A New Path-Following Algorithm for Nonlinear P_* Complementarity Problems*

Y. B. Zhao
Institute of Applied Mathematics, AMSS, Chinese Academy of Sciences, Beijing 100080, China (Email: ybzhao@amss.ac.cn).

D. Li
Department of Systems Engineering and Engineering Management, Chinese University of Hong Kong, Shatin, NT, Hong Kong (Email: dli@se.cuhk.edu.hk).

Abstract. Based on the recent theoretical results of Zhao and Li [*Math. Oper. Res.,* 26 (2001), pp. 119-146], we present in this paper a new path-following method for nonlinear P_* complementarity problems. Different from most existing interior-point algorithms that are based on the central path, this algorithm tracks the "regularized central path" which exists for any continuous P_* problem. It turns out that the algorithm is globally convergent for any P_* problem provided that its solution set is nonempty. By different choices of the parameters in the algorithm, the iterative sequence can approach to different types of points of the solution set. Moreover, local superlinear convergence of this algorithm can also be achieved under certain conditions.

Key words. Nonlinear complementarity problems, path-following algorithms, regularized central path, Tikhonov regularization, P_*-mappings.

*The research of the first author was supported by The National Natural Science Foundation of China under Grant No. 10201032 and Grant No. 70221001. The research of the second author was supported by Grant CUHK4214/01E, Research Grants Council, Hong Kong.

1. Introduction

We consider the nonlinear complementarity problem (NCP):

$$x \geq 0,\ f(x) \geq 0,\ x^T f(x) = 0,$$

where f is a mapping from R^n into itself. This problem is said to be a P_* NCP if f is a P_*-mapping. We recall that f is said to be a P_*-mapping if there exists a constant $\kappa \geq 0$ such that

$$(1+\kappa)\sum_{i\in I_+}(x_i - y_i)(f_i(x) - f_i(y)) + \sum_{i\in I_-}(x_i - y_i)(f_i(x) - f_i(y)) \geq 0$$

for any distinct vectors x, y in R^n, where

$$I_+ = \{i : (x_i - y_i)(f_i(x) - f_i(y)) > 0\},\ I_- = \{i : (x_i - y_i)(f_i(x) - f_i(y)) < 0\}.$$

It is evident that every monotone mapping is a P_*-mapping with $\kappa = 0$. We also recall that f is said to be a P_0-mapping (P-mapping) if $\max_{x_i \neq y_i}(x_i - y_i)(f_i(x) - f_i(y)) \geq\ (>)\ 0$ for any distinct vectors x and y in R^n (see, for example, [9, 26]). It is easy to see that every P_*-mapping is a P_0-mapping. The concept of the P_*-mapping is a straightforward extension of the P_*-matrix (sufficient matrix) introduced by Cottle et al. [8] and Kojima et al. [19]. In fact, when $f = Mx + q$ where M is an $n \times n$ matrix and $q \in R^n$, it is evident that f is a P_*-mapping if and only if M is a P_*-matrix. It is worth mentioning that the P_*-mapping can be equivalently defined as follows: There exists a nonnegative constant κ' such that

$$(1+\kappa')\max_{1\leq i\leq n}(x_i - y_i)(f_i(x) - f_i(y)) + \min_{1\leq i\leq n}(x_i - y_i)(f_i(x) - f_i(y)) \geq 0,$$

for any distinct vectors x, y in R^n (see Zhao and Isac [37]). P_* complementarity problems have been extensively studied in the area of interior-point methods (see, for instance, [3, 16, 19, 23, 24, 27, 28].

Many path-following methods for complementarity problems, in particular the interior-point methods (see, e.g. [19, 36]) and non-interior-point methods (see, e.g. [4, 5, 11, 14, 34]), are designed to follow the central path, i.e., $\{(x(\mu), v(\mu)) :\ \mu \in (0, \infty)\}$ where $(x(\mu), v(\mu))$ is the unique solution to the system:

$$x > 0,\ v = f(x) > 0,\ Xv = \mu e. \tag{1}$$

Interior-point algorithms usually iterate in the positive orthant, while non-interior-point algorithms allow negative iterates. It is known that for P_* complementarity problems the central path exists if and only if the problem has a strictly feasible point (see, for instance, Kojima et al. [19]). It is shown in [38] that a P_* problem has a strictly feasible point if and only if its solution set is nonempty and bounded. Therefore, we conclude that when the solution set is unbounded the P_* problem has no strictly feasible point, and hence the central path does not exist. This is why most of these central path based methods usually need the requirement of the existence of an interior point, or a nonempty and bounded solution set. For P_0 NCPs, the later requirement implies the existence of an interior point.

