Math. Program., Ser. A 86: 533–563 (1999)

Digital Object Identifier (DOI) 10.1007/s101079900097

Ji-Ming Peng · Zhenghua Lin

A non-interior continuation method for generalized linear complementarity problems*

Received September 3, 1997 / Revised version received April 27, 1999
Published online July 19, 1999

Abstract. In this paper, we propose a non-interior continuation method for solving generalized linear complementarity problems (GLCP) introduced by Cottle and Dantzig. The method is based on a smoothing function derived from the exponential penalty function first introduced by Kort and Bertsekas for constrained minimization. This smoothing function can also be viewed as a natural extension of Chen-Mangasarian's neural network smooth function. By using the smoothing function, we approximate GLCP as a family of parameterized smooth equations. An algorithm is presented to follow the smoothing path. Under suitable assumptions, it is shown that the algorithm is globally convergent and local Q-quadratically convergent. Few preliminary numerical results are also reported.

Key words. generalized linear complementarity problem – non-interior continuation method – Newton method – Q-quadratical convergence

1. Introduction

A matrix $N \in \Re^{m_0 \times n}$, $(m_0 \geq n)$ is a vertical block matrix of type $(m_1, ..., m_n)$ if it can be partitioned, row-wise, into n blocks so that the ith block, $N^i \in \Re^{m_i \times n}$ $(i = 1, ..., n)$ and $m_0 = \sum_{i=1}^{n} m_i$. If the constant vector $q \in \Re^{m_0}$ is partitioned conformably with N, i.e.

$$N = \begin{pmatrix} N^1 \\ N^2 \\ \vdots \\ N^n \end{pmatrix} \in \Re^{m_0 \times n},\ q = \begin{pmatrix} q^1 \\ q^2 \\ \vdots \\ q^n \end{pmatrix} \in \Re^{m_0}, \quad N^i \in \Re^{m_i \times n},\ q^i \in \Re^{m_i},\ \sum_{i=1}^{n} m_i = m_0.$$

The generalized linear complementarity problem (denoted by GLCP) associated with N and q is to find a vector $x \in \Re^n$ such that

$$N^i x + q^i \geq 0, \quad x \geq 0, \quad x_i \prod_{j=1}^{m_i} \left(N^i x + q^i\right)_j = 0,\ i = 1, 2, ..., n, \tag{1}$$

* The work of the first author was supported by Project 19601035 of NSFC. The work is finished when the second author visit State Key Laboratory of Scientific and Engineering Computing, Chinese Academy of Sciences

J.-M. Peng: State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific Computing, Academic Sinica, Beijing, P.O. Box 2719, China, 100080
Current address: Faculty of Technical Mathematics and Informatics, Delft University of Technology, P.O. Box 5031, 2628 CD, Mekelweg 4, Delft, The Netherlands, e-mail: `j.peng@its.tudelft.nl`

Z. Lin: Department of Mathematics, Jilin University, Changchun 130023, P.R. China

Mathematics Subject Classification (1991): 90C33

where x_i and $\left(N^i x + q^i\right)_j$ denote ith element of x and jth element of $\left(N^i x + q^i\right)$ respectively, $\prod_{j=1}^{m_i} z_j$ denotes $z_1 z_2 \cdots z_{m_i}$. This problem was first posed by Cottle and Dantzig [11]. It has many meaningful applications in different fields such as mathematical programming, game theory, control theory and economics [14,22,38,39]. Several authors have studied this problem and some numerical methods have been proposed to solve GLCP, the interested readers are referred to [12,13,31,32,39,42] and the references therein.

If $m_i = 1$ $(i = 1, ..., n)$, then the GLCP reduces to a linear complementarity problem (LCP), a special case of the following general complementarity problem (or CP for short)

$$x \geq 0,\ F(x) \geq 0,\ x_i F_i(x) = 0, \quad x \in \Re^n, \tag{2}$$

where $F(x)$ is a mapping from $\Re^n$ into itself. A useful way for solving CP is to reformulate it first as a system of nonsmooth (or smooth) equations and then try to find the solution of the complementarity problem by solving the correspondent system of equations. In recent years, reformulation of CPs has become a hot topic in the field of mathematical programming, and great progress has been made in this direction, for details see the survey paper [19] and the references therein. Generally speaking, there are two ways to transform a complementarity problem into a system of equations. The first one is to introduce some parameters or artificial variables, and then approximate the original problem as a family of parameterized smooth equations. For example, Chen and Mangasarian [7,8] introduced a class of smoothing functions and approximated CPs via a system of parameterized smooth equations, their approach was further studied by Chen and Xiu [9]. The interesting readers are referred to [5,6,10,24,36,44] and the references therein for more recent developments in this topic. In most cases, the solution set of the smooth equations system forms a path as the smooth parameter goes to zero, this path is usually called the smoothing path. The second way is to cast CPs as a system of nonsmooth equations via some equivalent transformations. For instance, by using the so-called *fischer-function* [16,17] or the minimum function [23,33], one can transform CP into a nonsmooth equations system, for more details, see [20,25]. A natural extension of the above mentioned results is to consider smoothing methods for GLCP.

We note (as pointed out by one referee) that in principle, one can also reformulate the GLCP as a linear complementarity problem in $\Re^{m_0}$ by introducing some artificial variables (see Sect. 2.3 in [15], and [21,32,36]) and then apply the smoothing methods for CPs to solve the reformulated problem. For instance, if $m = m_1 = m_2 \cdots = m_n$, then we get an artificial LCP in $\Re^{mn}$. In this situation, by using Chen-Mangasarian's smoothing function (or other smoothing functions) to approximate the reformulated CP, we get a system of equations in $\Re^{mn}$. It is of interests to approximate the GLCP as an equations system in the original space $\Re^n$. Different from the above mentioned approaches, we propose in this paper such a new reformulation for GLCP in $\Re^n$.

The continuation method is widely used in different fields, it is closely related to the homotopy method in numerical analysis [1], the path-following algorithms in interior point algorithm [28], and many smoothing methods for complementarity problems developed recently [6–9, 25, 26, 43]. The interior point path-following algorithm for CPs demands that all iterates stay in the positive orthant. In this case, the smoothing path reduces to the central path, a term often used in the interior point algorithm literature. Many results about the interior point algorithms for solving CPs have been reported, see [18] for a survey in this direction. On the other hand, the so-called noninterior continuation methods does not require the initial point and intermediate iterates to be in the positive orthant, so they are usually more flexible for numerical implementation.

Consider the constrained optimization problem as follows

$$\min\ f(x),\ \text{ s.t. } g_i(x) \leq 0,\ i = 1, 2, \ldots, m. \tag{3}$$

A useful approach for solving (3) is the penalty methods. In [29], Kort and Bertsekas proposed the following exponential penalty function

$$g(x,t) = t \ln \sum_{i=1}^{m} \exp\left(\frac{g_i(x)}{t}\right) \tag{4}$$

with a penalty parameter t and cast (3) as the following unconstrained optimization problem

$$\min\ f(x) + g(x,t).$$

Bertsekas further studied the function $g(x,t)$ and its applications [2–4]. On the other hand, it is easy to verify that the problem (3) is equivalent to the following mathematical programming problem with one constraint

$$\min\ f(x),\ \text{ s.t. } g(x) \leq 0, \tag{5}$$

where $g(x) = \max\{g_i(x) : i = 1, 2, \ldots, m\}$ is a piecewise smooth function. Since $g(x)$ is not usually differentiable even if all $g_i(x)$ are differentiable, it is difficult to use classical methods for constrained optimization problem to solve (5). As a remedy for this point, Li [30] proposed to approximate the problem (5) as a parameterized programming problems defined below

$$\min\ f(x),\ \text{ s.t. } g(x,t) \leq 0.$$

Some properties of $g(x,t)$ were also rediscovered by Li. In [41] Tang and Zhang also studied the properties of the function $g(x,t)$.

Our approach here follows the ideas of Li. As a consequence, we first use the function $g(x,t)$ to approximate the GLCP and reformulate it as a system of parameterized smooth equations. Then we discuss the regularity of the smoothing path under suitable conditions, and analyse the distance from the smoothing path to the solution set of the undertaking GLCP. Thirdly, we propose an algorithm to trace this smoothing path.

The paper is organized as follows: In Sect. 2, we first introduce some concepts and notations which will be used in this paper. Then some results about the vertical block matrix N are presented and the behavior of a system of nonsmooth equations equivalent to GLCP is explored. In Sect. 3, we discuss the interrelations between Chen-Mangasarian's smoothing functions and the smoothing function $g(x,t)$. Some properties of the function $g(x,t)$ (4) are also reported. We smooth the equivalent mapping $H(x)$ of GLCP via the function $g(x,t)$ and approximate the GLCP as a system of parameterized smooth equations in Sect. 4. It is shown that, when the parameter $t > 0$ is sufficiently small, the solution of the smooth equations can approximate the solution set of GLCP to any desired accuracy. The regularity of the smoothing path is also discussed. We propose in Sect. 5 an algorithm to follow the smoothing path and establish the global convergence of the algorithm under suitable assumptions. The local convergence property of the algorithm is studied in Sect. 6, and some numerical results are presented in Sect. 7. Finally we end this paper by some remarks.

Few words about our notations: throughout this work, $\Re^n_+$, $\Re^n_{++}$ denote the nonnegative orthant and the positive orthant in $\Re^n$, respectively. $\|x\|$ denotes the 2-norm of a vector $x \in \Re^n$ and $\|x\|_1$ the 1-norm. For any matrix $M \in \Re^{n\times n}$, $\|M\| = \max_{\|x\|=1} \|Mx\|$.

2. Preliminaries

First, we state some concepts and notations which will be used in this paper.

Definition 1. [23] *A matrix $M \in \Re^{n\times n}$ is said to be*
(1) a P-matrix, if there is an index i such that

$$s_i \neq 0, \text{ and } (Ms)_i \cdot s_i > 0 \text{ for all } s \neq 0 \in \Re^n.$$

(2) a P_0-matrix, if there is an index i such that

$$s_i \neq 0, \text{ and } (Ms)_i \cdot s_i \geq 0 \text{ for all } s \neq 0 \in \Re^n.$$

It is well known any positive definite matrix is a P-matrix, and any positive semi-definite matrix is a P_0-matrix. For the vertical block matrix N, we define

Definition 2. [32] *A square submatrix of N of order n is called a representative submatrix, if its ith row is drawn from the ith block N^i of N, for $i = 1, \ldots, n$. Hence, a vertical block matrix of type $(m_1, \ldots, m_n)$ has at most $\prod_{j=1}^n m_j$ distinct representative submatrices.*

Definition 3. [32] *A vertical block matrix N of type $(m_1, \ldots, m_n)$ is called a vertical block P-matrix (P_0-matrix), if all its representative submatrices are P-matrices (P_0-matrices).*

The properties of the matrix N play an important role in the analysis of GLCP. In [40], Sznajder and Gowda presented some properties of the vertical block P(or P_0) matrix in the case that $m_i = m_j$. In what follows we give some results about general vertical block P(or P_0) matrix N . Denote N_i^j be the ith row of the block matrix N^j, e_i^T be the ith row of the unit matrix I, $v_j = (v_{j,0}, v_{j,1}, v_{j,2}, \ldots, v_{j,m_j})^T \in \Re^{m_j+1}$. We have

Lemma 1. *Suppose that N is a vertical block P_0 matrix and the matrix G is defined as below*

$$G = \begin{pmatrix} v_{1,0}e_1^T + \sum_{i=1}^{m_1} v_{1,i} N_i^1 \\ v_{2,0}e_2^T + \sum_{i=1}^{m_2} v_{2,i} N_i^2 \\ \vdots \\ v_{n,0}e_n^T + \sum_{i=1}^{m_n} v_{n,i} N_i^n \end{pmatrix}. \tag{6}$$

If $v_j \in \Re^{m_j+1}_{++}$*, then G is nonsingular.*

Proof. Suppose to the contrary that G is singular. Then there exists $x \neq 0 \in \Re^n$ such that $Gx = 0$. It follows that

$$\sum_{i=1}^{m_j} x_j v_{j,i} N_i^j x = -v_{j,0} x_j^2 \leq 0, \;\; j = 1, 2, \ldots, n, \tag{7}$$

and the above inequality strictly holds if $x_j \neq 0$. Since $v_j \in \Re^{m_j+1}_{++}$, there exists an index $j_i \in \{1, 2, \ldots, m_j\}$ such that

$$x_j N_{j_i}^j x < 0, \quad \forall\, x_j \neq 0. \tag{8}$$

Denote $\bar{N} \in \Re^{n\times n}$ the matrix whose jth row is $N_{j_i}^j$. Then $\bar{N}$ is a representative submatrix of N. By the assumption of the lemma, $\bar{N}$ is a P_0 matrix. So for any $x \neq 0 \in \Re^n$, there exists an index j such that

$$x_j N_{j_i}^j x \geq 0, \;\; x_j \neq 0, \tag{9}$$

which contradicts to (8). This shows the lemma is true.

