^*) \end{pmatrix} \right|, \end{aligned}$$

from Lemma 5.2 we have:

$$\liminf_{r \rightarrow 0} \max((\nabla_x G_r^1(x^*, z^*))_{ii}, (\nabla_z G_r^1(x^*, z^*))_{ii}) \geq \frac{1}{4} > 0, \quad \forall i = 1, \dots, n, \quad (43)$$

hence $\lim_{r \rightarrow 0} (\nabla_x G_r^1(x^*, z^*) + \nabla_z G_r^1(x^*, z^*)) = \phi(Z^*) + \phi(X^*)$ is a positive diagonal matrix.

From Lemma 3.4, we take:

$$\begin{aligned} M &= \nabla F(x^*) - I_{n \times n}, \\ N_t &= \phi(X^*), \\ N_s &= \phi(Z^*) + \phi(X^*), \end{aligned} \quad (44)$$

where N_s is a positive diagonal matrix, and N_t is a nonnegative diagonale matrix. Therefore the matrix $\lim_{r \rightarrow 0} \nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbf{X}^*, r)$ exists and is invertible if $\nabla F(x^*) - I_{n \times n}$ is \mathcal{P}_0 -matrix.

Remark 5.1

The matrix $\lim_{r \rightarrow 0} \nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbf{X}^*, r)$ exists and is invertible if $\nabla F(x^*) - \beta I_{n \times n}$ is \mathcal{P}_0 -matrix for any $\beta > 0$. Since $\phi(Z^*) + \beta \phi(X^*)$ is a positive diagonal matrix for any $\beta > 0$.

Remark 5.2

If $\nabla F(x^*) - \beta I_{n \times n}$ is \mathcal{P}_0 -matrix for any $\beta > 0$ then $\nabla F(x^*)$ is a \mathcal{P} -matrix.

In the following, we focus our attention on the superlinear convergence rate of Algorithm 1.

Theorem 5.2

(Theorem 6.9, [5]). Let $\mathbb{H}_\theta : \mathbb{R}^{2n+1} \rightarrow \mathbb{R}^{2n+1}$ be a continuously differentiable function.

1. (Local analysis) Let \mathbb{X}^* be a regular zero of \mathbb{H}_θ . If \mathbb{X}^0 is close enough to \mathbb{X}^* , then $\zeta^k = 1$ for all k , and \mathbb{X}^k converge to \mathbb{X}^* super-linearly (and we recover the standard Newton method).

2. (Limit point) Let \mathbb{X}^* be a limit point of sequence $\{\mathbb{X}^k\}$. If $\nabla \mathbb{H}_\theta(\mathbb{X}^*)$ is invertible, then \mathbb{X}^* is a regular zero of \mathbb{H} . If \mathbb{X}^* is a regular zero of \mathbb{H}_θ , then $\zeta^k = 1$ for k big enough and \mathbb{X}^k converge to \mathbb{X}^* super-linearly.

Proof

We apply Theorem 5.1 and Lemma 5.4 under some condition on F . □

The next lemma measures the "additional coercivity" effect of the smoothing.

Lemma 5.5

Assume F is a \mathcal{P}_0 -function, then

1. \mathbb{H}_θ is a \mathcal{P} -function.
2. If $\mathbb{H}_\theta(\mathbf{X}, 0)$ exists then it is a \mathcal{P}_0 -function.

Proof

(1) Let X, Y be two distinct vectors of \mathbb{R}^{2n} . Since F is a \mathcal{P}_0 -function there exists an index $i \in \{1, \dots, 2n\}$ such that $X_i \neq Y_i$ and $(X_i - Y_i)(F_i(X) - F_i(Y)) \geq 0$. Without loss of generality, we can suppose that $X_i > Y_i$ and $F_i(X) > F_i(Y)$.

Since ψ and ψ^{-1} are decreasing, we obtain consecutively that for any $r > 0$,

$$\psi(X_i/r) + \psi(F_i(X)/r) < \psi(Y_i/r) + \psi(F_i(Y)/r), \tag{45}$$

so

$$G_r(X_i, F_i(X)) > G_r(Y_i, F_i(Y)).$$

Hence, \mathbb{H}_θ is a \mathcal{P} -function.

We will now deal with the case where $X_i = Y_i, \forall i < 2n + 1$. For $i = 2n + 1$, we can suppose that $X_{2n+1} > Y_{2n+1}$

$$(X_{2n+1} - Y_{2n+1})(\mathbb{H}_\theta(X)_{2n+1} - \mathbb{H}_\theta(Y)_{2n+1}) = (r_1 - r_2)(r_1^2 + \varepsilon r_1 - r_2^2 - \varepsilon r_2) > 0.$$

Hence, \mathbb{H}_θ is a \mathcal{P} function.

- (2) If $\mathbb{H}_\theta(\mathbb{X}, 0)$ exists, passing to the limit in (45) as $r \searrow 0$, we obtain that $\mathbb{H}_\theta(\mathbb{X}, 0)$ is a \mathcal{P}_0 -function. □

Now we would like to study the asymptotic behavior of the Jacobian matrix of our method with the Jacobian matrix of IPM when r goes to 0 and we need a lemma that is used to prove our main result.

Lemma 5.6

We consider the following system

$$\begin{aligned} Z.X &= 0 \\ Z \geq 0, \quad X \geq 0, \end{aligned} \tag{46}$$

where $Z = \text{diag}(z)$ and $X = \text{diag}(x)$.