On the other hand, Tikhonov regularization methods (see, e.g. [9, 10, 11, 15, 26, 29, 30, 39, 40]) follow the Tikhonov regularization path instead of the central path. Tikhonov regularization path, denoted by $\{(x(\mu), z(\mu)) : \mu \in (0, \infty)\}$, is a continuous trajectory on which each point $(x(\mu), z(\mu))$ is a unique solution to the system:

$$x > 0,\ z = f(x) + \mu x > 0,\ Xz = 0. \tag{2}$$

It is shown in [39] that the entire Tikhonov regularization path $\{(x(\mu), z(\mu)) : \mu \in (0, \infty)\}$ exists for any P_* NCP, and it is bounded as long as the P_* NCP has a solution. Therefore, the existence of a strictly feasible point is not necessary for the existence and boundedness of the Tikhonov regularization path. As a result, the Tikhonov regularization path based algorithms may not need the existence of an interior point. To our best knowledge, there is no path-following algorithm in the literature that employs the framework of interior-point methods to track the Tikhonov regularization path. It is worth mentioning that some existing non-Tikhonov regularization type algorithms can solve NCPs without requiring the existence of the interior point. For example, for P_* LCPs, the infeasible interior-point methods (see, for example, [3, 23, 24]) do not require the strict feasibility condition for the underlying problem. For monotone complementarity problems, this condition is also not needed for the interior-point algorithms using self-dual embedding models. Such a method was first proposed by Ye [35] for monotone LCPs, and was later generalized to monotone NCPs by Andersen and Ye [2]. Unfortunately, Ye's model cannot be applied to nonlinear P_* problems since there is no guarantee for the embedded problem being again a P_* problem. We also mention that Kojima et al. [19] proposed a big-$\mathcal{M}$ method to obtain a strictly feasible point for an artificial LCP. However, it remains to be seen whether their model can be applied to P_* or more general NCPs.

Motivated by the above observation, we define a new continuous path that uses both ideas of interior-point and regularization methods. Let $a \in R^n_{++}$ and $b \in R^n$ be given. We first define the following general system:

$$x > 0, \ \ s = f(x) + \mu' x + \mu b > 0, \ \ Xs = \mu a, \tag{3}$$

where μ' and μ are two positive parameters. In particular, if $b = 0$ and $\mu' = \mu$, then (3) reduces to $x > 0, \ s = f(x) + \mu x > 0, \ Xs = \mu a$ which is indeed a combination of (1) and (2).

Clearly, system (3) is quite different from (1) and (2). When f is a P_*-mapping, for each given $\mu' > 0$, it is easy to see that $g(x) = f(x) + \mu' x$ is a P-mapping in x. If b is restricted to be in R^n_{++}, it follows from the results in [18] that for every $(\mu', \mu) > 0$ system (3) has a unique solution. If b is chosen arbitrarily in R^n, it is shown in [39] that system (3) remains well-defined for some cases.

Notice that system (3) has two parameters μ' and μ. From the viewpoint of numerical implementation, however, it is convenient to consider the case with only one parameter. Thus, in this paper, we consider the case of $\mu = \frac{\theta}{1-\theta}$ and $\mu' = \theta^p$ in (3), where θ is a parameter in (0,1) and p is a constant in $(0, \infty)$. Then, by setting $y = (1-\theta)s$, system (3) reduces to the following one-parameter system:

$$x > 0, \ \ y = (1-\theta)(f(x) + \theta^p x) + \theta b > 0, \ \ Xy = \theta a,$$

which can be written as

$$\mathcal{H}(x, y, \theta) := \begin{pmatrix} Xy - \theta a \\ y - (1-\theta)(f(x) + \theta^p x) - \theta b \end{pmatrix} = 0, \ \ (x, y) > 0. \tag{4}$$

This system can be viewed as either a modified central path system by adding the Tikhonov regularization term "$\theta^p x$" to the function f or a modified Tikhonov regularization path system by introducing the centering term "θa". For each $\theta \in (0,1)$, we denote by $(x(\theta), y(\theta))$ the solution of system (4). The term *regularized central path* is used to refer the path $\{(x(\theta), y(\theta)) : \theta \in (0,1)\}$ throughout the paper.

System (4) makes it possible to design a new regularization algorithm for NCPs by employing the idea of interior-point algorithms. The purpose of this paper is to construct such a numerical method and to provide its global and local convergence analysis.

The organization of the paper is as follows. In Section 2, we describe the algorithm. In Section 3, we prove the global convergence of the algorithm, and characterize the accumulation points of the iterative sequence generated by our algorithm. In Section 4, we prove the local superlinear convergence of the algorithm. Some numerical results are reported in Section 5, and conclusions are given in the last section.

Notation: R^n denotes the n-dimensional Euclidean space, R^n_+ the nonnegative orthant, and R^n_{++} the positive orthant. We denote by $x \geq 0$ ($x > 0$) a vector $x \in R^n_+$ ($x \in R^n_{++}$). For any vector x, the capital X denotes the diagonal matrix whose (i,i)th entry is given by the ith component of x, i.e., $X = \mathrm{diag}(x)$, and x_I, where $I \subseteq \{1, ..., n\}$, denotes the vector with components x_i for $i \in I$ being arranged in the same order as in x. If $M = (m_{ij})_{n \times n}$ is an $n \times n$ matrix, then M_{II} denotes a submatrix of M, with entries m_{ij}, where $i \in I$ and $j \in I$, being arranged in the same order as in M. For any index set, $|I|$ denotes the cardinality of I. The symbols $\|\cdot\|$ and $\|\cdot\|_\infty$ denote, respectively, the 2-norm and the infinity norm of a vector or a matrix. Throughout the paper, e denotes the vector with all components equal to 1, and its dimension depends on the context; E denotes the identity matrix, i.e., $E = \mathrm{diag}(e)$.

2. A path-following algorithm

We first list here two known results about system (4) established in [39].