□

If N is a vertical block P-matrix, then we have

Lemma 2. *Suppose that N is a vertical block P matrix and the matrix G is defined by (6). If* $v_j \in \Re^{m_j+1}_{+}$ *and that* $\bar{v}_j = \sum_{i=0}^{m_j} v_{j,i} > 0$ *, then G is nonsingular.*

Proof. The proof of this lemma is similar to that of Lemma 1, for completeness, we give it as follows. Assume that G is singular. Then there must exist $x \neq 0 \in \Re^n$ such that $Gx = 0$. It follows that

$$\sum_{i=1}^{m_j} x_j v_{j,i} N_i^j x = -v_{j,0} x_j^2 \leq 0, \;\; j = 1, 2, \ldots, n, \tag{10}$$

Since $v_j \in \Re^{m_j+1}_{+}$ and that $\bar{v}_j = \sum_{i=0}^{m_j} v_{j,i} > 0$, there exists an index $j_i \in \{1, 2, \ldots, m_j\}$ such that

$$x_j N_{j_i}^j x \leq 0, \quad j = 1, 2, \ldots, n. \tag{11}$$

Let $\bar{N} \in \Re^{n\times n}$ be a matrix whose jth row is $N_{j_i}^j$, it is a representative submatrix of N. By the assumption of the lemma, $\bar{N}$ is a P matrix. So for any $x \neq 0 \in \Re^n$, there exists an index i such that

$$x_j N_{j_i}^j x > 0, \;\; x_j \neq 0, \tag{12}$$

which contradicts to (11). This completes the proof of the lemma.

□

Let us denote $W^i(x) = N^i x + q^i,\ i = 1, ..., n$. We can rewrite the GLCP (1) as follows

$$x_i \prod_{j=1}^{m_i} W_j^i(x) = 0, \quad x \geq 0,\ W^i(x) \geq 0,\ i = 1, ..., n. \tag{13}$$

Clearly, (13) is equivalent to the following nonsmooth equations

$$H(x) = \begin{pmatrix} \min\{x_1, W_1^1(x), ..., W_{m_1}^1(x)\} \\ \min\{x_2, W_1^2(x), ..., W_{m_2}^2(x)\} \\ \vdots \\ \min\{x_n, W_1^n(x), ..., W_{m_n}^n(x)\} \end{pmatrix}$$

$$= -\begin{pmatrix} \max\{-x_1, -W_1^1(x), ..., -W_{m_1}^1(x)\} \\ \max\{-x_2, -W_1^2(x), ..., -W_{m_2}^2(x)\} \\ \vdots \\ \max\{-x_n, -W_1^n(x), ..., -W_{m_n}^n(x)\} \end{pmatrix} = 0. \tag{14}$$

Because $W_j^i(x)$ are all linear functions, $H(x)$ is a piecewise linear system of equations. By the definition of $H(x)$, it is easy to see that all $-H_i(x), i \in \{1, ..., n\}$ are piecewise linear convex functions.

Denote

$$T = \{x \in \Re^n :\ H(x) = 0\} \tag{15}$$

the solution set of the GLCP. Let us define $dist(x, T)$ the distance from x to T as follows

$$dist(x, T) = \min_{y \in T} \|y - x\|. \tag{16}$$

Since the graphs of $H(x)$ are unions of finitely many polyhedral convex sets, $H(x)$ is a polyhedral multifunctions [37]. Our next result says that the norm of $H(x)$ plays as a local error bound for a GLCP if the solution set T of the GLCP is not empty .

Lemma 3. *Suppose T is not empty, then there exist constants $e, \tau > 0$ such that*

$$dist(x, T) \leq \tau ||H(x)||$$

for any $x \in \{x \in \Re^n :\ ||H(x)|| \leq e\}$.

Proof. This lemma is a direct consequence of Proposition 1 in [37], thus the proof is omitted here.

□

We next study the growth behavior of the norm of $H(x)$ under certain conditions. To continue our analysis, we need the following definition which is a generalization of the LCP with a R_0 matrix.

Definition 4. [21] *We say that the GLCP with the vertical block matrix N is of type R_0 if*

$$\bar{H}(x)=\begin{pmatrix}\min\{x_1,\, N_1^1 x,\, ...,\, N_{m_1}^1 x\}\\ \vdots \\ \min\{x_n,\, N_1^n x,\, ...,\, N_{m_n}^n x\}\end{pmatrix}=0 \Longleftrightarrow x=0.$$

Now we have

Lemma 4. *GLCP (1) is of type R_0 if and only if*

$$\lim_{\|x\|\to\infty}\frac{\|H(x)\|}{\|x\|}\geq C_0 \tag{17}$$

holds for some constant $C_0>0$.

Proof. $\Leftarrow$: Since the GLCP is of type R_0, it holds $\bar{C_0}=\inf_{||x||=1}||\bar{H}(x)||>0$. If (17) does not hold, i.e., there exists a sequence $\{x^k\}$ such that

$$\lim_{k\to\infty}\|x^k\|\to\infty,\ \text{and}\quad \lim_{k\to\infty}\frac{||H(x^k)||}{\|x^k\|}=0. \tag{18}$$

Let us denote $\bar{x}^k=\frac{x^k}{\|x^k\|}$, then $\{\bar{x}^k\}$ is a bounded sequence. By choosing a subsequence if necessary, we can assume that the sequence $\{\bar{x}^k\}$ converges to an accumulation point $\bar{x}^*$. It follows that

$$\lim_{k\to\infty}\frac{||H(x^k)||}{\|x^k\|}=\lim_{k\to\infty}||H(x^k/||x^k||)||=\|\bar{H}(\bar{x}^*)\|\geq\bar{C_0}. \tag{19}$$

The above inequality contradicts to (18).

$\Rightarrow$: Now we assume that (17) is true. It follows

$$\lim_{||x||\to\infty}||H(x)||=\infty. \tag{20}$$

Suppose that the GLCP is not of type R_0, i.e., there exists an $x\neq 0\in\Re^n$ such that

$$\min\{x_i,\, N_1^i x,\, ...,\, N_{m_i}^i x\}=0,\ i=1,2,...,n.$$

Then we have

$$\lim_{t\to\infty}||H(tx)||<\infty$$

which contradicts (20). This completes the proof of the lemma.

□

The following assumptions will be used numerously in the rest part of this paper.

Assumptions.
(A1): *T is nonempty.*
(A2): *GLCP is of type R_0.*

Now we can give one of our main results in this section.

Theorem 1. *Suppose the Assumptions (A1) and (A2) are true. Then there exists a constant* $\tau_1 > 0$ *such that*

$$dist(x, T) \leq \tau_1 \|H(x)\|, \quad \forall x \in \Re^n. \tag{21}$$

Proof. Suppose to the contrary that the theorem is false. There exists a point sequence $\{x^k\}$ such that (21) is violated, i.e.,

$$dist(x^k, T) > b^k \|H(x^k)\|, \tag{22}$$

where $\{b^k\}$ is a constant sequence satisfying

$$\lim_{k\to\infty} b^k = \infty. \tag{23}$$

Hence there exists a sufficiently large K such that

$$b^k > \tau, \quad \forall k \geq K,$$

where τ is the same constant as defined in Lemma 3. By Lemma 3, we have

$$\|H(x^k)\| > e, \quad \forall k > K, \tag{24}$$

where e is also the constant defined in Lemma 3. Since T is nonempty, for any fixed point $\bar{x} \in T$ we obtain from the definition of $dist(x, T)$ that

$$||x^k - \bar{x}|| \geq dist(x^k, T) > b^k \|H(x^k)\| \geq b^k e. \tag{25}$$

From (23) and (25) we obtain

$$\lim_{k\to\infty} \|x^k - \bar{x}\| = +\infty.$$

This means $\{x^k\}$ is an unbounded point sequence. Now it follows immediately from (25) that

$$\frac{\|H(x^k)\|}{\|x^k\|} = \frac{\|H(x^k)\|}{\|x^k - \bar{x}\|} \cdot \frac{\|x^k - \bar{x}\|}{\|x^k\|} \leq \frac{\|x^k - \bar{x}\|}{b^k \|x^k\|}. \tag{26}$$

The above inequality gives

$$\lim_{k\to\infty} \frac{\|H(x^k)\|}{\|x^k\|} \leq \lim_{k\to\infty} \frac{\|x^k - \bar{x}\|}{b^k \|x^k\|} = 0, \tag{27}$$

where the last equality follows from (23). (27) contradicts to Lemma 4 because the GLCP is of type R_0. So (22) is not true. This completes the proof of the theorem. □

Since a GLCP with a vertical block P-matrix is of type R_0 and its solution is also unique [32], we get the following result as a direct consequence of the above theorem.

Corollary 1. *Suppose the vertical matrix* N *in GLCP (1) is a* P*-matrix and* x^* *is the unique solution of (1). Then there exists a constant* $\tau_2 > 0$ *such that*

$$\|x - x^*\| \leq \tau_2 \|H(x)\|, \quad \forall x \in \Re^n. \tag{28}$$

3. Chen-Mangasarian's neural network smooth function and the function $g(x, t)$

The plus function

$$z_+ = \max\{z, 0\}, \quad z \in \Re$$

is widely used in the reformulations of complementarity problems. In [7], Chen and Mangasarian introduced a class of smoothing function to approximate this function by twice integrating a parameterized probability density function. Chen-Mangasarian's function is defined as below

$$p_t(z) = \int_{-\infty}^{z} \int_{-\infty}^{t} \frac{1}{t} p\left(\frac{\xi}{t}\right) d\xi dt,$$

where $t \in [0, \infty)$ is a parameter, and $p(\xi)$ is a probability density function. Using their smoothing function, Chen and Mangasarian approximate a complementarity problem as a system of smoothing equations

$$x - p_t(x - F(x)) = 0, \quad t > 0.$$

It is easy to verify that when $t \to 0$, $\frac{1}{t} p\left(\frac{\xi}{t}\right)$ is δ-function with all masses concentrated at origin, hence $p_0(z) = \lim_{t \to 0} p_t(z) = z_+$. The following lemma summarizes some properties of the smoothing function $p_t(z)$ [7,9].