Assume that Z, X are strictly complementary (i.e. $Z + X > 0_{n \times n}$). Then J is singular if and only if T is singular, where

$$J = \begin{pmatrix} \nabla F(x) & -I \\ Z & X \end{pmatrix}, \quad \text{and} \quad T = \begin{pmatrix} \nabla F(x) & -I \\ \phi(Z) & \phi(X) \end{pmatrix},$$

where $\phi(\cdot)$ is defined in Lemma 5.3, here ϕ operates component-wise on Z and X .

Proof

By the strict complementarity hypothesis, we range the rows and the columns of J and T as follows

$$J_\sigma = \begin{pmatrix} \nabla F(x)_\sigma & -I_\sigma \\ \begin{pmatrix} Z_1 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & X_2 \end{pmatrix} \end{pmatrix},$$

where $X_2 > 0$ and $Z_1 > 0$, and

$$(T)_\sigma = \begin{pmatrix} \nabla F(x)_\sigma & & & -I_\sigma \\ \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{pmatrix} & \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{pmatrix} & & \end{pmatrix}.$$

The determinant of the two matrices J_σ and $(T)_\sigma$ are equal to

$$\det(J_\sigma) = \begin{vmatrix} \nabla F(x)_\sigma & -I_\sigma \\ \begin{pmatrix} Z_1 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & X_2 \end{pmatrix} \end{vmatrix} = \pm \prod_{i \in \mathbb{I}_1} x_i \prod_{i \in \mathbb{I}_2} z_i \det(C),$$

$$\det(T_\sigma) = \begin{vmatrix} \nabla F(x)_\sigma & & & -I_\sigma \\ \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{pmatrix} & \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{pmatrix} & & \end{vmatrix} = \pm \prod_{i \in \mathbb{I}_1} \phi(x_i) \prod_{i \in \mathbb{I}_2} \phi(z_i) \det(C),$$

where C is a certain matrix, $\mathbb{I}_1 = \{i \mid x_i > 0\}$ and $\mathbb{I}_2 = \{i \mid z_i > 0\}$. Since

$$\pm \prod_{i \in \mathbb{I}_1} x_i \prod_{i \in \mathbb{I}_2} z_i \quad \text{and} \quad \prod_{i \in \mathbb{I}_1} \phi(x_i) \prod_{i \in \mathbb{I}_2} \phi(z_i),$$

are nonzeros, then we can conclude that J and T are invertibles and singulars at the same time. □

Theorem 5.3

Suppose that $\mathbf{X}^* = (x^*, z^*)$ is a solution of NCP which satisfies the strict complementarity, and $\nabla H_0(\mathbf{X}^*)$ be defined by (28) (the Jacobian matrix of the interior-point method) is invertible. Then $\lim_{r \rightarrow 0} \nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbf{X}^*, r)$ is invertible, i.e., the two Jacobian matrices are singular or nonsingular at the same time.

Proof

In view of Lemma 5.6, and thanks to the assumption $\mathbf{X}^* = (x^*, z^*)$ is a solution of NCP, we have $x^* \geq 0$ and $z^* \geq 0$, so that $x^- = z^- = 0$. Hence

$$\lim_{r \rightarrow 0} \det(\nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbf{X}^*, r)) = \left| \begin{pmatrix} \begin{pmatrix} \nabla F(x^*) & -I_{n \times n} \\ \phi(Z^*) & \phi(X^*) \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ \varepsilon \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ \varepsilon \end{pmatrix} \end{pmatrix} \right| = \varepsilon \left| \begin{pmatrix} \nabla F(x^*) & -I_{n \times n} \\ \phi(Z^*) & \phi(X^*) \end{pmatrix} \right|,$$

where $\phi(\cdot)$ is defined in Lemma 5.3, $Z^* = \text{diag}(z^*)$ and $X^* = \text{diag}(x^*)$.

From Lemma 5.6, we conclude that if $\nabla_{\mathbf{X}} H_0(\mathbf{X}^*)$ is invertible then $\lim_{r \rightarrow 0} \nabla_{\mathbb{X}} \mathbb{H}_\theta(\mathbf{X}^*, r)$ is invertible. This means, that if the IPM converges our method converges. □

Hypothesis H_0 (F is a \mathcal{P}_0 -function) assures us that our method is well defined and the Theorem 5.3 shows that the domain of convergence of our method is at least as large as that of the IPM.

6. Numerical experiments and applications

In this section, we present some numerical experiments for the two smoothing functions. Our aim is just to verify the theoretical assertions for these two “extreme” cases.

First, we study eight test problems with various sizes and characteristics. Then, we present a comparison of some randomly generated problems of our method and other approaches that have been suggested recently in [6, 14]. We also present numerical results for two concrete examples. All the codes are written in MATLAB 2020R and run in the system of Windows 10 with PC i5 8-th Gen and 16.00 GB RAM. We take the precision $\varepsilon = 10^{-9}$ (the termination criterion).

Example 6.1

We consider eight test problems (that can be found in [9, 16, 21, 26, 38, 40]) with various sizes and characteristics. In some cases, F is monotone or strongly monotone whereas others can have a non-connected solution set, in this case, F is at most a \mathcal{P}_0 -function. A precise description of each test problem is given in the appendix.