Theorem 2.1. [39] *Let f be a continuous P_*-mapping from R^n into itself. Let $(a,b) \in R^n_{++} \times R^n$ and $p \in (0,\infty)$ be given.*

(i) System (4) has a unique solution $(x(\theta), y(\theta))$ for each $\theta \in (0,1)$, and $(x(\theta), y(\theta))$ is continuous on (0,1). If f is continuously differentiable, $(x(\theta), y(\theta))$ is also continuously differentiable on (0,1).

(ii) Moreover, if $p \le 1$ and the solution set of NCP is nonempty, then for any $\delta \in (0,1)$ the section of the path $\{(x(\theta), y(\theta)) : \theta \in (0,\delta]\}$ is bounded. Thus, any accumulation point of the path, as $\theta \to 0$, is a solution to the NCP.

In particular, for monotone NCPs, we have the following result.

Theorem 2.2. [39] *Let $(a,b) \in R^n_{++} \times R^n$ and $p \in (0,1)$. Let f be a continuous monotone mapping from R^n into itself. Suppose that the solution set of NCP is nonempty. Then the entire trajectory $\{(x(\theta), y(\theta)) : \theta \in (0,1)\}$ generated by system (4) converges, as $\theta \to 0$, to $(\hat{x}, \hat{y})$ where $\hat{x}$ is the least 2-norm solution, i.e., $\|\hat{x}\| \le \|x^*\|$ where x^* is an arbitrary solution of the NCP.*

It is worth mentioning that the part (ii) in Theorem 2.1 does not cover the case of $p > 1$. In fact, for $p > 1$ and $b \le 0$, it remains elusive whether this result holds or not. However, if the vector b is restricted to be in R^n_{++}, then by using the proof idea in [20], we can show the following result: *Let f be given as in Theorem 2.1. For any given $(a,b) > 0$ and $p > 1$, if the NCP has a nonempty solution set, then the result of part (ii) in Theorem 2.1 remains valid.*

Throughout the paper, p is a fixed scalar in $(0,\infty)$, and f is assumed to be continuously differentiable. ∇f denotes the Jacobian matrix of f. Thus, for a given vector $(a,b) \in R^n_{++} \times R^n$, the mapping $\mathcal{H}$ defined by (4) is continuously differentiable. Denote by $\nabla_{(x,y)}\mathcal{H}(x,y,\theta)$ the Jacobian matrix of $\mathcal{H}$ with respect to (x,y), i.e.,

$$\nabla_{(x,y)}\mathcal{H}(x,y,\theta) = \begin{pmatrix} Y & X \\ -(1-\theta)(\nabla f(x) + \theta^p E) & E \end{pmatrix}.$$

The following fact is very useful.

Lemma 2.1. *Let f be a continuously differentiable P_0-mapping. Let $(\bar{x}, \bar{y}) \in R^{2n}_{+}$ and $\bar{\theta} \in (0,1)$. Then the matrix $\nabla_{(x,y)}\mathcal{H}(\bar{x}, \bar{y}, \bar{\theta})$ is nonsingular if and only if $\bar{x} + \bar{y} > 0$.*

Proof. Notice that $(\bar{x}, \bar{y}) \in R^n_+$. If for some i, $\bar{x}_i + \bar{y}_i = 0$, then all elements in the ith row of the matrix $\nabla_{(x,y)}\mathcal{H}(\bar{x}, \bar{y}, \bar{\theta})$ are zeroes. Hence, if $\nabla_{(x,y)}\mathcal{H}(\bar{x}, \bar{y}, \bar{\theta})$ is nonsingular, we must have that $\bar{x}_i + \bar{y}_i > 0$ for $i = 1, ..., n$, i.e., $\bar{x} + \bar{y} > 0$. Conversely, if $\bar{x} + \bar{y} > 0$, we prove that $\nabla_{(x,y)}\mathcal{H}(\bar{x}, \bar{y}, \bar{\theta})$ is nonsingular. By contrary, we assume that there exists an vector $(u,v) \neq 0$ such that $\nabla_{(x,y)}\mathcal{H}(\bar{x}, \bar{y}, \bar{\theta})(u,v) = 0$, i.e.,

$$\bar{Y}u + \bar{X}v = 0, \tag{5}$$

$$v = (1-\bar{\theta})(\nabla f(\bar{x}) + \bar{\theta}^p E)u. \tag{6}$$

From the above, we see that $u \neq 0$ since otherwise v also equals to zero. Since $\nabla f(\bar{x}) + \bar{\theta}^p E$ is a P-matrix, there exists an index i such that $u_i[(\nabla f(\bar{x}) + \bar{\theta}^p E)u]_i > 0$. Thus, from (6) we have $v_i u_i = (1-\bar{\theta})u_i[(\nabla f(\bar{x}) + \bar{\theta}^p E)u]_i > 0$. Since $\bar{y}_i \ge 0$, $\bar{x}_i \ge 0$, $\bar{y}_i + \bar{x}_i > 0$ and $v_i u_i > 0$, it follows that $\bar{y}_i u_i + \bar{x}_i v_i \neq 0$, which contradicts (5). $\square$

Central to many aspects of path-following algorithms is the concept of the neighborhood associated with a continuous or smooth path to be considered. Notice that the *regularized central path* is given by

$\{(x,y) > 0 :\ \mathcal{H}(x,y,\theta) = 0,\ \theta \in (0,1)\}$. In this paper, we choose the neighborhood around the regularized central path as follows:

$$\mathcal{N}(\beta) := \{(x,y) \geq 0 :\ \|\mathcal{H}(x,y,\theta)\|_\infty \leq \beta\theta,\ \theta \in (0,1)\}.$$

For each given $\theta \in (0,1)$, we denote

$$\mathcal{N}_\beta(\theta) = \{(x,y) \geq 0 : \|\mathcal{H}(x,y,\theta)\|_\infty \leq \beta\theta\}.$$

Clearly, $\mathcal{N}(\beta) = \cup_{\theta\in(0,1)}\mathcal{N}_\beta(\theta)$.