Lemma 5. *Suppose the probability density function $p(\xi)$ satisfies the following conditions:*

(A1) *$p(\xi)$ is continuous, symmetric, and has an infinite support, i.e.,*

$$0 < p(\xi) \le \bar{\xi}_1 < \infty, \; p(\xi) = p(-\xi), \quad \forall \xi \in (-\infty, +\infty);$$

(A2) $\int_0^{\infty} \xi p(\xi) d\xi = \bar{\xi}_2 < \infty$.
Then the smooth function $p_t(z)$, which defined in Definition 1 with parameter $t > 0$, has the following properties:
(1) *$p_t(z)$ is continuously differentiable, increasing, and strictly convex with respect to z;*
(2) $0 < p_t'(z) < 1$ *and* $0 < p_t'(-z) = 1 - p_t'(z)$ *for all z;*
(3) $0 < p_t''(z) < \bar{\xi}_1/t$ *for all z;*
(4) $|p_{t_2}(z) - p_{t_1}(z)| \le \bar{\xi}_2 |t_2 - t_1|$ *for all z and $t_1, \; t_2 \ge 0$;*
(5) *If $z \neq 0$, then p_0 is differentiable at z. In addition, $|p_t'(z) - p_0'(z)| \le \bar{\xi}_2 t/|z|$ for all z.*

Particularly, if the probability density function $p(\xi) = \exp(-\xi)/(1 + \exp(-\xi))^2$ is applied, then one has [7,8]

$$p_t(z) = t \ln\left(1 + \exp\left(\frac{z}{t}\right)\right). \tag{29}$$

This function is also known as neural network smooth function. Our next result establish a relation between the function $g(x, t)$ and Chen-Mangasarian's neural network smooth function.

Lemma 6. *Suppose that the aggregate function $g(x,t)$ is defined by (4) with $m = 2$ and that $g_1(x) \equiv 0,\ g_2(x) = x$. If the smoothing function $p_t(x)$ is defined with $p(\xi) = \exp(-\xi)/(1+\exp(-\xi))^2$, then it holds $p_t(x) = g(x,t)$.*

Proof. By (4), we have

$$g(x,t) = t\ln\left(\exp\left(\frac{g_1(x)}{t}\right) + \exp\left(\frac{g_2(x)}{t}\right)\right) = t\ln\left(1+\exp\left(\frac{x}{t}\right)\right). \tag{30}$$

The above equality and (29) imply the lemma is true.

□

Using the smoothing function $p_t(z)$, Chen and Mangasarian [7,8] reformulated CP (2) as the following smooth equations

$$x_i - p_t(x_i - F_i(x)) = 0,\ i = 1,\dots,n.$$

In the case that $p_t(z)$ is given by (29), one has

$$\begin{aligned} x_i - p_t(x_i - F_i(x)) &= x_i - t\ln\left(1+\exp\left(\frac{x_i - F_i(x)}{t}\right)\right) \\ &= -t\left(\ln\exp\left(\frac{-x_i}{t}\right) + \ln\left(1+\exp\left(\frac{x_i - F_i(x)}{t}\right)\right)\right) \\ &= -t\ln\left(\exp\left(\frac{-x_i}{t}\right) + \exp\left(\frac{-F_i(x)}{t}\right)\right). \end{aligned} \tag{31}$$

Since $\min\{x_i, F_i(x)\} = -\max\{-x_i, -F_i(x)\} = 0$. Let $g_1(x) = -x_i$ and $g_2(x) = -F_i(x)$. By applying the function $g(x,t)$ to approximate the function $\min\{x_i, F_i(x)\}$ we get

$$-g(x,t) = -t\ln\left(\exp\left(\frac{-x_i}{t}\right) + \exp\left(\frac{-F_i(x)}{t}\right)\right). \tag{32}$$

It follows from (31) and (32) that, if we use Chen-Mangasarian's neural network smooth function (29) and the function $g(x,t)$ (4) to reformulate a CP, then the same system of equations is derived.

Our following theorem summarizes some interesting properties of the function $g(x,t)$ (4).

Lemma 7. *Suppose $g_i(x)$ are all twice continuously differentiable functions,*

$$g(x) = \max_{i\in\{1,\dots,m\}} g_i(x)$$

and $g(x,t)$ is defined by (4), then we have:

(i) *$g(x,t)$ is increasing with respect to t, and $g(x) \le g(x,t) \le g(x) + t\ln m$;*

(ii) *$g(x,t)$ is twice continuously differentiable for all $t > 0$, and*

$$\nabla_x g(x,t) = \sum_{i=1}^{m} \lambda_i(x,t)\nabla g_i(x),$$

$$\nabla_x^2 g(x,t) = \sum_{i=1}^{m}\left(\lambda_i(x,t)\nabla^2 g_i(x) + \frac{1}{t}\lambda_i(x,t)\nabla g_i(x)\nabla g_i(x)^T\right) - \frac{1}{t}\left(\sum_{i=1}^{m}\lambda_i(x,t)\nabla g_i(x)\right)\left(\sum_{i=1}^{m}\lambda_i(x,t)\nabla g_i(x)\right)^T,$$

where

$$\lambda_i(x,t) = \frac{\exp(g_i(x)/t)}{\sum_{j=1}^{m}\exp(g_j(x)/t)} \in (0,1), \quad \sum_{i=1}^{m}\lambda_i(x,t) = 1,$$

Particularly, if $g_i(x)$ are all linear functions, then $g(x,t)$ is an infinite order differentiable convex function for all $t > 0$, and that

$$\nabla_x^2 g(x,t) = \frac{1}{t}\left(\sum_{i=1}^{m}\lambda_i(x,t)\nabla g_i(x)\nabla g_i(x)^T - \left[\sum_{i=1}^{m}\lambda_i(x,t)\nabla g_i(x)\right]\left[\sum_{i=1}^{m}\lambda_i(x,t)\nabla g_i(x)\right]^T\right).$$

(iii) *For any fixed $x \in \Re^n$,*

$$\nabla_x g(x,0+) = \lim_{t\to 0^+}\nabla_x g(x,t) = \sum_{i\in B(x)}\nabla g_i(x)/\bar{k},$$

where $B(x) = \{i \in \{1,\ldots,m\} : \ g_i(x) = \max_{i\in\{1,\ldots,m\}} g_i(x)\}$, $\bar{k}$ is the element number of the index set $B(x)$.

(iv) *Suppose $g_i(x)$ are all linear functions. For any $x \in \Re^n$, there exists a constant $C_1 > 0$ such that*

$$t\|\nabla_x^2 g(x,t)\| \le C_1, \quad \forall t > 0, \tag{33}$$

and that

$$\lim_{t\to 0^+}\|\nabla_x^2 g(x,t)\| = 0 \tag{34}$$

if $\bar{k} = 1$.

(v) *For any fixed $x \in \Re^n$, $g(x,t)$ is a continuously differentiable, increasing and convex function of t if $t > 0$. Furthermore, we have*

$$g_t'(x,0+) = \lim_{t\to 0^+} g_t'(x,t) = \ln\bar{k}, \tag{35}$$

and that

$$\lim_{t\to 0^+}\frac{g_t'(x,t) - \ln\bar{k}}{t} = 0. \tag{36}$$

(vi) *For any $x \in \Re^n$ and $t > 0$, it holds $\ln \bar{k} \le g_t'(x,t) \le \ln m$.*

Proof. The conclusions **(i)** and **(ii)** have been proven in [30, 41], we need only to prove the statements **(iii)-(vi)**.

By the definition of $\lambda_i(x,t)$, we have

$$\lambda_i(x,t) = \frac{\exp\left(\frac{g_i(x)}{t}\right)}{\sum_{j=1}^m \exp\left(\frac{g_j(x)}{t}\right)} = \frac{\exp\left(\frac{g_i(x)-g(x)}{t}\right)}{\sum_{j=1}^m \exp\left(\frac{g_j(x)-g(x)}{t}\right)}.$$

For any fixed $x \in \Re^n$, it holds

$$\exp\left(\frac{g_i(x)-g(x)}{t}\right) \equiv 1, \quad \forall i \in B(x),$$

and

$$\lim_{t\to 0} \exp\left(\frac{g_i(x)-g(x)}{t}\right) = 0, \quad \forall i \notin B(x). \tag{37}$$

It follows from (37) that

$$\lim_{t\to 0+} \nabla_x g(x,t) = \sum_{i\in B(x)} \nabla g_i(x)/\bar{k}$$

which implies the statement **(iii)**.

We next prove the assertion **(iv)**. By the definition of $\lambda_i(x,t)$, we have

$$\lim_{t\to 0^+} \lambda_i(x,t) = \frac{1}{\bar{k}}, \ \forall i \in B(x), \tag{38}$$

and that

$$\lim_{t\to 0^+} \frac{\lambda_i(x,t)}{t} = 0, \ \forall i \notin B(x). \tag{39}$$

Since all $\lambda_i(x,t) \in (0,1)$, and all $g_i(x)$ are linear, it follows directly from (38), (39) and **(ii)** that the result **(iv)** is true.

Now we turn to case **(v)**. By direct algebraic calculus, we have

$$g_t''(x,t) = \frac{1}{t^3}\left(\sum_{i=1}^m \lambda_i(x,t)\cdot g_i^2(x) - \left(\sum_{i=1}^m \lambda_i(x,t)\cdot g_i(x)\right)^2\right) \ge 0 \tag{40}$$

where the inequality follows from the convexity of the function t^2 and the fact that

$$\sum_{i=1}^m \lambda_i(x,t) = 1, \quad \lambda_i(x,t) \ge 0.$$

Hence $g'_t(x,t)$ is an increasing function of t, and $g(x,t)$ is convex with respect to t. To prove (36), we observe

$$\begin{aligned} g'_t(x,t) &= \ln \sum_{i=1}^m \exp\left(\frac{g_i(x)}{t}\right) - \frac{\sum\limits_{i=1}^m \exp\left(\frac{g_i(x)}{t}\right)\cdot \frac{g_i(x)}{t}}{\sum\limits_{j=1}^m \exp\left(\frac{g_j(x)}{t}\right)} \\ &= \ln[\bar{k} + \eta_1(x,t)] + \frac{g(x)\frac{\eta_1(x,t)}{t} - \eta_2(x,t)}{\bar{k} + \eta_1(x,t)} \end{aligned}$$

where

$$\eta_1(x,t) = \sum_{i\notin B(x)} \exp\left(\frac{g_i(x) - g(x)}{t}\right),\ \eta_2(x,t) = \sum_{i\notin B(x)} \frac{g_i(x)}{t}\exp\left(\frac{g_i(x) - g(x)}{t}\right).$$

For any fixed $x \in \Re^n$, one can easily verify that

$$\lim_{t\to 0^+} \eta_1(x,t)/t^2 = 0, \tag{41}$$

and

$$\lim_{t\to 0^+} \eta_2(x,t)/t = 0. \tag{42}$$

It follows from (41) and (42) that

$$\begin{aligned} \lim_{t\to 0^+} \frac{g'_t(x,t) - \ln \bar{k}}{t} &= \lim_{t\to 0^+} \left\{ \frac{\ln(1 + \eta_1(x,t)/\bar{k})}{t} + \frac{g(x)\frac{\eta_1(x,t)}{t^2} - \frac{\eta_2(x,t)}{t}}{\bar{k} + \eta_1(x,t)} \right\} \\ &= \lim_{t\to 0^+} \frac{\ln(1 + \eta_1(x,t)/\bar{k})}{t} \\ &= \lim_{t\to 0^+} \frac{\ln(1 + \eta_1(x,t)/\bar{k})}{\eta_1(x,t)/\bar{k}} \cdot \frac{\eta_1(x,t)}{\bar{k}t} = 0. \end{aligned} \tag{43}$$

This proves (36). (35) follows directly from (36). (35) and (40) imply that $g'_t(x,t) \geq 0$ for all $t > 0$. It follows that $g(x,t)$ is also an increasing function of t. This completes the proof of the statement **(v)**.

Now we turn to the last conclusion of the lemma. For any $x \in \Re^n$, it is easy to see

$$\lim_{t\to +\infty} \eta_1(x,t) = m - \bar{k},\ \lim_{t\to +\infty} \eta_2(x,t) = 0, \tag{44}$$

which implies that

$$\lim_{t\to +\infty} g'_t(x,t) = \ln m. \tag{45}$$

The above equality and the fact that $g'_t(x,t)$ is an increasing function of t give the conclusion **(vi)**. The proof of the lemma is finished.