Table 1. Results for θ_1 and θ_2

Pb	size	Iter (θ_1, θ_2)	Opt. (θ_1, θ_2)	Feas. (θ_1, θ_2)	cpu time(s) (θ_1, θ_2)	r (θ_1, θ_2)
P1	10	(114, 47)	(7.58e-11, 7.66e-13)	(0, 8.23e-11)	(0.04, 0.0073)	(2.96e-04, 0.0014)
	100	(134, 65)	(9.31e-12, 1.15e-15)	(0, 1.51e-12)	(0.14, 0.07)	(1.06e-04, 9.21e-04)
	500	(148, 66)	(1.87e-12, 6.39e-16)	(0, 4.28e-12)	(5.24, 2.52)	(4.90e-05, 8.95e-04)
	1000	(153, 68)	(1.05e-12, 7.82e-14)	(0, 1.04e-09)	(25.30, 11.41)	(3.70e-05, 0.0012)
P2	10	(116, 47)	(7.53e-11, 7.66e-13)	(0, 8.23e-11)	(0.01, 0.02)	(2.83e-04, 1.41e-03)
	100	(133, 74)	(7.64e-12, 2.64e-13)	(0, 3.47e-10)	(0.19, 0.10)	(8.74e-05, 0.0013)
	500	(147, 84)	(1.60e-12, 1.29e-12)	(0, 8.62e-08)	(5.25, 3.14)	(4.01e-05, 0.0014)
	1000	(153, 115)	(8.22e-13, 2.25e-15)	(0, 3.02e-10)	(24.47, 19.89)	(2.86e-05, 9.53e-04)
P3	10	(14, 16)	(3.46e-10, 1.65e-19)	(0, 3.99e-18)	(0.006, 0.02)	(0.001, 4.49e-04)
	100	(108, 44)	(4.99e-10, 3.63e-20)	(0, 4.75e-21)	(0.22, 0.05)	(0.0014, 0.0042)
	500	(353, 140)	(7.58e-10, 3.97e-11)	(9.68e-13, 7.54e-11)	(35.03, 5.14)	(0.0011, 0.002)
	1000	(675, 265)	(8.94e-10, 2.32e-10)	(1.31e-12, 1.58e-11)	(91.41, 24.77)	(0.0011, 0.002)
P4	4	(53, 58)	(2.97e-10, 1.20e-10)	(1.24e-07, 4.19e-08)	(0.008, 0.0242)	(6.06e-04, 1.72e-04)
P5	4	(16, 14)	(3.02e-10, 1.92e-10)	(0, 3.84e-10)	(0.003, 0.009)	(0.0026, 0.018)
P6	7	(10, 13)	(1.06e-10, 7.16e-11)	(0, 0)	(0.1264, 0.0044)	(0.0016, 1.33e-04)
P7	5	(33, 30)	(2.23e-11, 1.16e-10)	(0, 3.44e-14)	(0.011, 0.016)	(0.004, 0.003)
P8	10	(65, 45)	(7.21e-11, 2.27e-11)	(1.34e-12, 5.18e-11)	(0.18, 0.16)	(0.0018, 3.76e-04)

In this table, **Size** stands for the number of variables, **Iter** corresponds to the total number of Jacobian evaluations, **Opt.** and **Feas.** correspond to the following optimality and feasibility measures

$$\text{Opt.} := \max_{1 \leq i \leq n} |x_i F_i(x)| \quad \text{and} \quad \text{Feas.} := \|\min(x, 0)\|_1 + \|\min(F(x), 0)\|_1.$$

The results clearly show that our methods are efficient. We also remark that the second smoothing function is much more efficient and powerful than the first one.

In the next table, we compare the results of θ_2 -smoothing approach to three state-of-the-art methods (Namely: Fischer Burmeister (FB-Alg) [6], Newton Min (Min-Alg) and projection method (PM-Alg) [10]). We make a comparison among Algorithm 1, FB-Alg, Min-Alg, and PM-Alg by implementing these algorithms to solve the same benchmark test problems available in the literature. Since, we can not compare the iterative numbers, we only present the optimality measures, and cpu time(s).

Table 2. Comparison of Algorithm 1 (θ_2) with FB-Alg, Min-Alg and PM-Alg.

Pb	size	Algorithm 1 (θ_2)		FB-Alg		Min-Alg		PM-Alg	
		Opt.	time(s)	Opt.	time(s)	Opt.	time(s)	Opt.	time(s)
P1	10	7.66e-13	0.0073	4.23e-11	0.0273	3.11e-07	0.0092	2.5e-09	0.63
	100	1.15e-15	0.07	1.55e-10	0.1052	1.55e-14	0.0134	9.1e-11	5.48
	500	6.39e-16	2.52	4.47e-11	10.3936	1.55e-14	0.0282	4.7e-10	96.37
	1000	7.82e-14	11.41	4.47e-11	96.5009	1.55e-14	0.0557	8.8e-11	224.04
P2	10	7.66e-13	0.02	2.47e-11	0.0512	1.76e-12	0.0136	2.2e-10	1.19
	100	2.64e-13	0.10	1.25e-10	1.7285	5.06e-17	0.0193	7.1e-12	5.76
	500	1.29e-12	3.14	6.04e-11	5.6432	2.06e-12	0.6265	5.3e-12	112.41
	1000	2.25e-15	19.89	9.51e-12	25.1059	2.20e-13	3.2308	2.9e-11	336.20
P3	10	1.65e-19	0.02	6.58e-11	0.0272	4.24e-17	0.0078	6.4e-11	1.03
	100	3.63e-20	0.05	5.90e-11	0.0327	4.24e-17	0.0080	1.8e-12	5.19
	500	3.97e-11	5.14	1.30e-11	1.6040	4.24e-17	0.020	5.8e-13	90.22
	1000	2.32e-10	24.77	5.56e-11	11.0915	4.24e-17	0.0314	2.4e-11	350.06
P4	4	1.20e-10	0.0242	2.94e-14	0.0435	6.12e-10	0.0019	3.1e-12	0.19
P5	4	1.92e-10	0.009	2.37e-10	0.2752	2.22e-16	0.0120	1.4e-12	0.34
P6	7	7.16e-11	0.0044	5.70e-10	0.1158	6.05e-17	0.0341	2.3e-11	0.31