As we will see below, at most two linear systems with the same matrix $\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k)$ are needed to be solved at each iteration of the algorithm proposed in this paper. In order to guarantee the nonsingularity of the matrix $\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k)$, by Lemma 2.1, it is sufficient to maintain the positivity of x^k and y^k when $\theta_k \in (0,1)$. Thus, all iterates generated by our algorithm are confined to be in the positive orthant. Before stating the algorithm, we first give some strategies about the choice of initial points and parameters required by the algorithm. Such initial points and parameters can be easily constructed.

Strategy 2.1. Let $a \in R^n_{++}, b \in R^n$, $\theta_0 \in (0,1)$, and $(x^0, y^0) > 0$ be given. Choose $\beta > 0$ such that

$$\beta \geq \|\mathcal{H}(x^0, y^0, \theta_0)\|_\infty/\theta_0.$$

Strategy 2.2. Let $(x^0, y^0) > 0$ and $\theta_0 \in (0,1)$ be given. Set $a = X^0y^0/\theta_0$ and

$$b = \left[y^0 - (1-\theta_0)(f(x^0) + (\theta_0)^p x^0)\right]/\theta_0.$$

Let β be an arbitrary scalar in $(0,\infty)$.

Strategy 2.3. Let $\theta_0 \in (0,1)$ and $x^0 > 0$ be given, and let $y^0 > 0$ be chosen such that $y^0 > (1-\theta_0)(f(x^0) + (\theta_0)^p x^0)$. Set $a = X^0y^0/\theta_0$, and

$$b = \left[y^0 - (1-\theta_0)(f(x^0) + (\theta_0)^p x^0)\right]/\theta_0.$$

Let β be an arbitrary scalar in $(0,\infty)$.

Remark 2.1. All the above strategies guarantee that $(x^0, y^0) > 0$ and $(x^0, y^0) \in \mathcal{N}_\beta(\theta_0)$ which are the requirement in the initial step of our algorithm. We note that both Strategies 2.2 and 2.3 imply that $\mathcal{H}(x^0, y^0, \theta_0) = 0$. Thus, the initial condition $(x^0, y^0) \in \mathcal{N}_\beta(\theta_0)$ is satisfied for any given $\beta \in (0,\infty)$. In particular, β can be taken such that $0 < \beta < \min_{1\leq i\leq n} a_i$. Also, since $b > 0$ in Strategy 2.3, β can be taken such that $0 < \beta < \min_{1\leq i\leq n} \min\{a_i, b_i\}$. Such choices will be used in the analysis for the global convergence of the algorithm (see Section 3 for details).

We now state the algorithm as follows.

Algorithm 2.1. Let $p \in (0,\infty)$ be a fixed positive scalar. Let $\varepsilon > 0$ be the termination tolerance. Assign scalars $\alpha_1, \alpha_2, \sigma$ and η in (0,1).

Step 1 (Initial Step). Let $(x^0, y^0) \in R^{2n}_{++}, a \in R^n_{++}, b \in R^n$, $\theta_0 \in (0,1)$, and $\beta > 0$ such that $(x^0, y^0) \in \mathcal{N}_\beta(\theta_0)$.

Step 2. If $\|\mathcal{H}(x^k, y^k, 0)\| \leq \varepsilon$, stop; otherwise, let $(\Delta\hat{x}^k, \Delta\hat{y}^k)$ solve the equation

$$\mathcal{H}(x^k, y^k, 0) + \nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k)(\Delta x, \Delta y) = 0. \tag{7}$$

Let

$$t_k = \sup\{t > 0 : x^k + \lambda\Delta\hat{x}^k \geq 0,\ y^k + \lambda\Delta\hat{y}^k \geq 0 \text{ for all } \lambda \in (0,t]\}.$$

Step 3. Set

$$\hat{\lambda}_k = \min\{1, (1-\theta_k)t_k\},\ (\hat{x}^{k+1}, \hat{y}^{k+1}) = (x^k, y^k) + \hat{\lambda}_k(\Delta\hat{x}^k, \Delta\hat{y}^k),$$

and $\hat{\theta}_{k+1} = \|\mathcal{H}(\hat{x}^{k+1}, \hat{y}^{k+1}, 0)\|_\infty$. If $(\hat{x}^{k+1}, \hat{y}^{k+1}) \in \mathcal{N}_\beta(\hat{\theta}_{k+1})$ and $\hat{\theta}_{k+1} \le \eta\theta_k$, then set

$$(x^{k+1}, y^{k+1}) = (\hat{x}^{k+1}, \hat{y}^{k+1}), \ \theta_{k+1} = \hat{\theta}_{k+1}, \ k := k+1$$

and go to Step 2; otherwise, go to Step 4.