□

4. A smooth reformulation of GLCP

In the previous section, we have studied the relations between Chen-Mangasarian's neural network smooth function and the function $g(x,t)$. In this section, we will use the function $g(x,t)$ (4) to smooth $H(x)$. In this way we get the following parameterized smooth equations system

$$H(x,t) = -\begin{pmatrix} t\ln\left(\exp\left(-\frac{x_1}{t}\right) + \sum_{j=1}^{m_1}\exp\left(-\frac{W_j^1(x)}{t}\right)\right) \\ t\ln\left(\exp\left(-\frac{x_2}{t}\right) + \sum_{j=1}^{m_2}\exp\left(-\frac{W_j^2(x)}{t}\right)\right) \\ \vdots \\ t\ln\left(\exp\left(-\frac{x_n}{t}\right) + \sum_{j=1}^{m_n}\exp\left(-\frac{W_j^n(x)}{t}\right)\right) \end{pmatrix}. \tag{46}$$

By the conclusion (i) of Lemma 7, we have

Lemma 8. *Suppose $H(x)$ and $H(x,t)$ are defined by (14) and (46) respectively. Then it holds*

$$H_i(x) \le H_i(x,t) \le H_i(x) + t\ln(m_i+1), \quad i \in \{1,2,...,n\}. \tag{47}$$

By the conclusion (ii) of Lemma 7, we know that $-H_i(x,t)$ are all infinite order differentiable convex function for all $t > 0$.

Let $T(t)$ be the solution set of the equations (46) defined by

$$T(t) = \{x \in \Re^n : \; H(x,t) = 0\}. \tag{48}$$

We have

Theorem 2. *Suppose the Assumptions (A1) and (A2) are true. Then there exist constants $\tau_3, \tau_4 > 0$ such that*

$$\min_{y\in T} ||x-y||_1 \le n\tau_3 t\ln\bar{m}, \quad dist(x,T) \le \sqrt{n}\tau_4 t\ln\bar{m}, \quad \forall x \in T(t), t > 0, \tag{49}$$

where $\bar{m} = \max\{m_1,...,m_n\} + 1$.

Proof. By Theorem 1, for any $x \in T(t)$, there exists a constant τ_1 such that

$$dist(x,T) \le \tau_1||H(x)||. \tag{50}$$

Because all the norms of a vector (or matrix) are in the same order, there exists a constant $\tau_3 > 0$ such that

$$\min_{y\in T} ||x-y||_1 \le \tau_3||H(x)||_1.$$

Since $x \in T(t)$, it follows from (47) that

$$\begin{aligned}\min_{y\in T} ||x-y||_1 &\le \tau_3 \sum_{i=1}^{n} |H_i(x)| = \tau_3 \sum_{i=1}^{n} (-H_i(x)) \\ &\le \tau_3 \sum_{i=1}^{n} (-H_i(x,t) + t \ln \bar{m}) = n\tau_3 t \ln \bar{m}.\end{aligned}$$

By using (50) again, we get

$$\begin{aligned} dist(x,T) &\le \tau_1 \sqrt{\sum_{i=1}^{n} (-H_i(x))^2} \\ &\le \tau_1 \sqrt{\sum_{i=1}^{n} (-H_i(x,t) + t\ln \bar{m})^2} = \sqrt{n}\tau_1 t \ln \bar{m}.\end{aligned}$$

Therefore, the theorem is true.

□

Let us define the smoothing path $\Gamma = \{(x,t) \in \Re^n \times \Re_{++} : \ H(x,t) = 0\}$. We next consider the properties of the Jacobian $\nabla_x H(x,t)$ on the smoothing path Γ. For any $t > 0$, it follows directly from (46) that

$$\nabla_x H(x,t) = \begin{pmatrix} \dfrac{\exp\left(-\dfrac{x_1}{t}\right) e_1^T + \sum_{j=1}^{m_1} \exp\left(-\dfrac{W_j^1(x)}{t}\right) \cdot W_j^1(x)'}{\exp\left(-\dfrac{x_1}{t}\right) + \sum_{j=1}^{m_1} \exp\left(-\dfrac{W_j^1(x)}{t}\right)} \\ \vdots \\ \dfrac{\exp\left(-\dfrac{x_n}{t}\right) e_n^T + \sum_{j=1}^{m_n} \exp\left(-\dfrac{W_j^n(x)}{t}\right) \cdot W_j^n(x)'}{\exp\left(-\dfrac{x_n}{t}\right) + \sum_{j=1}^{m_n} \exp\left(-\dfrac{W_j^n(x)}{t}\right)} \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_0^1 e_1^T + \sum_{j=1}^{m_1} \lambda_j^1 N_j^1 \\ \lambda_0^2 e_2^T + \sum_{j=1}^{m_2} \lambda_j^2 N_j^2 \\ \vdots \\ \lambda_0^n e_n^T + \sum_{j=1}^{m_n} \lambda_j^n N_j^n \end{pmatrix}$$

where

$$\lambda_0^j = \frac{\exp\left(-\frac{x_j}{t}\right)}{\exp\left(-\frac{x_j}{t}\right) + \sum_{l=1}^{m} \exp\left(-\frac{W_l^j(x)}{t}\right)} \in (0,1), \;\; j \in \{1,2,...,n\}, \tag{51}$$

$$\lambda_i^j = \frac{\exp\left(-\frac{W_i^j(x)}{t}\right)}{\exp\left(-\frac{x_j}{t}\right) + \sum_{l=1}^{m_j} \exp\left(-\frac{W_l^j(x)}{t}\right)} \in (0,1), \;\; i \in \{1,2,...,m_i\}, \;\; j \in \{1,2,...,n\}. \tag{52}$$

By Lemma 1, we have

Lemma 9. *Suppose GLCP is defined with a vertical block P_0-matrix N. If $t > 0$, then the matrix $\nabla_x H(x,t)$ is nonsingular.*

Let κ be the closed set of matrices defined by

$$\kappa = co\{\nabla_x H(x,t), t > 0\} \tag{53}$$

where *co* denotes the convex hull of a set. Since all $\lambda_i^j \in (0,1)$ and that

$$\sum_{i=0}^{m_j} \lambda_i^j = 1, \quad j = 1, 2, \ldots, n. \tag{54}$$

It follows directly from Lemma 2 that

Lemma 10. *Suppose GLCP is defined with a vertical block P-matrix N. Then any $M \in \kappa$ is nonsingular.*

Remark. The above lemma implies that if GLCP is defined with a vertical block P-matrix N, then the matrix $\nabla_x H(x,t)$ is nonsingular for all $t > 0$. Furthermore, it is easy to see that κ is a closed bounded convex set. Since all the matrices in κ are nonsingular, it follows that $\|M^{-1}\|$ is bounded above for any $M \in \kappa$.

In what follows we consider the derivatives of $H(x,t)$ with respect to t. Let $W_0^i(x) = x_i$. Denote $B^i(x)$ the active index set defined by

$$B^i(x) = \{j | H_i(x) = W_j^i(x), \;\; j = 0, 1, ..., m_i.\} \tag{55}$$

and $\bar{k}_i$ the element number of $B^i(x)$. Then by the result (v) and (vi) of Lemma 7, we get

Lemma 11. *For any $i \in \{1, 2, \ldots, n\}$, $H_i(x,t)$ is a continuously differentiable, decreasing and concave function of t. Furthermore, it holds*

$$\lim_{t \to 0^+} \frac{1}{t} \left(\frac{dH_i(x,t)}{dt} + \ln \bar{k}_i \right) = 0. \tag{56}$$

and that

$$-\ln(m_i + 1) \le \frac{dH_i(x,t)}{dt} \le -\ln \bar{k}_i. \tag{57}$$

A direct consequence of the above lemma is the following corollary which will be used repeatedly in the rest of the paper.

Corollary 2. *For any $x \in \Re^n$ and $t_1, t_2 \geq 0$, it holds*

$$|t_1 - t_2| \ln \bar{k}_i \leq |H_i(x, t_1) - H_i(x, t_2)| \leq |t_1 - t_2| \ln(m_i + 1) \\ \leq |t_1 - t_2| \ln \bar{m}, \quad i = 1, 2, \ldots, n. \tag{58}$$

Our next result consider the relationships between the smoothing path Γ and the mapping $H(x)$.

Theorem 3. *Suppose $H(x)$ and $H(x, t)$ are defined by (14) and (46) respectively. If $(x(t), t) \in \Gamma$, then we have*

$$t \ln \bar{k}_i \leq H_i(x(t)) \leq -t \frac{dH_i(x(t), t)}{dt} \leq t \ln(m_i + 1), \; i \in \{1, 2, ..., n\}, \tag{59}$$

and that

$$\lim_{t \to 0^+} \frac{H_i(x(t)) - t \ln \bar{k}_i}{t^2} = 0. \tag{60}$$

Particularly , if $\bar{k}_i \equiv 1$, then

$$\lim_{t \to 0^+} \frac{\|H(x(t))\|}{t^2} = 0.$$

Proof. Since $(x(t), t) \in \Gamma$, we have $H_i(x(t), t) = 0$. By Lemma 11, all $H_i(x, t)$ are concave functions of t. For any $i \in \{1, 2, \ldots, n\}$, it follows

$$H_i(x(t), \bar{t}) + \frac{dH_i(x(t), \bar{t})}{dt}(t - \bar{t}) \geq H_i(x(t), t) \\ \geq H_i(x(t), \bar{t}) + \frac{dH_i(x(t), t)}{dt}(t - \bar{t}), \quad \forall 0 < \bar{t} \leq t.$$

Taking limits $\bar{t} \to 0$ in both sides of the above inequality, we get

$$H_i(x(t)) + t \frac{dH_i(x(t), 0+)}{dt} \geq 0 \geq H_i(x(t)) + \frac{dH_i(x(t), t)}{dt} t \geq H_i(x(t)) - t \ln(m_i + 1), \tag{61}$$

where the last inequality follows from (57). (61) and (56) gives (59). (60) follows from (59) and (56).

□

5. A path-following algorithm and its global convergence

In last section, we have reformulated an equivalent system of GLCP as a system of smooth parameterized equations. It is easy to see that, when the parameter t tends to zero, any accumulation point of the smoothing path is a solution point of the GLCP. In this section, we present a noninterior continuation method to follow this smoothing path Γ and analyse the global convergence of the algorithm. First we introduce some neighborhoods around the smoothing path.

Definition 5. *(1) A β neighbourhood around Γ is defined as $\mathcal{N}(\beta) = \{(x,t) \in \Re^n \times (0, t_0] : ||H(x,t)|| \le \beta \min\{t, 1\}\}$, where $\beta > 0$ is called the width of the neighborhood $\mathcal{N}(\beta)$, and $t_0 > 0$ is an initial parameter.*
(2) For any $t > 0$, $\mathcal{N}(\beta, t) = \{(x,t) \in \Re^n \times \Re_+ : ||H(x,t)|| \le \beta \min\{t, 1\}\}$.

The algorithm can be stated as follows.