The results clearly show that our methods are efficient, competitive, and superior to the Fischer Burmeister method and projection method.

Example 6.2

This example is described in [16, 40]. The corresponding function $F(x)$ is of the form:

$$F(x) = (AA^T + B + D)x + q,$$

where the matrices A , B and D are randomly generated as: any entry of the square $n \times n$ matrix A and of the $n \times n$ skew-symmetric matrix B is uniformly generated from $] - 5, 5[$, and any entry of the diagonal matrix D is uniformly generated from $]0, 3[$. The vector q is uniformly generated from $] - 500, 0[$.

The matrix $AA^T + B + D$ is a positive definite and the function F is strongly monotone. We used the M-files proposed in [40] to generate A , B , D and q .

In this example, we will compare our methods already mentioned in sections 3 and 4, named: Algorithm 1 (θ_1), Algorithm 1 (θ_2) to some other methods (Newton min method (Min-Alg), Fischer-Burmeister's method [6] (FB-Alg) and the classical interior-point method [14] (IPM-Alg)). In order to complete this experiment, we propose the performance profiles, developed by E. D. Dolan and J. J. Moré [8], as a tool for the comparative analysis of these methods.

We set " $n_s = 5$ " as the number of methods and we have chosen " $n_p = 100$ " (problems to be tested). We are interested in the comparison of the computation time and the number of iterations.

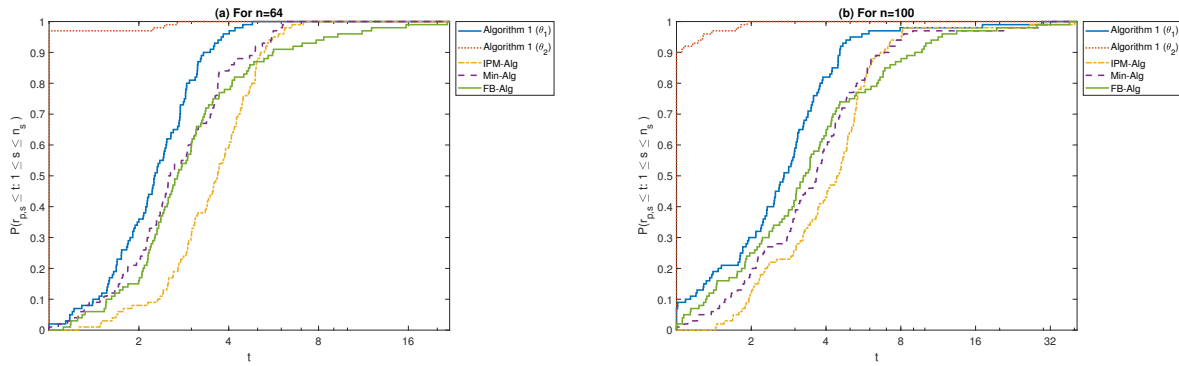


Figure 3. Performance profiles where $t_{p, s}$ represents the average computation time.

The figure above shows the performance profiles of five methods where the performance measure is execution time. It is clear that our method with the θ_2 -function captures our attention (admits the highest probability value). In fact, in the interval $[0, 1]$, our method is able to solve 99% of the problems, while the other methods do not reach 20% and require more time. We also notice that IPM-Alg is the slowest compared to others. However, for $t > 4$, the three algorithms FB-Alg, Min-Alg, and Algorithm 1 with θ_1 -function confirm their robustness. Figure 3 also indicates that, with respect to the computation time, with the same initial points and under the same stopping criterion, our method with θ_2 -function (resp. θ_1 -function) is the fastest method, followed respectively by Min-Alg, FB-Alg, and IPM-Alg.

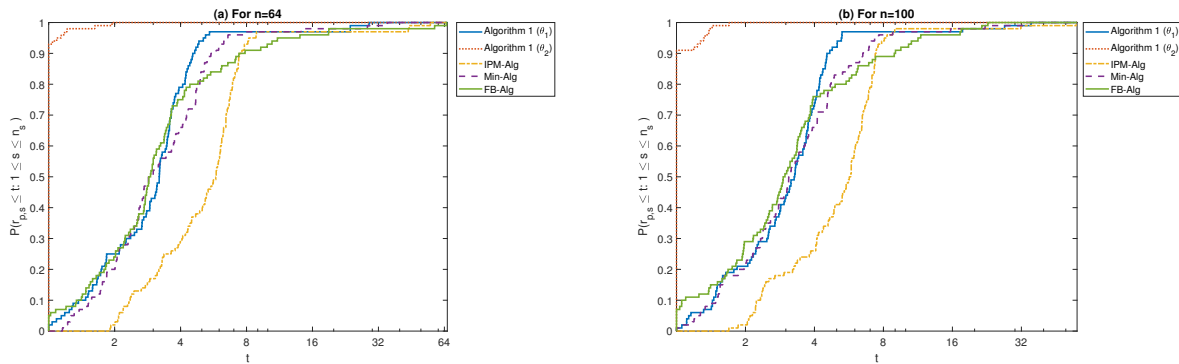


Figure 4. Performance profiles where $t_{p, s}$ represents the average number of iterations.