Step 4. If $\mathcal{H}(x^k, y^k, \theta_k) = 0$, set $(x^{k+1}, y^{k+1}) = (x^k, y^k)$ and go to Step 6; otherwise, let $(\Delta x^k, \Delta y^k)$ be the solution to equation

$$\mathcal{H}(x^k, y^k, \theta_k) + \nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k)(\Delta x, \Delta y) = 0. \tag{8}$$

Let

$$t'_k = \sup\{t > 0 : x^k + \lambda\Delta x^k \ge 0, \ y^k + \lambda \Delta y^k \ge 0 \text{ for all } \lambda \in (0, t]\}.$$

Step 5. Set $t''_k = \{1, (1-\theta_k)t'_k\}$. Let λ_k be the maximum among values of t''_k, $\alpha_1 t''_k$, $\alpha_1^2 t''_k$, ... such that

$$\|\mathcal{H}(x^k + \lambda_k\Delta x^k, y^k + \lambda_k \Delta y^k, \theta_k)\|_\infty \le (1 - \sigma\lambda_k)\|\mathcal{H}(x^k, y^k, \theta_k)\|_\infty. \tag{9}$$

Set $(x^{k+1}, y^{k+1}) = (x^k, y^k) + \lambda_k(\Delta x^k, \Delta y^k)$.

Step 6. Let γ^k be the maximum among the values of $\alpha_2, \alpha_2^2, \alpha_2^3, ...$ such that $(x^{k+1}, y^{k+1}) \in \mathcal{N}_\beta((1-\gamma^k)\theta_k)$. Set $\theta_{k+1} = (1-\gamma^k)\theta_k$ and $k := k+1$. Go to Step 2.

A feature of Algorithm 2.1 is that it uses a modified predictor-corrector strategy. A combination of Step 2 and Step 3 in this algorithm can be viewed as the 'predictor step', actually an approximate Newton step. The two steps are used to accelerate the iteration when iterates are close to the solution set such that a local rapid convergence can be achieved. At each iteration, the criterion '$(\hat{x}^{k+1}, \hat{y}^{k+1}) \in \mathcal{N}_\beta(\hat{\theta}_{k+1})$ and $\hat{\theta}_{k+1} \le \eta\theta_k$' is checked to decide whether the iterate $(\hat{x}^{k+1}, \hat{y}^{k+1})$ generated by the 'predictor step' can be accepted or not. If it is accepted, the 'predictor step' is repeated; otherwise, the algorithm proceeds to the 'corrector step' that consists of Step 4 through Step 6, which can guarantee a desirable global convergence of iterates. Modified predictor-corrector strategies are also used by several authors such as Wright and Ralph [33], and Chen and Chen [6]. It should be pointed out that the factor '$1-\theta_k$' is used in the formulas of $\hat{\lambda}_k$ and t''_k in order to guarantee the positiveness of the iterate (x^k, y^k). Other factors can be used provided that they are less than 1 and tend to zero as $k \to \infty$. However, such factors are not needed for maintaining the positiveness of iterates in the cases such as $0 < \beta < \min_{1\le i\le n} a_i$.

We have the following result on which Step 5 and Step 6 of Algorithm 2.1 are based.

Proposition 2.1. *Let $\sigma \in (0,1)$ be given as in Algorithm 2.1. Suppose that $(x,y) > 0$, $\theta \in (0,1)$, and $(x,y) \in \mathcal{N}_\beta(\theta)$. Let $(\Delta x, \Delta y)$ be the solution to the system:*

$$\mathcal{H}(x, y, \theta) + \nabla_{(x,y)}\mathcal{H}(x,y,\theta)(\Delta x, \Delta y) = 0.$$

(i) If $\mathcal{H}(x,y,\theta) = 0$, then there exists a scalar $\gamma^* \in (0,1)$ such that $(x,y) \in \mathcal{N}_\beta((1-\gamma)\theta)$ for all $\gamma \in (0, \gamma^*]$.

(ii) If $\mathcal{H}(x,y,\theta) \ne 0$, then there exists a $\lambda^* \in (0,1)$ such that

$$\|\mathcal{H}(x + \lambda\Delta x, y + \lambda\Delta y, \theta)\|_\infty \le (1-\sigma\lambda)\|\mathcal{H}(x,y,\theta)\|_\infty \quad \text{for all } \lambda \in (0, \lambda^*]. \tag{10}$$

Let $\mathcal{B} := \{(x', y') > 0 : x' = x + \lambda'\Delta x, y' = y + \lambda'\Delta y, \lambda' \in (0, \lambda^*]\}$. For any $(x', y') \in \mathcal{B}$, there exists a $\gamma^* \in (0,1)$ such that $(x', y') \in \mathcal{N}_\beta((1-\gamma)\theta)$ for all $\gamma \in (0, \gamma^*]$.