Algorithm 1. Given constant numbers $\epsilon_0 \ge 0, \sigma \in (0,1), \ \alpha_i \in (0,1), \ i = 1,2, \ \alpha_3 \in (0, 1-\alpha_2), \ \beta > 0$, an initial parameter $t_0 > 0$, initial point $(x^0, t_0) \in \mathcal{N}(\beta, t_0)$ and iterative number $k := 0$;

Step 1. *The Newton step of $H(x, t_k) = 0$ at x^k:*
If $\nabla_x H(x^k, t_k)$ is singular, stop (The algorithm fails);
else if $\|H(x^k)\| \le \epsilon_0$, stop, x^k is an approximate solution of GLCP;
Otherwise, compute a Newton step Δx^k satisfying

$$\nabla_x H(x^k, t_k)\Delta x^k + H(x^k, t_k) = 0; \tag{62}$$

Step 2. *Compute x^{k+1}: Let h_k be the maximum value of $\{1, \alpha_1, \alpha_1^2, \dots\}$ such that*

$$||H(x^k + h_k \Delta x^k, t_k)|| \le (1 - \sigma h_k) \min\{t_k, 1\}\beta, \tag{63}$$

and $x^{k+1} = x^k + h_k \Delta x^k$;

Step 3. *Compute t_{k+1}:*
If $(x^{k+1}, \min\{\alpha_3, t_k\}t_k) \in \mathcal{N}(\beta, \min\{\alpha_3, t_k\}t_k)$, then we set $\nu_k = 1 - \min\{\alpha_3, t_k\}$;
Otherwise ν_k be the maximum value of $\{\alpha_2, \alpha_2^2, \dots\}$ such that

$$\left(x^{k+1}, (1-\nu_k)t_k\right) \in \mathcal{N}(\beta, (1-\nu_k)t_k), \tag{64}$$

set $t_{k+1} = (1 - \nu_k)t_k$;

Step 4. *$k := k + 1$, return to Step (1).*

Remark (a). It is very easy to initialize the above method. One may simply choose any $t_0 > 0$, $x^0 \in \Re^n$, and $\beta \ge \frac{\|H(x^0, t_0)\|}{\min\{t_0, 1\}}$.
Remark (b). In Step 3, by $\alpha_3 \in (0, 1-\alpha_2)$, we have $1 - \min\{\alpha_3, t_k\} \in (1-\alpha_3, 1) \subset (\alpha_2, 1)$, then

$$1 - \min\{\alpha_3, t_k\} > \alpha_2 > \alpha_2^2 > \cdots;$$

Remark (c). The Step 1 and 2 of our method is similar to that of the method presented by Chen and Xiu [9]. But we do not calculate approximate Newton step such as in [9]. The iteration of t is also different from that in [9]. As we will see later, this new reducing step of t plays an important role in the analysis of the local convergence of the algorithm.

Next, we discuss the global convergence of Algorithm 1. First we give a result about the mapping $H(x, t)$.

Lemma 12. *Let $H(x, t)$ is defined by (46). Then for any $x, y \in \Re^n$ and $t > 0$, there exists a constant $C_2 > 0$ such that*

$$||H(y, t) - H(x, t) - \nabla_x H(x, t)(y - x)|| \le \frac{nC_2}{2t}||y - x||^2. \tag{65}$$

Proof. By Lemma 7, we know that all $H_i(x, t)$ are twice differentiable whenever $t > 0$. It follows that

$$H(y, t) = H(x, t) + \nabla_x H(x, t)(y - x) + \frac{1}{2}\begin{pmatrix} (y - x)^T \nabla_x^2 H_1(x + \bar{\xi}_1(y - x), t)(y - x) \\ (y - x)^T \nabla_x^2 H_2(x + \bar{\xi}_2(y - x), t)(y - x) \\ \vdots \\ (y - x)^T \nabla_x^2 H_n(x + \xi_n(y - x), t)(y - x) \end{pmatrix}. \tag{66}$$

where $\xi_i \in (0, 1)$. By the result (iv) of Lemma 7, there exists a constant $C_2 > 0$ such that

$$\|\nabla_x^2 H_i(x + \xi_i(y - x))\| \le \frac{C_2}{t}, \quad i = 1, 2, \ldots, n. \tag{67}$$

The above inequality and (66) prove (65).

□

Our next results study the properties about the direction Δx updated by (62).

Lemma 13. *Suppose $\nabla_x H(x, t)$ is nonsingular for some $x \in \Re^n$ with $t > 0$ and Δx is a solution of (62) at x. Then for any $h \in (0, 1]$,*

$$\|H(x + h\Delta x, t)\| \le (1 - h)\|H(x, t)\| + \frac{nC_2}{2t}h^2\|\Delta x\|^2. \tag{68}$$

Proof. Let $y = x + h\Delta x$. Since Δx is a solution of (62) at x, it holds

$$H(y, t) - H(x, t) - \nabla_x H(x, t)(y - x) = H(y, t) - (1 - h)H(x, t).$$

The above equation and (65) yield (68).

□

The following assumption play an important role in the analysis of the global linear convergence of the algorithm.

Assumption. (A3) The matrix $\nabla_x H(x^k, t_k)$ is nonsingular for all k. In addition, there exists a constant C such that $\|\nabla_x H(x^k, t_k)^{-1}\| \le C$ for all k.

Remark. If GLCP (1) is defined with a vertical block P_0-matrix, it follows from Lemma 9 that all $\nabla_x H(x^k, t^k)$ are nonsingular if $t^k > 0$. For GLCP with a vertical block P-matrix, Lemma 10 implies (A3) is true.

If (A3) is true, then

$$\|\Delta x^k\| = \|\nabla_x H(x^k, t_k)^{-1} H(x^k, t_k)\| \le C\|H(x^k, t_k)\| \le \beta C t_k. \tag{69}$$

Our next result consider the line search step h_k under the condition (A3).

Lemma 14. *Let (x^k, t_k) be the kth iteration of the algorithm. If (A3) is true, then there exists an independent constant $\bar{h} > 0$ such that $h_k \ge \bar{h}$.*

Proof. If $H(x^k, t_k) = 0$, then $\Delta x^k = 0$. It follows from the line search rule in Step 2 that $h_k = 1$.

Assume that $H(x^k, t_k) \ne 0$. For any $h \in (0, 1]$, it follows from (68) that

$$\begin{aligned} ||H(x^k + h\Delta x^k, t_k)|| &\le (1-h)||H(x^k, t_k)|| + \frac{nC_2}{2t_k} h^2 ||\Delta x^k||^2 \\ &\le \left(1 - h + \frac{nC_2}{2t_k} C h^2 ||\Delta x^k||\right) ||H(x^k, t_k)|| \\ &\le \left(1 - \left(1 - \frac{nC_2}{2}\beta C^2 h\right) h\right) ||H(x^k, t_k)||, \end{aligned}$$

where the last two inequalities follow from (69). Let $\hat{h} = \frac{2(1-\sigma)}{nC_2 \beta C^2}$. It is easy to see that for all $h \in (0, \hat{h}]$, we have

$$||H(x^k + h\Delta x^k, t_k)|| \le (1 - \sigma h)||H(x^k, t_k)||. \tag{70}$$

Let $\bar{h} = \min\{1, \alpha_1 \hat{h}\}$, it is independent of k. Clearly, we have $h_k \ge \bar{h}$.

□

We next show the step length ν_k for reducing t is also bounded below by a positive constant.

Lemma 15. *Let (x^k, t_k) be the kth iteration of the algorithm. If (A3) is true, then there exists an independent constant $\bar{\nu} > 0$ such that $\nu_k \ge \bar{\nu}$.*

Proof. From Corollary 2 and triangle inequality we get

$$\begin{aligned} \|H(x^{k+1}, t)\| &\le \|H(x^{k+1}, t_k)\| + \|H(x^{k+1}, t_k) - H(x^{k+1}, t)\| \\ &\le \|H(x^{k+1}, t_k)\| + \mu(t_k - t), \end{aligned} \tag{71}$$

where $\mu > 0$ is a constant derived from Corollary 2. It follows from (63) that

$$||H(x^{k+1}, t)|| \leq (1 - \sigma h_k) \min\{t_k, 1\}\beta + \mu(t_k - t). \quad (72)$$

By (72) and the line search rule , there exists $\nu_k > 0$ satisfying to

$$||H(x^{k+1}, (1 - \nu_k)t_k)|| \leq (1 - \sigma h_k) \min\{t_k, 1\}\beta + \mu\nu_k t_k \leq \min\{(1 - \nu_k)t_k, 1\}\beta.$$

There are three cases:

(i) If $(x^{k+1}, \min\{\alpha_3, t_k\}t_k) \in \mathcal{N}(\beta, \min\{\alpha_3, t_k\}t_k)$, then $\nu_k = 1 - \min\{\alpha_3, t_k\} \geq 1 - \alpha_3 > 0$;
(ii) If $(x^{k+1}, (1 - \alpha_2)t_k) \in \mathcal{N}(\beta, (1 - \alpha_2)t_k)$, then we have $\nu_k = \alpha_2$;
(iii) Otherwise, by the line search rule in the algorithm, we have

$$(x^{k+1}, t_{k+1}) \in \mathcal{N}(\beta, t_{k+1}), \ (x^{k+1}, t_k - (t_k - t_{k+1})/\alpha_2) \notin \mathcal{N}(\beta, t_k - (t_k - t_{k+1})/\alpha_2),$$

i.e.,

$$||H(x^{k+1}, t_k - (t_k - t_{k+1})/\alpha_2)|| \geq \min\{t_k - (t_k - t_{k+1})/\alpha_2, 1\}\beta.$$

By (72), it holds

$$\min\{t_k - (t_k - t_{k+1})/\alpha_2, 1\}\beta \leq (1 - \sigma h_k) \min\{t_k, 1\}\beta + \mu\frac{t_k - t_{k+1}}{\alpha_2}, \quad (73)$$

It follows that

$$t_k - t_{k+1} \geq \begin{cases} \dfrac{\alpha_2\sigma h_k\beta}{\mu}, & \mathit{if} \quad t_k - (t_k - t_{k+1})/\alpha_2 \geq 1, \ t_k > 1; \\ \dfrac{(t_k - 1 + \sigma h_k)\alpha_2\beta}{\beta + \mu}, & \mathit{if} \quad t_k - (t_k - t_{k+1})/\alpha_2 \leq 1, \ t_k > 1; \\ \dfrac{t_k\alpha_2\sigma h_k\beta}{\beta + \mu}, & \mathit{if} \quad t_k \leq 1, \end{cases} \quad (74)$$

Since $h_k \geq \bar{h}$, by (74) , we have

$$\frac{t_k - t_{k+1}}{t_k} \geq \frac{t_k - t_{k+1}}{t_0} \geq \frac{\sigma\bar{h}\alpha_2\beta}{t_0(\beta + \mu)}, \quad \mathit{whenever} \ \ t_k > 1, \quad (75)$$

and that

$$\frac{t_k - t_{k+1}}{t_k} \geq \frac{\alpha_2\sigma\bar{h}\beta}{\beta + \mu}, \quad \mathit{if} \quad t_k \leq 1. \quad (76)$$

Let

$$\bar{\nu} = \min\left\{\frac{\alpha_2\sigma\bar{h}\beta}{t_0(\beta + \mu)}, \frac{\alpha_2\sigma\bar{h}\beta}{\beta + \mu}, 1 - \alpha_3, \alpha_2\right\}$$

which is a constant independent of k. It is easy to see that $\nu_k \geq \bar{\nu}$.

□

Now we are ready to show the global linear convergence of the continuation method. We assume the algorithm does not terminate finitely and that $\epsilon_0 = 0$, then we have

Theorem 4. *Suppose (A3) is true for the infinite sequence* $\{x^k, t_k\}$ *generated by the algorithm with* $\epsilon_0 = 0$. *Then*
1. For all $k = 1, 2, \ldots,$ *we have*

$$t_k \leq t_0(1 - \bar{\nu})^k. \tag{77}$$

Namely the sequence $\{t_k\}$ *converges to zero global Q-linearly.*
2. The sequence $\{\|H(x^k)\|\}$ *converges to zero globally and r-linearly.*
3. The sequence $\{x^k\}$ *is bounded and converges to a solution of the GLCP.*

Proof. By Lemma 15, $\nu_k \geq \bar{\nu}$ at each iteration. Therefore,

$$t_{k+1} = (1 - \nu_k)t_k \leq (1 - \bar{\nu})t_k, \quad \forall k = 1, 2, \ldots, \tag{78}$$

Result 1 then follows immediately.