In Figure 4, we illustrate the performance profiles of five methods considering the number of iterations required as a performance measure. We notice that our method with the θ_2 -function is the winner (admits the highest probability value) followed by our method with the θ_1 -function, Min-Alg, and FB-Alg. We also note that IPM-Alg needs more iterations to resolve problems. The performance of our method with the θ_1 -function becomes interesting beyond $t = 3$.

Example 6.3

(Geochemical Models [35]) The problem comes from Geochemistry. We introduce a model which are 2-salts. The main idea is that we need to find a way to reformulate a general problem to a problem which has a form like $G(x) = 0$. We show the numerical results by applying several iteration methods.

Let T, K are constant vectors which have meaning in chemistry. We define the problem as follows:

Let $x = (x_1, x_2, x_3)$ and $p = (p_1, p_2)$,

$$\begin{aligned}
 H : \mathbb{R}^5 &\rightarrow \mathbb{R}^3 & F : \mathbb{R}^3 &\rightarrow \mathbb{R}^2 & G : \mathbb{R}^5 &\rightarrow \mathbb{R}^5 \\
 (x, p) \rightarrow H(x, p) &= \begin{bmatrix} T_1 - x_1 - p_1 \\ T_2 - x_2 - p_2 \\ x_3 - x_2 - x_1 \end{bmatrix}, & x \rightarrow F(x) &= \begin{bmatrix} K_1 - x_1x_3 \\ K_2 - x_2x_3 \end{bmatrix}, & (x, p) \rightarrow G(x, p) &= \begin{bmatrix} H(x, p) \\ P^T F(x) \\ p \geq 0, F(x) \geq 0 \end{bmatrix}.
 \end{aligned}$$

We want to solve the equation

$$G(x, p) = 0 \iff \begin{bmatrix} H(x, p) \\ P^T F(x) \\ p \geq 0, F(x) \geq 0 \end{bmatrix} = 0. \tag{47}$$

Using our approach with $\theta_r = \theta_r^1$ (resp. $\theta_r = \theta_r^2$), we reformulate (47) and we get

$$G_{\theta_r^1}(x, p) = \begin{bmatrix} K_1 - x_1x_3 - z_1 \\ K_2 - x_2x_3 - z_2 \\ T_1 - x_1 - x_4 \\ T_2 - x_2 - x_5 \\ x_3 - x_2 - x_1 \\ \frac{x_4z_1 - r^2}{x_4 + z_1 + 2r} \\ \frac{x_5z_2 - r^2}{x_5 + z_2 + 2r} \\ \frac{1}{2}\|x^-\|^2 + \frac{1}{2}\|z^-\|^2 + r^2 + \varepsilon r \end{bmatrix} = 0, \quad G_{\theta_r^2}(x, p) = \begin{bmatrix} K_1 - x_1x_3 - z_1 \\ K_2 - x_2x_3 - z_2 \\ T_1 - x_1 - x_4 \\ T_2 - x_2 - x_5 \\ x_3 - x_2 - x_1 \\ -r \log(e^{-x_4/r} + e^{-z_1/r}) \\ -r \log(e^{-x_5/r} + e^{-z_2/r}) \\ \frac{1}{2}\|x^-\|^2 + \frac{1}{2}\|z^-\|^2 + r^2 + \varepsilon r \end{bmatrix} = 0$$

and considering $x_4 = p_1, x_5 = p_2$.

Since our focus is on the effect of different smoothing approaches in solving (47), we replace the complementarity constraint by the Min function, and by the Fischer-Burmeister’s function [6]. The corresponding two algorithms are referred to as Min-Alg and FB-Alg, respectively. We make a comparison among Algorithm 1, Min-Alg, and FB-Alg, IPM-Alg (the classical interior-point method [14]) by implementing these algorithms to solve problem (47).

The Table 3, and Figure 5 show the results with the intial point $x_0 = (3, 1, 4, 5, 6)^T$, $T = (2, 6)^T$, and $K = (37.5837, 7.6208)^T$.

Table 3. Comparison of Algorithm 1 (θ_1) and Algorithm 1 (θ_2) with Min-Alg, FB-Alg, and IPM-Alg.