Proof. If $\mathcal{H}(x,y,\theta) = 0$, by continuity it is evident that $\|\mathcal{H}(x, y, (1-\gamma)\theta\|_\infty \le \beta(1-\gamma)\theta$ for all small $\gamma > 0$. Thus result (i) follows. We now prove the case that $\mathcal{H}(x,y,\theta) \ne 0$. By Lemma 2.1, the matrix $\nabla_{(x,y)}\mathcal{H}(x,y,\theta)$ is nonsingular, and thus $(\Delta x, \Delta y)$ is well-defined. By differentiability of $\mathcal{H}$, for all sufficiently small λ we have

$$\begin{aligned}
\|\mathcal{H}(x + \lambda\Delta x, y + \lambda\Delta y, \theta)\|_\infty &= \|\mathcal{H}(x,y,\theta) + \lambda\nabla_{(x,y)}\mathcal{H}(x,y,\theta)(\Delta x, \Delta y)\|_\infty + o(\lambda) \\
&= (1-\lambda)\|\mathcal{H}(x,y,\theta)\|_\infty + o(\lambda) \\
&= (1-\sigma\lambda)\|\mathcal{H}(x,y,\theta)\|_\infty + (\sigma - 1)\lambda\|\mathcal{H}(x,y,\theta)\|_\infty + o(\lambda).
\end{aligned}$$

Since $\sigma < 1$, there exists a small $\lambda^* \in (0,1)$ such that for all $\lambda \in (0, \lambda^*]$ the inequality (10) holds. We now define $\varphi : (0,1] \to R$ by

$$\varphi(\gamma) = \frac{[1 - \theta - (1 - (1-\gamma)\theta)(1-\gamma)^p]\theta^p}{\theta}.$$

For any fixed $(x', y') \in \mathcal{B}$, we have

$$\begin{aligned}
\|\mathcal{H}(x', y', (1-\gamma)\theta)\|_\infty &\leq \|\mathcal{H}(x', y', (1-\gamma)\theta) - \mathcal{H}(x', y', \theta)\|_\infty + \|\mathcal{H}(x', y', \theta)\|_\infty \\
&\leq \left\| \begin{pmatrix} x'y' - (1-\gamma)\theta a \\ y' - (1 - (1-\gamma)\theta)(f(x') + (1-\gamma)^p \theta^p x') - (1-\gamma)\theta b \end{pmatrix} \right. \\
&\quad \left. - \begin{pmatrix} x'y' - \theta a \\ y' - (1-\theta)(f(x') + \theta^p x') - \theta b \end{pmatrix} \right\|_\infty + \|\mathcal{H}(x', y', \theta)\|_\infty \\
&= \left\| \begin{pmatrix} \theta\gamma a \\ -\theta\gamma f(x') + \theta\varphi(\gamma)x' + \theta\gamma b \end{pmatrix} \right\|_\infty + \|\mathcal{H}(x', y', \theta)\|_\infty \\
&\leq \theta[\gamma(\|a\|_\infty + \|b\|_\infty + \|f(x')\|_\infty) + \|\varphi(\gamma)x'\|_\infty] + \|\mathcal{H}(x', y', \theta)\|_\infty.
\end{aligned}$$

By (10) and $(x, y) \in \mathcal{N}_\beta(\theta)$, we have

$$\|\mathcal{H}(x', y', \theta)\|_\infty \leq (1 - \sigma\lambda')\|\mathcal{H}(x, y, \theta)\|_\infty \leq (1 - \sigma\lambda')\beta\theta.$$

Combining the above two inequalities yields

$$\begin{aligned}
&\|\mathcal{H}(x', y', (1-\gamma)\theta)\|_\infty \\
&\leq \beta(1-\gamma)\theta \left[\frac{\gamma(\|a\|_\infty + \|b\|_\infty + \|f(x')\|_\infty) + \|\varphi(\gamma)x'\|_\infty}{\beta(1-\gamma)} + \frac{1 - \sigma\lambda'}{1-\gamma} \right].
\end{aligned} \tag{11}$$

Since $1 - \sigma\lambda' < 1$ and $\varphi(\gamma) \to 0$ as $\gamma \to 0$, the term in the bracket of (11) is less than 1 provided that γ is sufficiently small. Therefore, for all sufficiently small $\gamma > 0$ we have $\|\mathcal{H}(x', y', (1-\gamma)\theta_k)\|_\infty \leq \beta(1-\gamma)\theta$, that is, $(x', y') \in \mathcal{N}_\beta((1-\gamma)\theta)$. □

We now prove that the algorithm is well-defined. It is sufficient to prove the proposition below.

Proposition 2.2. For any k, if the current iterate (x^k, y^k, θ_k) satisfies that $(x^k, y^k) > 0$, $\theta_k \in (0,1)$ and $(x^k, y^k) \in \mathcal{N}_\beta(\theta_k)$, then the algorithm can arrive at $(k+1)$th iteration, and the new iterate has the same feature, that is, $(x^{k+1}, y^{k+1}) > 0$, $\theta_{k+1} \in (0,1)$, and $(x^{k+1}, y^{k+1}) \in \mathcal{N}_\beta(\theta_{k+1})$.