For result 2, we have

$$\|H(x^k)\| \leq \|H(x^k, t_k)\| + \|H(x^k) - H(x^k, t_k)\| \leq \beta t_k + nt_k \ln \bar{m}, \tag{79}$$

where the last inequality follows from Lemma 8 and $\bar{m} = \max\{m_1, m_2, \ldots, m_n\} + 1$ as defined in Section 4. Then results follows from result 1.

For result 3, since

$$\begin{aligned} \|x^{k+1} - x^k\| &= \|h_k \Delta x^k\| \leq \|\Delta x^k\| \leq C\|H(x^k, t_k)\| \\ &\leq C\beta t_k \leq C\beta t_0(1 - \bar{\nu})^k, \end{aligned} \tag{80}$$

where the first inequality follows from the fact that $h_k \leq 1$. (80) implies that $\{x^k\}$ is a Cauchy sequence, so it is bounded and has a unique accumulation point x^*. By result 1, we have $H(x^*) = 0$. Therefore x^* is a solution of GLCP.

□

Remark. If the GLCP is defined with a vertical block P-matrix which implies (A3) is true, the above theorem shows that our method is well-defined and has a global linear convergence.

If the GLCP is defined with a vertical block P_0-matrix and of type R_0, then we have the following result.

Theorem 5. *Suppose the GLCP is defined with a vertical block* P_0*-matrix and of type* R_0, *Then the sequence* $\{x^k, t_k\}$ *updated by the algorithm with* $\epsilon_0 = 0$ *is bounded , and any accumulation point is a solution of GLCP.*

Proof. Since the GLCP has a vertical block p_0-matrix, it follows from Lemma 9 that the matrix $\nabla_x H(x, t)$ is nonsingular for all $t > 0$ and $x \in \Re^n$. Hence the algorithm is well-defined.

We now show the slice of neighborhood $\mathcal{N}(\beta)$ is bounded for GLCP of type R_0. For any $(x, t) \in \mathcal{N}(\beta)$, it follows from Lemma 8 that

$$\begin{aligned} \|H(x)\| &\leq \|H(x) - H(x, t)\| + \|H(x, t)\| \\ &\leq nt \ln \bar{m} + \beta t \leq (n \ln \bar{m} + \beta)t_0. \end{aligned} \tag{81}$$

It follows from Lemma 4 that $\mathcal{N}(\beta)$ is bounded. So the sequence $\{x^k, t_k\}$ has at least an accumulation point, say (x^*, t^*). Since t_k decreases monotonically, we have $t_k \to t^* \geq 0$. If $t^* > 0$ and $\{x^k\}$ is bounded, it follows from Lemma 9 that the matrix $\nabla_x H(x^k, t_k)$ is uniformly bounded above which implies (A3) is true and so $t_k \to 0$. This leads to a contradiction. So it holds $t^* = 0$. It follows that $H(x^*) = 0$ and x^* is a solution of GLCP.

□

6. Local convergence result

In this section, we discuss the local convergence properties of the algorithm. Assume x^* is a solution point of GLCP. We say x^* is a strictly complementarity solution if $k_i^* = 1$ for any $i \in \{1, 2, \ldots, n\}$ where k_i^* is the element number of the index set $B^i(x^*)$ defined by (55) in Sect. 4. If x^* is a strictly complementarity solution of GLCP, then there exists a neighborhood $\Omega(x^*, \epsilon) = \{x | \|x - x^*\| \leq \epsilon\}$ of x^* such that

$$\bar{k}_i = 1, \quad \forall x \in \Omega(x^*, \epsilon),\ \forall i \in \{1, 2, \ldots, n\}$$

where $\bar{k}_i$ is the same as defined in Sect. 4. It follows immediately that $H(x)$ is differentiable for any $x \in \Omega(x^*, \epsilon)$ and that $\nabla_x H(x) = \nabla_x H(x^*)$. Now we give a result about the mapping $H(x, t)$ and the Jacobian $\nabla_x H(x, t)$ in the neighborhood $\Omega(x^*, \epsilon)$ of x^*.

Lemma 16. *Suppose x^* is a strictly complementarity solution and $\Omega(x^*, \epsilon)$ be a neighborhood of x^* such that all $\bar{k}_i \equiv 1$ for any $x \in \Omega(x^*, \epsilon)$. If $\nabla H(x^*)$ is nonsingular, there exists a constant $c_1 \in (0, 1)$ such that*
(a) For any $x \in \Omega(x^*, \epsilon)$ and $t \in (0, c_1)$,

$$\|\nabla_x H(x, t)^{-1}\| \leq C. \tag{82}$$

(b) For any $x \in \Omega(x^*, \epsilon)$ and $t \in (0, c_1)$, we have $\|H_t'(x, t)\| \leq \frac{\beta t}{2}$;
(c) For any $t \in (0, c_1)$, then $||H(x, t) - H(x)|| \leq \frac{\beta}{2} t^2$.

Proof. Denote the Jacobian $\nabla_x H(x) = \nabla_x H(x, 0)$, then it follows from (38), (39) and the definition of $\nabla_x H(x, t)$ that for any $x \in \Omega(x^*, \epsilon)$,

$$\lim_{t \to 0^+} \|\nabla_x H(x, t) - \nabla_x H(x)\| = 0. \tag{83}$$

It follows that if $\nabla_x H(x)$ is nonsingular, then $\nabla_x H(x, t)$ is also uniformly nonsingular for any $x \in \Omega(x^*, \epsilon)$ and sufficiently small $t > 0$. For simplicity, we use the same constant C in (82) as in Assumption (A3). This proves the result (a).

To prove result (b), we note that for any $x \in \Omega(x^*, \epsilon)$ and any $i \in \{1, 2, \ldots, n\}$, $\bar{k}_i = 1$. It follows from Lemma 11 that

$$\lim_{t \to 0^+} \frac{\frac{dH_i(x, t)}{dt} + \ln \bar{k}_i}{t} = 0,$$

which implies the result (b).

Now we consider the last result of the lemma. When t is small enough, we have $H_i(x,t) = H_i(x,0+) + \frac{dH_i(x,\xi t)}{dt}t$ for some $\xi \in (0,1)$. By Lemma 11 we know that $H_i(x,t)$ is a concave function of t if $t > 0$. Further, one has

$$\frac{dH_i(x,0+)}{dt} = 0, \quad \forall x \in \Omega(x^*,\epsilon).$$

It follows that

$$\begin{aligned}\lim_{t\to 0^+} \frac{\|H(x,t)-H(x)\|}{t^2} &\le \lim_{t\to 0^+} \left\{ \sum_{i=1}^n \frac{|H_i(x,t)-H_i(x,0)|}{t^2} \right\} \\ &\le \lim_{t\to 0^+} \left\{ \sum_{i=1}^n \frac{\left|\frac{dH_i(x,t)}{dt}\right|}{t} \right\} = 0.\end{aligned}$$

The above relation gives the result (c).

□

Lemma 17. *Suppose x^* is a strictly complementarity solution and $\Omega(x^*,\epsilon)$ is a neighborhood of x^* such that all $\bar{k}_i \equiv 1$ for any $x \in \Omega(x^*,\epsilon)$. Then there exists a constant $c_2 \in (0,1)$ such that for any $x, y \in \Omega(x^*,\epsilon)$ and $t \in (0,c_2)$,*

$$||H(y,t) - H(x,t) - \nabla_x H(x,t)(y-x)|| \le \min\left\{\frac{1}{2C^2\beta}, \frac{1-\sigma}{C^2\beta}, 1\right\} ||y-x||^2. \quad (84)$$

Proof. Since $\bar{k}_i \equiv 1$ for any $x \in \Omega(x^*,\epsilon)$. It follows from result (iv) of Lemma 7 that $\lim_{t\to 0^+} \|\nabla_x^2 H(x,t)\| = 0$ for any $x \in \Omega(x^*,\epsilon)$. Following the proof of Lemma 12, one can easily show that the above lemma is true.

□

Theorem 6. *Suppose x^* is a strictly complementarity solution of GLCP and $\nabla_x H(x^*,0)$ is nonsingular. Suppose that $\Omega(x^*,\epsilon)$ is a neighborhood of x^* such that all $\bar{k}_i \equiv 1$ for any $x \in \Omega(x^*,\epsilon)$. Then there exists a constant $\bar{c}$ such that*

1. *If $x \in \Omega(x^*,\epsilon)$ and $t \in (0,\bar{c})$, then $x + \Delta x \in \Omega(x^*,\epsilon)$.*
2. *The step length updated by the algorithm take the values $h_k = 1,\ \nu_k = 1 - t_k$ and that $x^{k+1} \in \Omega(x^*,\epsilon)$.*

Proof. We first prove the result 1. From the definition of Δx we get

$$\begin{aligned}
\|x + \Delta x - x^*\| &= \|x + \nabla_x H^{-1}(x,t)H(x,t) - x^*\| \\
&\leq \|x + \nabla_x H^{-1}(x)H(x) - x^*\| + \|[\nabla_x H^{-1}(x,t) - \nabla_x H^{-1}(x)]H(x)\| \\
&\quad + \|\nabla_x H^{-1}(x,t)[H(x,t) - H(x)]\| \\
&= \|[\nabla_x H^{-1}(x,t) - \nabla_x H^{-1}(x)]H(x)\| \\
&\quad + \|\nabla_x H^{-1}(x,t)[H(x,t) - H(x)]\| \\
&= \|[\nabla_x H^{-1}(x,t) \cdot \nabla_x H(x) - I](x - x^*)\| \\
&\quad + \|\nabla_x H^{-1}(x,t)[H(x,t) - H(x)]\| \\
&\leq \|\nabla_x H^{-1}(x,t)\| \cdot \|\nabla_x H(x,t) - \nabla_x H(x)\| \cdot \|x - x^*\| \\
&\quad + \|\nabla_x H^{-1}(x,t)\| \cdot \|H(x,t) - H(x)\| \qquad (85)
\end{aligned}$$

where the first inequality is given by triangle inequality, the second and third equalities follow from the nonsingularity of $\nabla_x H(x^*)$ and the fact that $\nabla_x H(x) = \nabla_x H(x^*)$ for all $x \in \Omega(x^*, \epsilon)$, and the last inequality given by the definition of the matrix norm. Now recall (82) and (83), we get

$$\lim_{t \to 0^+} \left\{ \|\nabla_x H^{-1}(x,t)\| \cdot \|\nabla_x H(x,t) - \nabla_x H(x)\| \cdot \|x - x^*\| \right\} = 0.$$

Further, by using Lemma 8 and (82), one has

$$\lim_{t \to 0^+} \left\{ \|\nabla_x H^{-1}(x,t)\| \cdot \|H(x,t) - H(x)\| \right\} = 0.$$

The above two relations mean the right side of (85) reduces to zero as t goes to zero. Hence there is a constant $c_3 > 0$ such that

$$x + \Delta x \in \Omega(x^*, \epsilon), \quad \text{if} \quad t \leq c_3,\ x \in \Omega(x^*, \epsilon). \qquad (86)$$

This proves the first result of the theorem.

We next consider the second statement of the theorem. Since $\nabla_x H(x^*, 0)$ is nonsingular, by result (a) of Lemma 16, for all $x \in \Omega(x^*, \epsilon)$ and $t_k \in (0, c_1)$, the Assumption (A3) is true. So the algorithm is well-defined. If $t_k \leq c_4 = \min\{c_1, c_2, c_3\}$ and $x^k \in \Omega(x^*, \epsilon)$, then it holds $x^k + \Delta x^k \in \Omega(x^*, \epsilon)$. Now from (69) and Lemma 17 we obtain

$$\begin{aligned}
\|H(x^k + \Delta x^k, t_k)\| &= \|H(x^k + \Delta x^k, t_k) - H(x^k, t^k) - \nabla_x H(x^k, t_k)\Delta x^k\| \\
&\leq \frac{1-\sigma}{C^2 \beta} \|\Delta x^k\|^2 \\
&\leq \frac{1-\sigma}{\beta} \|H(x^k, t_k)\|^2 \\
&\leq (1-\sigma) t_k \|H(x^k, t_k)\| \leq (1-\sigma)\|H(x^k, t_k)\|, \qquad (87)
\end{aligned}$$

where the first inequality follows from (84), the second inequality follows from (82), the third by $(x^k, t_k) \in \mathcal{N}(\beta, t_k)$ and the last by $t_k \leq c_4 < 1$. It follows from the line search rule in the algorithm that we can choose $h_k = 1$.