Iter	Algorithm 1 (θ_1)	Algorithm 1 (θ_2)	Min-Alg	FB-Alg	IPM-Alg
k	$\ G_{\theta_1}(x^k)\ $	$\ G_{\theta_2}(x^k)\ $	$\ G_{Min}(x^k)\ $	$\ G_{FB}(x^k)\ $	$\ G_{IPM}(x^k)\ $
0	24.5837	24.5837	24.5837	24.5837	24.5837
1	8.2104	13.2212	13.3538	8.9798	8.3720
2	2.5518	12.0213	12.1412	4.2253	6.2486
3	4.1723	9.6610	11.0112	1.2774	1.8581
4	1.9497	6.1377	9.9651	0.4355	1.2035
5	0.0382	1.3388	9.0020	0.2001	0.8359
6	0.0063	0.3347	8.1192	0.1049	0.2406
7	0.0012	0.0838	7.1321	0.0537	0.0032
8	2.4936e-04	0.0210	6.0534	0.0272	5.1935e-05
9	5.0533e-05	0.0052	4.6609	0.0137	1.6166e-05
10	1.0487e-05	0.0013	2.9998	0.0068	4.9426e-06
11	2.1474e-06	3.2750e-04	1.2778	0.0034	1.5108e-06
12	4.4258e-07	8.1875e-05	0.0563	0.0017	4.6172e-07
13	9.1435e-08	2.0469e-05	1.2410e-05	8.5782e-04	1.4110e-07
14	1.8834e-08	5.1172e-06	4.4658e-12	4.2899e-04	4.3119e-08
15	4.5031e-09	1.2793e-06		2.1451e-04	1.3176e-08
16	9.6186e-10	3.1982e-07		1.0726e-04	4.0264e-09
17		7.9956e-08		5.3631e-05	1.2304e-09
18		1.9989e-08		2.6816e-05	3.7597e-10
19		4.9973e-09		1.3408e-05	
20		1.2493e-09		6.7041e-06	
21		3.1233e-10		3.3520e-06	
⋮				⋮	
34				4.0918e-10	

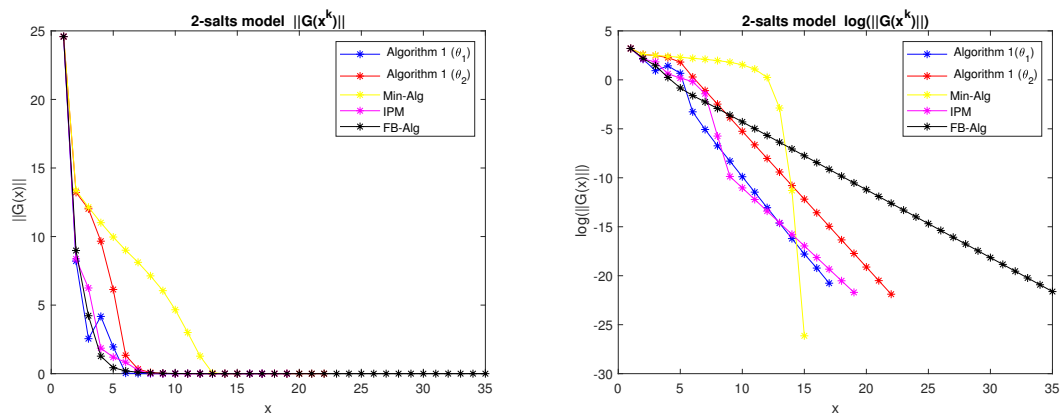


Figure 5. 2-salts model of $\|G(x^k)\|_\infty$ and $\log(\|G(x^k)\|_\infty)$.

According to Figure 5, the two best methods based on the number of iterations are the semi-smooth Newton min method (14 iters) followed by our method (for θ_1 -function) with (16 iters). Next, come the classical interior-point method and our method (for θ_2 -function) with (18 iters) and (21 iters), respectively. The last is the Fischer-Burmeister's method with (34 iters).

7. Conclusion

In this paper, we have presented a new smoothing approach for solving the nonlinear complementarity problem. For such an approach, some useful properties have been analyzed, which were employed to develop a well-defined and efficient Jacobian Newton algorithm for solving the nonlinear complementarity problem with \mathcal{P}_0 -function. We have established a result of global convergence and a super-linear convergence rate for the developed algorithm. Numerical experiments prove the efficiency of our study in the following aspects:

1. It can find the solution of NCP either with less number of iterations or with higher precision than the other.
2. It is relatively more robust when the dimension of the test problems increases. In particular, it seems to be more suitable for large-scale problems.
3. It is more efficient to find the nondegenerate solution of NCP with less iteration number than the others.

Appendix

We give in this appendix a brief description of each test example.

1. The two first examples **P1** and **P2** [21] correspond to strongly monotone function

$$F(x) = (F_1(x), \dots, F_n(x))^T \quad \text{with} \quad F_i(x) = -x_{i+1} + 2x_i - x_{i-1} + \frac{1}{3}x_i^3 - b_i, \quad i = 1, \dots, n.$$

where $x_0 = x_{n+1} = 0$ and $b_i = (-1)^i$ (resp. $b_i = \frac{(-1)^i}{\sqrt{i}}$), $i = 1, \dots, n$, for **P1** (resp. **P2**).