Proof. We prove the result by induction. From the initial step, we have $(x^0, y^0) > 0$, $\theta_0 \in (0,1)$, and $(x^0, y^0) \in \mathcal{N}_\beta(\theta_0)$. We now assume that $(x^k, y^k) > 0$, $\theta_k \in (0,1)$, and $(x^k, y^k) \in \mathcal{N}_\beta(\theta_k)$. We prove that the algorithm can proceed to $(k+1)$th iteration and the next iterate satisfies the desired properties. Indeed, by Lemma 2.1, systems (7) and (8) are well defined. Since $(x^k, y^k) > 0$, the scalar t_k in Step 2 is positive. We consider two cases:

Case 1: The criterion '$(\hat{x}^{k+1}, \hat{y}^{k+1}) \in \mathcal{N}_\beta(\hat{\theta}_{k+1})$ and $\hat{\theta}_{k+1} \leq \eta\theta_k$' holds. In this case, (x^{k+1}, y^{k+1}) is generated by Step 3. By the definition of t_k and construction of the algorithm, we see that $(x^{k+1}, y^{k+1}) > 0$, $(x^{k+1}, y^{k+1}) \in \mathcal{N}_\beta(\theta_{k+1})$, and $\theta_{k+1} \leq \eta\theta_k < \theta_k$. The fact $\theta_{k+1} > 0$ follows from the positivity of (x^{k+1}, y^{k+1}). In fact, if $\theta_{k+1} = 0$, then from $(x^{k+1}, y^{k+1}) \in \mathcal{N}_\beta(0)$ it follows that $X^{k+1}y^{k+1} = 0$ which contradicts the positivity of (x^{k+1}, y^{k+1}). Thus the new iterate (x^{k+1}, y^{k+1}) satisfies all desired properties.

Case 2: The above-mentioned criterion does not hold. In this case, since $(x^k, y^k) > 0$, it is evident that $t'_k > 0$, and hence by Proposition 2.1 and construction of Step 5 the new iterate (x^{k+1}, y^{k+1}) maintains positivity. Step 6 indicates that $(x^{k+1}, y^{k+1}) \in \mathcal{N}_\beta(\theta_{k+1})$ where $\theta_{k+1} = (1 - \gamma^k)\theta_k < \theta_k$. By the same proof as in Case 1, we have $\theta_{k+1} > 0$. □

We note that θ_k is monotonically decreasing. In fact, by the construction of the algorithm, we have either $\theta_{k+1} \leq \eta\theta_k$ or $\theta_{k+1} = (1 - \gamma^k)\theta_k$. From the above discussion, we have actually proved the following result.

Lemma 2.2. *If $\{(x^k, y^k, \theta_k)\}$ is generated by Algorithm 2.1, then θ_k is monotonically decreasing, the sequence $\{(x^k, y^k)\}$ is confined to be in R^{2n}_{++}, and*

$$\|\mathcal{H}(x^k, y^k, \theta_k)\|_\infty \leq \beta\theta_k \text{ for every } k. \tag{12}$$

3. Global convergence and properties of accumulation points.

In this section, we address the following questions: Under what condition the sequence $\{(x^k, y^k)\}$ generated by Algorithm 2.1 is globally convergent? and what can be said about accumulation points of this sequence? For convenience, we introduce two auxiliary sequences u^k and v^k:

$$(u^k, v^k) = \mathcal{H}(x^k, y^k, \theta_k)/\theta_k.$$

By (12), the sequence (u^k, v^k) is bounded, i.e., $\|(u^k, v^k)\|_\infty \leq \beta$. By the definition of $\mathcal{H}(\cdot)$, the above equation can be written as

$$X^k y^k = \theta_k(a + u^k), \tag{13}$$

$$y^k = (1-\theta_k)(f(x^k) + (\theta_k)^p x^k) + \theta_k(b + v^k). \tag{14}$$

Clearly, $a + u^k > 0$ since $\theta_k > 0$ and $(x^k, y^k) > 0$. Before showing the global convergence, we first give some useful results.

Lemma 3.1. *Let $f : R^n \to R^n$ be a continuously differentiable P_*-mapping. Assume that the solution set of the NCP is nonempty. Let the sequence $\{(x^k, y^k, \theta_k)\}$ be generated by Algorithm 2.1. Then for any solution x^* of the NCP, the following holds:*

$$\begin{aligned} 0 &\leq (x^*)^T y^k + f(x^*)^T x^k \\ &\leq \theta_k(1+\kappa)e^T(a+u^k) - (1-\theta_k)(\theta_k)^p\kappa|I_+| \min_{1\leq j\leq n} \pi_j^k \\ &\quad -(1-\theta_k)(\theta_k)^p (x^k)^T(x^k - x^*) - \theta_k(x^k - x^*)^T(b + v^k - f(x^*)), \end{aligned} \tag{15}$$

where

$$I_+ = \{i : (x_i^k - x_i^*)(f_i(x^k) - f_i(x^*)) \geq 0\}, \tag{16}$$

$$\pi_j^k = (x_j^k - x_j^*)\left(x_j^k + \frac{(\theta_k)^{1-p}(b_j + v_j^k - f_j(x^*))}{1-\theta_k}\right), \ i = 1, ..., n. \tag{17}$$

Proof. Notice that $(x^k, y^k) > 0$ and x^* is a solution to the NCP. For each i, we have

$$(x_i^k - x_i^*)(y_i^k - f_i(x^*)) = x_i^k y_i^k - x_i^k f_i(x^*) - x_i^* y_i^k + x_i^* f_i(x^*) \leq x_i^k y_i^k.$$