Now let us denote $\Delta t_k = t_k - t_{k+1}$ and assume that $t_{k+1} \le t_k \le c_4$. It follows

$$\begin{aligned} ||H(x^{k+1}, t_k - \Delta t_k)|| &\le \|H(x^{k+1}, t_k)|| + \|H(x^{k+1}, t_k) - H(x^{k+1}, t_k - \Delta t_k)|| \\ &= \|H(x^{k+1}, t_k) - H(x^k, t_k) - \nabla_x H(x^k, t_k)\Delta x^k|| \\ &\quad + \|H(x^{k+1}, t_k) - H(x^{k+1}, t_k - \Delta t_k)|| \\ &\le \frac{1}{2C^2\beta}||\Delta x^k||^2 + \frac{\beta t_k \Delta t_k}{2} \le \frac{1}{2\beta}||H(x^k, t_k)||^2 + \frac{\beta t_k}{2}\Delta t_k \\ &\le \frac{\beta}{2}t_k^2 + \frac{\beta t_k}{2}\Delta t_k, \end{aligned} \tag{88}$$

where the first inequality is given by the triangle inequality, the equality is derived from the choice of Δx^k, the second and third inequalities follows from results (a), (b) in Lemma 16 and Lemma 17, and the last one implied by the fact $(x^k, t_k) \in \mathcal{N}(\beta, t)$. Now let us define $\bar{c} = \min\{\alpha_3, c_4\}$. If $t_k < \bar{c}$, one has $\nu_k = 1 - \min\{\alpha_3, t_k\} = 1 - t_k$. It follows $\Delta t_k = t_k - t_k^2$. In this situation (88) reduces to

$$||H(x^{k+1}, t_k - \Delta t_k)|| \le \beta t_k^2. \tag{89}$$

Therefore, if $t_k \le \bar{c}$, we can set $\nu_k = 1 - t_k$ which means $t_{k+1} = t_k^2$. Furthermore, it still holds $||H(x^{k+1}, t_{k+1})|| \le \min\{t_{k+1}, 1\}\beta$. The proof of the theorem is completed.

Now we are ready to give our main result in this section.

Theorem 7. *Suppose $(x^*, 0)$ is an accumulation point of the sequence $\{(x^k, t_k)\}$ generated by the algorithm with $\epsilon_0 = 0$. If x^* is a strictly complementarity solution of GLCP and $\nabla H(x^*)$ is nonsingular, then the sequence $\{x^k, t^k\}$ converges local Q-quadratically to $(x^*, 0)$.*

Proof. By Theorem 6, there exists a neighborhood $\Omega(x^*, \epsilon)$ and a constant $\bar{c} \in (0, 1)$ such that, when $x^k \in \Omega(x^*, \epsilon)$ and $t_k \le \bar{c}$, the algorithm is well defined and that $t_{k+1} = t_k^2$. For simplicity, we assume that $\|\nabla_x H^{-1}(x^*)\| \le C$. From this assumption, result (c) of Lemma 16 and the fact $(x^{k+1}, t_{k+1}) \in \mathcal{N}(\beta, t)$ we obtain

$$\begin{aligned} \|x^{k+1} - x^*\| &\le C\|H(x^{k+1})\| \le C\|H(x^{k+1}, t_{k+1}) - H(x^{k+1}, 0)\| + C\|H(x^{k+1}, t_{k+1})\| \\ &\le C\frac{3\beta}{2}t_k^2, \end{aligned}$$

where the second inequality is given by triangle inequality. This relation implies

$$||(x^{k+1}, t_{k+1}) - (x^*, 0)|| \le C\frac{3\beta}{2}||(x^k, t_k) - (x^*, 0)||^2. \tag{90}$$

Since $(x^*, 0)$ is an accumulation point of the sequence $\{(x^k, t_k)\}$, there exists a point (x^k, t_k) such that $x^k \in \Omega(x^*, \epsilon)$ and $t_k \le \bar{c}$. The theorem follows immediately from (90).

□

7. Numerical results

We have tested the Algorithm 1 for several problems. The algorithm was implemented in double precision C. For all test problems, we choose $\epsilon_0 = 10^{-10}$, $\sigma = 0.005$, $\alpha_1 = 0.9, \alpha_2 = 0.85$, $\alpha_3 = 0.001$, $t_0 = 1.0$. The initial point $x^0 = a_0 e$, where $e = (1, \ldots, 1)^T \in \Re^n$ and $a_0 \in \Re$ is a constant. The set band width $\beta = ||H(x^0, t_0)||_1$. To make the algorithm robust, we terminate if $h_k \leq \alpha_1^{50}$.

Our first test problem was taken from Example 7.5 of [12], and other problems are constructed by ourselves. All the vertical block matrices of tested GLCPs are P_0-matrices. The solution points of Problems 1 to 5 are strictly complementarity. We also modify the Problems 2 a little so that the solution points of the new problems (Problems 6, 7 and 8) may be not strictly complementarity. The numerical results are listed in the table below, where a_0 is a real number associated with initial points, IT denotes totally iteration number. (x^ϵ, t_ϵ) denotes the approximate solution of GLCP satisfying the stop condition $||H(x^\epsilon)||_1 < \epsilon_0$. For convenience, we define $W_i = (W_i^1, \ldots, W_i^n)^T$, $N_i = (N_i^1, \ldots, N_i^n)^T$, $q_i = (q_i^1, \ldots, q_i^n)^T$ for the case $m_i \equiv m$.

Problem 1. N and q are given as Example 7.5 of [12].

Problem 2. $n = 6$, $m = 3$ and $W_i(x) = N_i x + q_i$, $i = 1, 2, 3$, where

$$N_1 = \begin{pmatrix} 4 & -2 & & \\ -2 & 4 & \ddots & \\ & \ddots & \ddots & -2 \\ & & -2 & 4 \end{pmatrix}, N_2 = \begin{pmatrix} 2 & -2 & & & \\ -2 & 4 & \ddots & & \\ & \ddots & \ddots & -2 & \\ & & -2 & 4 & -2 \\ & & & -2 & 2 \end{pmatrix}, N_3 = \begin{pmatrix} 1 & 1 & & \\ 1 & 2 & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 2 \end{pmatrix},$$

and $q_1 = (–2, 0, 5,–3, 6,–4)^T$, $q_2 = (1,–2, 10,–3, 6,–1)^T$, $q_3 = (0, 0, 0,–2, 0,–1)^T$.

Problem 3. $m = 3$ and $W_1^i(x) = x_i - 1$, $i = 1, \ldots, n$, $W_2^i(x) = \frac{i}{n} x_i - 1$, $i = 1, \ldots, n-1$ and $W_2^n(x) = x_n + 1$, and $W_3^1(x) = 4x_1 - 2x_2 - 1$, $W_3^i(x) = x_{i-1} + 4x_i - 2x_{i+1}$, $i = 2, \ldots, n-1$, and $W_3^n(x) = x_{n-1} + 4x_n$.

Problem 4. $m = 3$ and $W_i(x) = N_i x + q_i$, $i = 1, 2, 3$, where

$$N_1 = \begin{pmatrix} 4 & -2 & & \\ -2 & 4 & \ddots & \\ & \ddots & \ddots & -2 \\ & & -2 & 4 \end{pmatrix}, N_2 = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & \ddots & \\ \ddots & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}, N_3 = \begin{pmatrix} 1 & 1 & & & \\ 1 & 2 & & \ddots & \\ & \ddots & \ddots & 1 & \\ & & 1 & 2 & 1 \\ & & & 1 & 1 \end{pmatrix},$$

and $q_1 = q_2 = q_3 = -e \in \Re^n$.

Problem 5. $m = 6$ and $W_1^i(x) = x_i - 1$, $i = 1, \ldots, n$, $W_2^i(x) = \frac{i}{n} x_i - 1$, $i = 1, \ldots, n-1$ and $W_2^n(x) = x_n + 1$, and $W_3^1(x) = 4x_1 - 2x_2 - 1$, $W_3^i(x) = x_{i-1} + 4x_i -$

$2x_{i+1},\ i = 2, \ldots, n-1$, and $W_3^n(x) = x_{n-1} + 4x_n$, and $W_i(x) = N_i x + q_i,\ i = 4, 5, 6$, where

$$N_4 = \begin{pmatrix} 4 & -2 & & \\ -2 & 4 & \ddots & \\ & \ddots & \ddots & -2 \\ & & -2 & 4 \end{pmatrix},\ N_5 = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix},\ N_6 = \begin{pmatrix} 1 & 1 & & & \\ 1 & 2 & \ddots & & \\ & \ddots & \ddots & 1 & \\ & & 1 & 2 & 1 \\ & & & 1 & 1 \end{pmatrix},$$

and $q_4 = q_5 = q_6 = -e \in \Re^n$.

Problem 6. $n = 6,\ m = 3,\ W_i(x),\ i = 1, 2, 3$ are the same as that defined in Problem 3. Denote $\overline{W_i}(x)$ be the functions of the new problem. We choose $\overline{W_1}(x) = W_1(x) - W_1(x^*)$ where $x^* = (1, 1, 0, 1, 0, 1)^T \in \Re^6$ is the solution point of Problem 2, and that $\overline{W_2}(x) = W_2(x), \overline{W_3}(x) = W_3(x).$;

Problem 7. $n = 6,\ m = 3,\ W_i(x),\ i = 1, 2, 3$ are the same as that defined in Problem 2. We choose $\overline{W_1}(x) = W_1(x)$, $\overline{W_2}(x) = W_2(x) - W_2(x^*)$ and $\overline{W_3}(x) = W_3(x)$.

Problem 8. $n = 6, m = 3,\ W_i(x),\ i = 1, 2, 3$ are the same as that defined in Problem 2. We choose $\overline{W_1}(x) = W_1(x)$, $\overline{W_2}(x) = W_2(x)$ and $\overline{W_3}(x) = W_3(x) - W_3(x^*)$.