2. **P3** is another strongly monotone test problem from [9] where $F(x) = (F_1(x), \dots, F_n(x))^T$ with

$$F_i(x) = -x_{i+1} + 2x_i - x_{i-1} + \arctan(x_i) + \left(i - \frac{\pi}{2}\right), \quad i = 1, \dots, n, \quad (x_0 = x_{n+1} = 0).$$

3. **P4** and **P5** are known as the degenerate and non-degenerate examples of Kojima-Shindo [26]. **P4** and **P5** are respectively defined by

$$F_4(x) = \begin{bmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{bmatrix}, \quad F_5(x) = \begin{bmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 3x_4 - 1 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{bmatrix}.$$

P5 has a unique solution $x^* = \left(\frac{\sqrt{6}}{2}, 0, 0, \frac{1}{2}\right)$ with $F(x^*) = \left(0, 2 + \frac{\sqrt{6}}{2}, 5, 0\right)$ while **P4** has two optimal solutions $x^* = \left(\frac{\sqrt{6}}{2}, 0, 0, \frac{1}{2}\right)$ with $F(x^*) = \left(0, 2 + \frac{\sqrt{6}}{2}, 0, 0\right)$ and $x^{**} = (1, 0, 3, 0)$ with $F(x^*) = (0, 31, 0, 4)$. The first optimal solution of **P4** is degenerate since $x_3^* = F_3(x^*) = 0$.

4. [38] In problem (4), $x \in \mathbb{R}^7$ and $F(x) : \mathbb{R}^7 \rightarrow \mathbb{R}^7$ is given by

$$F_6(x) = \begin{bmatrix} 2x_1 - x_3 + x_5 + 3x_6 - 1 \\ x_2 + 2x_5 + x_6 - x_7 - 3 \\ -x_1 + 2x_3 + x_4 + x_5 + 2x_6 - 4x_7 + 1 \\ x_3 + x_4 + x_5 - x_6 - 1 \\ -x_1 - 2x_2 - x_3 - x_4 + 5 \\ -3x_1 - x_2 - 2x_3 + x_4 + 4 \\ x_2 + 4x_3 - 1.5 \end{bmatrix}.$$

P6, has a non-degenerate solution

$$x^* = (0.2727, 2.0909, 0, 0.54545, 0.4545, 0, 0)^T.$$

5. A complete description of **P7** and **P8** can be found in [40, 16]. These two examples correspond to the Nash-Cournot test problem with $N = 5$ and $N = 10$.

Let $x \in \mathbb{R}^N$, $Q = \sum x_i$ and define the functions $C_i(x_i)$ and $p(Q)$ as follows:

$$P(Q) = 5000^{\frac{1}{\gamma}} Q^{-\frac{1}{\gamma}}, \quad C_i(x_i) = c_i x_i + \frac{b_i}{1 + b_i} L_i^{\frac{1}{b_i}} x_i^{\frac{b_i+1}{b_i}}.$$

The NCP function is given by

$$F_i(x) = C'_i(x_i) - P(Q) - x_i p'(Q), \quad i = 1, \dots, N,$$

with $c_i, L_i, b_i > 0$ and $\gamma \geq 1$. For our numerics, we used:

- $N = 5$, $c = [10, 8, 6, 4, 2]^T$, $b = [1.2, 1.1, 1, 0.9, 0.8]^T$, $L = [5, 5, 5, 5, 5]^T$, $e = [1, 1, 1, 1, 1]^T$ and $\gamma = 1.1$.
- $N = 10$, $c = [5, 3, 8, 5, 1, 3, 7, 4, 6, 3]^T$, $b = [1.2, 1, 0.9, 0.6, 1.5, 1, 0.7, 1.1, 0.95, 0.75]^T$, $L = [10, 10, 10, 10, 10, 10, 10, 10, 10, 10]^T$, $e = [1, 1, 1, 1, 1, 1, 1, 1, 1, 1]^T$ and $\gamma = 1.2$.

REFERENCES

1. L. Abdallah, M. Haddou, and T. Migot, *Solving absolute value equation using complementarity and smoothing functions*, Journal of Computational and Applied Mathematics, vol. 327, pp. 1196–207, 2018.
2. A. Auslender, R. Cominetti, and M. Haddou, *Asymptotic analysis for penalty and barrier methods in convex and linear programming*, Mathematics of Operations Research, vol. 22, no. 1, pp. 43–62, 1997.
3. M. Bergounioux, and M. Haddou, *A new relaxation method for a discrete image restoration problem*, Journal of Convex Analysis, vol. 17, no. 3, pp. 861–883, 2010.
4. F. Bonnans, *Optimisation continue: cours et problèmes corrigés, Mathématiques appliquées pour le Master*, Dunod, 2006.
5. F. Bonnans, *Optimisation continue: cours et problèmes corrigés, Mathématiques appliquées pour le Master*, Dunod, 2006.
6. X. Chen, B. Chen, and C. Kanzow, *A penalized Fischer-Burmeister NCP function: theoretical investigation and numerical results*, Int. fur Angewandte Mathematik der Univ. Mathematical Programming, 1997.
7. C. Chen, and O. L. Mangasarian, *A class of smoothing functions for nonlinear complementarity problems*, Computational Optimization and Applications, vol. 5, pp. 97–138, 1996.
8. E. D. Dolan, and J. J. Moré, *Benchmarking optimization software with performance profiles*, Mathematical Programming, vol. 91, pp. 201–213, 2002.
9. L. Dong-hui, and Z. Jin-ping, *A penalty technique for nonlinear problems*, Journal of Computational Mathematics, vol. 16, no. 1, pp. 40–50, 1998.
10. F. Facchinei, and J. S. Pang, *Finite-dimensional Variational Inequalities and Complementarity Problem*, Springer Series in Operations Research, Springer-Verlag, New York, vol. 1-2, 2003.
11. M. C. Ferris, and J. S. Pang, *Engineering and economic applications of complementarity problems*, Society of Inian Automobile Manufacturers Review, vol. 39, no. 4, pp. 669–713, 1997.
12. A. Fischer, *Solution of monotone complementarity problems with locally Lipschitz functions*, Mathematical Programming, vol. 76, no. 2, pp. 513–532, 1997.
13. A. Fischer, *A special Newton-type optimization method*, Optimization, vol. 24, no. 3-4, pp. 269–284, 1992.
14. M. E. Ghami, *New primal-dual interior-point methods based on kernel functions*, PhD thesis, 2005.
15. C. Geiger, and C. Kanzow, *On the solution of monotone complementarity problems*, Computational Optimization and Applications, vol. 5, pp. 155–173, 1996.