Numerical Results								
Example	DIM	a_0	β	$\|\|H(x^0)\|\|_1$	IT	t_ϵ	$\|\|H(x^\epsilon)\|\|_1$	$\|\|H(x^\epsilon, t_\epsilon)\|\|_1$
1	2	1.0	2.2696	1.0	3	7.2e-5	8.7e-12	1.2e-7
		10.0	9.1381	10.0	3	4.2e-5	6.2e-15	1.4e-10
		-10.0	23.1753	22.0	4	3.4e-8	0.0	0.0
2	6	1.0	10.7386	9.0	3	1.8e-4	2.8e-11	1.5e-7
		10.0	18.1043	18.0	4	4.4e-5	8.8e-16	1.9e-11
		-10.0	215.1278	215.0	4	2.2e-8	8.8e-16	3.9e-8
3	50	5.0	76.2907	86.5	4	2.0e-5	1.0e-12	5.4e-8
	100	5.0	151.0446	171.5	4	2.0e-5	2.6e-12	1.3e-7
	200	5.0	300.5496	341.5	5	4.2e-6	6.4e-15	1.5e-9
	100	-5.0	1605.0162	1605.0	4	1.1e-5	4.0e-14	3.4e-9
	200	-5.0	3205.0233	3205.0	4	1.1e-5	4.3e-13	3.6e-8
4	50	5.0	88.6844	56.0	4	2.3e-5	6.6e-13	2.8e-8
	100	5.0	173.4037	106.0	4	2.8e-5	1.1e-12	4.1e-8
	200	5.0	342.8423	206.0	5	5.0e-6	4.3e-11	8.5e-6
	100	-5.0	2081.3955	208.0	6	5.1e-6	1.5e-11	2.9e-6
	200	-5.0	4181.3955	418.0	7	1.2e-6	2.3e-11	1.9e-5
5	50	5.0	89.3458	52.9	4	3.3e-5	3.4e-12	1.0e-7
	100	5.0	178.6970	102.95	5	1.4e-5	1.3e-11	9.3e-7
	200	5.0	357.3514	202.97	6	2.2e-6	3.4e-11	1.5e-5
	100	-5.0	2096.7621	2095.0	5	1.9e-5	9.4e-11	4.7e-6
	200	-5.0	4197.4337	4195.0	6	2.9e-6	3.7e-12	1.2e-6
6	6	1.0	11.8312	10.0	22	2.7e-4	9.7e-7*	3.2e-3
		10.0	16.8723	16.0	24	1.8e-4	6.1e-7*	3.1e-3
		-10.0	215.1278	215.0	11	9.4e-6	1.9e-8*	2.0e-3
7	6	1.0	11.9236	10.0	4	1.0e-5	1.7e-15	1.6e-10
		10.0	15.9961	16.0	3	7.7e-8	0.0e0	0.0e0
		-10.0	215.1278	215.0	6	5.1e-7	1.3e-15	2.6e-9
8	6	1.0	10.6568	9.0	12	4.4e-4	2.2e-6*	4.7e-3
		10.0	18.1043	18.0	13	3.0e-4	1.7e-6*	5.6e-3
		-10.0	223.6935	223.0	6	1.9e-5	8.2e-8*	4.2e-3

The above table indicates that for Problems 1-5, our method finds the strictly complementarity solution points of the problems after few iterations. It is easy to see that all the points $x = (1, 1, 0, 1, 0, 1)^T + t(1, 1, 1, 1, 1, 1)^T$, $t \geq 0$ are solution points of Problem 7 and the solution points are strictly complementarity if $t > 0$. Our algorithm detects the solution point $x^* = (1, 1, 0, 1, 0, 1)^T$ once, and finds many others solution point from different initial points.

However, for Problems 6 and 8, the algorithm not only takes more iterations but also terminates at a point which does not satisfy the stop criterion of the method. The possible reason of the relatively poor behavior of the algorithm for Problems 6 and 8 is that the solution points are not strictly complementarity.

8. Conclusion remarks

The generalized complementarity problems are transformed into an equivalent system of nonsmooth equations. It is shown that, the norm of the nonsmooth equations act as a local error bound for GLCP. If some additional conditions are satisfied, then the norm of the equations also provides a global error bound for GLCP.

The relations between Chen-Mangasarian's neural network smooth function and the function $g(x, t)$ is also explored. It is shown that, for some special choice, these two different class of smoothing functions act precisely in the same way. By using the smoothing function, we approximate the GLCP via a system of parameterized smooth equations. The distance from the smoothing path of the equations to the solution set of the GLCP is considered. Our results show that, when the parameter tends to zero, we can approximate the solution set of the GLCP to any desired accuracy. A non-interior continuation method is proposed to follow the smoothing path of the smooth equations. Under certain assumptions, the method is globally convergent and locally quadratically convergent. Preliminary numerical results show the method is promising.

To the authors' knowledge, this work is the first one which cast GLCP as a system of smooth equations in $\Re^n$. There are several ways to extend our results. The first one is to study the behavior of $H(x)$ for generalized complementarity problems(GCP). Under what conditions, $\|H(x)\|$ provides a local or global error bound for GCP? Some results about this problem have been reported in the first author's recent work [34]. However, this is still an interesting topic deserve further study. The second is to design new and efficient method for solving GCP via the approach proposed in this paper. We note that, in a recent work [35], Qi and Liao used the same approach proposed in our paper to reformulate GCP as a system of parameterized equations. Different from our method, they try to solve the equations $H(x, t) = 0$ and $t = 0$. A Newton-type method was also presented to solve their system.

We observe that if the solution points are not strictly complementarity, then our algorithm may behave poorly. How to design efficient algorithms without the strictly complementarity conditions is also an interesting topics.

Acknowledgements. The authors would like to thank the anonymous referees for their useful suggestions, particularly for one referee's remark about one mistake in the proof of Theorem 6 in a early version of this paper. They are also grateful to Dr. D. Sun for drawing their attention to exponential penalty function.

References

1. Allgower, E.L., Georg, K. (1990): Numerical continuation Methods: An introduction. Springer, Berlin
2. Bertsekas, D.P. (1976): A new algorithm for solution of nonlinear resistive networks involving diodes. IEEE Trans. Circuit Theory, Vol. CAS23, 599–608
3. Bertsekas, D.P. (1976): Minimax Methods Based on Approximation. Proceedings of the 1976 Johns Hopkins Conference on Information Sciences and Systems, Baltimore
4. Bertsekas, D.P. (1977): Approximation procedures based on the method of multipliers. J. Optim. Theory Appl. **23**, 487–510
5. Burke, J.V., Xu, S. (1998): The global linear convergence of a noninterior path-following algorithm for linear complementarity problem. Math. Oper. Res. **23**, 719–734
6. Chen, B., Harker, P.T. (1997): Smooth approximations to nonlinear complementarity problems. SIAM J. Optim. **7**, 403–420
7. Chen, C., Mangasarian, O.L. (1995): Smoothing methods for convex inequalities and linear complementarity problems. Math. Program. **71**, 51–69
8. Chen, C., Mangasarian, O.L. (1996): A class of smoothing functions for nonlinear and mixed complementarity problems. Comput. Optim. Appl. **5**, 97–138
9. Chen, B., Xiu, N. (1999): A global linear quadratic noninterior continuation method for nonlinear complementarity problems based on Chen-Mangasarian smoothing functions. SIAM J. Optim. **9**, 624–645
10. Chen, B., Chen, X. (1999): A global and local superlinear continuation-smoothing method for $P_0 + R_0$ and monotone NCP. SIAM J. Optim. **9**, 605–623
11. Cottle, R.W., Dantzig, G.B. (1970): A generalization of the linear complementarity problem. J. Comb. Theory **8**, 79–90
12. Ebiefung, A. (1995): Nonlinear mappings associated with the generalized linear complementarity problem. Math. Program. **69**, 255–268
13. Ebiefung, A.A., Kostreva, M.M. (1992): Global solvability of generalized linear complementarity problems and a related class of polynomial complementarity problems. In: Floudas, C., Pardalos, P., eds., Recent Advances in Global Optimization, pp. 102–124. Princeton University Press, Princeton, NJ
14. Ebiefung, A.A., Kostreva, M.M. (1993): The generalized Leontief input-output model and its application to the choice of new technology. Ann. Oper. Res. **44**, 161–172
15. Ferris, M.C., Pang, J.S. (1997): Engineering and Economic Applications of Complementarity Problems. SIAM Rev. **39**, 669–713
16. Fischer, A. (1992): A Special Newton-type optimization method. Optimization **24**, 269–284
17. Fischer, A. (1995): A Newton-type method for positive-semidefinite linear complementarity problems. J. Optim. Theory Appl. **86**, 585–608
18. Freund, R.M., Mizuno, S. (1996): Interior point methods: current status and future directions. OPTIMA Newsletter **51**, 1–9
19. Fukushima, M. (1996): Merit functions for variational inequality and complementarity problems. In: Di Pillo, P., Giannessi, F., eds., Nonlinear Optimization and Applications, pp. 155–170. Plenum Press, New York
20. Geiger, C., Kanzow, C. (1996): On the Resolution of Monotone complementarity problems. Comput. Optim. Appl. **5**, 155–173
21. Gowda, M.S., Sznajder, R. (1994): The generalized order linear complementarity problem. SIAM J. Matrix Anal. Appl. **159**, 779–795
22. Gowda, M.S., Sznajder, R. (1996): A generalization of the Nash equilibrium theorem on bimatrix games. Int. J. Game Theory **25**, 1–12
23. Harker, P.T., Pang, J.S. (1990): Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications. Math. Program. **48**, 339–357
24. Hotta, K., Yoshise, A. (1996): Global convergence of a class of non-interior point algorithms using Chrn-harker-Kanzow function for nonlinear complementarity problems. Technical Report 708, Institute of Policy and Planning Sciences, University of Tsukuba, Tsukuba, Japan
25. Kanzow, C. (1996): Some noninterior continuation methods for linear complementarity problems. SIAM J. Matrix Anal. Appl. **17**, 851–868
26. Kanzow, C., Jiang, H. (1998): A continuation method for the solution of monotone variational inequality problems. Math. Program. **81**, 103–125
27. Kojima, M., Megiddo, N., Mizuno, S. (1993): A general framework of continuation methods for complementarity problems. Math. Oper. Res. **18**, 945–963
28. Kojima, M., Megiddo, N., Noma, T. (1991): Homotopy continuation methods for nonlinear complementarity problems. Math. Oper. Res. **16**, 754–774

29. Kort, B.W., Bertsekas, D.P. (1972): A new penalty function algorithm for constrained minimization. Proceedings of the 1972 IEEE Conference on Decision and Control, New Orleans, Louisiana
30. Li, X.S. (1991): An aggregate function method for nonlinear programming. Sci. China Ser. A **34**, 1467–1473
31. Mangasarian, O.L. (1979): Generalized linear complementarity problems as linear programming. Oper. Res. Verfahren **31**, 393–402
32. Mohan, S.R., Neogy, S.K., Sridhar, R. (1996): The generalized linear complementarity problem revisited. Math. Program. **74**, 197–218
33. Pang, J.S. (1995): Complementarity problems. In: Horst, R., Pardalos, P., eds., Handbook of Global Optimization, pp. 271-338. Kluwer Academic Publishers, Boston
34. Peng, J.M. (1998): A smoothing function and its applications. In: Fukushima, M., Qi, L., eds., Reformulation-Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, pp. 293–321. Kluwer Academic Publisher
35. Qi, H.D., Liao, L.Z. (1997): A smoothing Newton method for extended vertical linear complementarity problems. Research Report, Institute of Computational Mathematics and Scientific Computing, Chinese Academy of Sciences, P.O. Box 2719, Beijing, China
36. Qi, L., Sun, D. (1997): Globally linearly, and locally superlinearly convergent versions of the Hotta-Yoshise non-interior point algorithm for nonlinear complementarity problems. Technical Report, School of Mathematics, The University of New South Wales, Sydney, Australia
37. Robinson, S.M. (1981): Some continuity properties of polyhedral multifunctions. Math. Program. Study **14**, 206–214
38. Sun, M. (1987): Singular control problems in bounded intervals. Stochastics **21**, 303–344
39. Sun, M. (1989): Monotonicity of Mangasarian's iterative algorithm for the generalized linear complementarity. J. Math. Anal. Appl. **144**, 474–485
40. Sznajder, R., Gowda, M.S. (1995): Generalizations of P_0- and P-properties; extended vertical and horizonta. Linear Algebra Appl. **223/224**, 695–715
41. Tang, H., Zhang L. (1994): A maximum entropy method for convex programming. Chin. Sci. Bull. **39**, 682–684
42. Vandenberghe, L., DeMoor, B.L., Vanderwalle, J. (1989): The generalized linear complementarity applied to the complete analysis of resistive piecewise-linear circuits. IEEE Trans. Circuits Syst. **11**, 1382–1391
43. Wright, S., Ralph, D. (1996): A superlinear infeasible-interior-point algorithm for monotone complementarity problems. Math. Oper. Res. **21**, 815–834
44. Xu, S. (1996): The global linear convergence of an infeasible non-interior path-following algorithm for complementarity problems with uniform P-functions. Preprint, Department of Mathematics, University of Washington, Seattle, WA 98195