16. P. T. Harker, *Accelerating the convergence of the diagonalization and projection algorithms for finite-dimensional variational inequalities*, *Mathematical Programming*, vol. 41, pp. 29–59, 1988.
17. M. Haddou, *A new class of smoothing methods for mathematical programs with equilibrium constraints*, *Pacific Journal of Optimization*, vol. 5, pp. 87–95, 2009.
18. M. Haddou, and P. Maheux, *Smoothing methods for nonlinear complementarity problems*, *Journal of Optimization Theory and Applications*, vol. 160, pp. 711–729, 2014.
19. M. Haddou, T. Migot, and J. Omer, *A generalized direction in interior-point method for monotone linear complementarity problems*, *Optimization Letters*, vol. 13, pp. 35–53, 2019.
20. P. T. Harker, N. Yamashita, *Finite dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications*, *Mathematical Programming*, vol. 48, pp. 161–220, 1990.
21. C. Huang, and S. Wang, *A power penalty approach to a Nonlinear Complementarity Problem*, *Operations Research Letters*, vol. 38, no. 1, pp. 72–76, 2010.
22. C. Kanzow, N. Yamashita, and M. Fukushima, *New NCP functions and their properties*, *Journal of Optimization Theory and Applications*, vol. 94, pp. 115–135, 1997.
23. C. Kanzow, *Some equation-based methods for the nonlinear complementarity problem*, *Optimization Methods and Software*, vol. 3, pp. 327–340, 1994.
24. C. Kanzow, and H. Pieper, *Jacobian smoothing methods for nonlinear complementarity problems*, *SIAM Journal on Optimization*, vol. 9, no. 2, pp. 342–373, 1999.
25. N. Krejic, and S. Rapaji, *Globally convergent Jacobian smoothing inexact Newton methods for NCP*, *Computational Optimization and Applications*, vol. 41, pp. 243–261, 2008.
26. M. Kojima, and S. Shindo, *Extensions of Newton and quasi-Newton methods to systems of PCl equations*, *Journal of the Operations Research Society of Japan*, vol. 29, pp. 352–374, 1986.
27. C.F. Ma, P. Y. Nie, and G. P. Liang, *A new smoothing equations approach to the nonlinear complementarity problems*, *Journal of Computational Mathematics*, vol. 21, pp. 747–758, 2003.
28. MATLAB, version R2020. Natick, Massachusetts: The Math Works Inc, 2019.
29. J. J. Moré, *Global methods for nonlinear complementarity problems*, *Mathematics of Operations Research*, vol. 21, no. 3, pp. 589–614, 1996.
30. P. Y. Nie, *A null space approach for solving nonlinear complementarity problems*, *Acta Mathematicae Applicatae Sinica*, vol. 22, no. 1, pp. 9–20, 2006.
31. J. S. Pang, *A B-differentiable equations based, globally and locally quadratically convergent algorithm for nonlinear programming, complementarity, and variational inequality problems*, *Mathematical Programming*, vol. 51, pp. 101–131, 1991.
32. J. S. Pang, and S. A. Gabriel, *NE-SQP: A robust algorithm for nonlinear complementarity problem*, *Mathematical Programming*, vol. 60, pp. 295–337, 1993.
33. L. Qi, and D. H. Li, *A smoothing Newton method for nonlinear complementarity problems*, *Advanced Modeling and Optimization*, vol. 13, no. 2, pp. 141–152, 2011.
34. T. Migot, *Contributions aux méthodes numériques pour les problèmes de complémentarité et problèmes d’optimisation sous contraintes de complémentarité*, Ph.D. thesis, INSA Rennes, 2017.
35. T. Migot, and Jocelyne Erhel, *Analyse mathématique de modèles géochimiques*, Rapport de recherche, INRIA Rennes, équipe SAGE, 2014.
36. D. T. S. Vu, I.B. Gharbia, M. Haddou, and Q. H. Tran, *A new approach for solving nonlinear algebraic systems with complementarity conditions*, *Mathematics and Computers in Simulation*, vol. 190, pp. 1243–1274, 2021.
37. N. Yamashita, and M. Fukushima, *Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems*, *Mathematical Programming*, vol. 76, pp. 469–491, 1997.
38. W. Zhong, Y. Min, and W. Chang, *A partially smoothing Jacobian method for nonlinear complementarity problems with P0 function*, *Journal of Computational and Applied Mathematics*, vol. 286, pp. 158–171, 2015.
39. L. Zhang, J. Han, and Z. Huang, *Superlinear quadratic one step smoothing Newton method for P0 NCP*, *Acta Mathematica Sinica*, vol. 21, no. 2, pp. 117–128, 2005.
40. <http://dm.unife.it/pn2o/software/Extragradient/test-problems.html>.