Therefore, by (14), for each i we have

$$\begin{aligned} &(x_i^k - x_i^*)(f_i(x^k) - f_i(x^*)) \\ &= (x_i^k - x_i^*)\left(\frac{y_i^k}{1-\theta_k} - \frac{\theta_k(b_i + v_i^k)}{1-\theta_k} - (\theta_k)^p x_i^k - f_i(x^*)\right) \\ &= \frac{(x_i^k - x_i^*)(y_i^k - f_i(x^*))}{1-\theta_k} - (x_i^k - x_i^*)\left((\theta_k)^p x_i^k + \frac{\theta_k(b_i + v_i^k - f_i(x^*))}{1-\theta_k}\right) \\ &\leq \frac{x_i^k y_i^k}{1-\theta_k} - (\theta_k)^p \min_{1\leq j\leq n} \pi_j^k, \end{aligned} \tag{18}$$

where π_j^k is given by (17). By the P$_*$ property of f, we have

$$(x^k - x^*)^T(f(x^k) - f(x^*)) \geq -\kappa \sum_{i\in I_+} (x_i^k - x_i^*)(f_i(x^k) - f_i(x^*)), \tag{19}$$

where I_+ is defined by (16). Notice that $\sum_{i\in I_+} x_i^k y_i^k \leq (x^k)^T y^k = \theta_k e^T(a + u^k)$. By using (13), (14), (18), and (19), we have

$$\begin{aligned}
0 &\leq (x^*)^T y^k + f(x^*)^T x^k \\
&= -(x^k - x^*)^T(y^k - f(x^*)) + (x^k)^T y^k \\
&= -(x^k - x^*)^T((1-\theta_k)(f(x^k) + (\theta_k)^p x^k) + \theta_k b + \theta_k v^k - f(x^*)) \\
&\quad +\theta_k e^T(a + u^k) \\
&= -(1-\theta_k)(x^k - x^*)^T(f(x^k) - f(x^*)) \\
&\quad -(x^k - x^*)^T[(1-\theta_k)(\theta_k)^p x^k + \theta_k(b + v^k - f(x^*))] + \theta_k e^T(a + u^k) \\
&\leq (1-\theta_k)\kappa \sum_{i\in I_+} (x_i^k - x_i^*)(f_i(x^k) - f_i(x^*)) \\
&\quad -(x^k - x^*)^T[(1-\theta_k)(\theta_k)^p x^k + \theta_k(b + v^k - f(x^*))] + \theta_k e^T(a + u^k) \\
&\leq (1-\theta_k)\kappa \sum_{i\in I_+} \left(\frac{x_i^k y_i^k}{1-\theta_k} - (\theta_k)^p \min_{1\leq j\leq n} \pi_j^k\right) \\
&\quad -(x^k - x^*)^T[(1-\theta_k)(\theta_k)^p x^k + \theta_k(b + v^k - f(x^*))] + \theta_k e^T(a + u^k) \\
&\leq (1-\theta_k)\kappa \left(\theta_k e^T(a + u^k)/(1-\theta_k) - |I_+|(\theta_k)^p \min_{1\leq j\leq n} \pi_j^k\right) \\
&\quad -(x^k - x^*)^T[(1-\theta_k)(\theta_k)^p x^k + \theta_k(b + v^k - f(x^*))] + \theta_k e^T(a + u^k) \\
&= \theta_k(1+\kappa)e^T(a + u^k) - (1-\theta_k)(\theta_k)^p \kappa |I_+| \min_{1\leq j\leq n} \pi_j^k \\
&\quad -(1-\theta_k)(\theta_k)^p (x^k)^T(x^k - x^*) - \theta_k(x^k - x^*)^T(b + v^k - f(x^*)).
\end{aligned}$$

The proof is complete. $\square$

The next result will be utilized to prove the boundedness of the iterative sequence. For a given scalar $\mu > 0$, let $\mathcal{F}_\mu : R^{2n} \to R^{2n}$ be the mapping defined by

$$\mathcal{F}_\mu(x, y) = \begin{pmatrix} Xy \\ y - (f(x) + \mu x) \end{pmatrix}. \tag{20}$$

For any vector $x \leq y$, we denote by $[x, y]$ the rectangle $[x_1, y_1] \times [x_2, y_2] \times ... \times [x_n, y_n]$.

Lemma 3.2 (Lemma 4.1 and Proposition 4.2 in [39]). *If f is a continuous P_0-mapping, then f satisfies the following properties:*

(i) Let z^k be an arbitrary sequence satisfying $\|z^k\| \to \infty$ and $z^k \geq \bar{z}$ for all k, where $\bar{z} \in R^n$ is a vector. Then there exist a subsequence of $\{z^k\}$, denoted by $\{z^{k_j}\}$, and a fixed index i_0 such that $z_{i_0}^{k_j} \to \infty$ and $f_{i_0}(z^{k_j})$ is bounded from below.

(ii) If the solution set of the NCP is nonempty and bounded, then there exists a scalar $\bar{\mu} \in (0, 1)$ such that

$$\bigcup_{\mu\in(0,\bar{\mu}]} \mathcal{F}_\mu^{-1}(D_{\bar{\mu}})$$

is bounded, where $D_{\bar{\mu}} := [0, \bar{\mu}e] \times [-\bar{\mu}e, \bar{\mu}e]$, and

$$\mathcal{F}_\mu^{-1}(D_{\bar{\mu}}) = \{(x, y) \in R_+^{2n} : \mathcal{F}_\mu(x, y) \in D_{\bar{\mu}}\},$$

where $\mathcal{F}_\mu$ is defined by (20).

We are ready to show the boundedness of the sequence $\{(x^k, y^k)\}$.

Theorem 3.1. *Let f be a continuously differentiable P_*-mapping. Assume that the solution set of the NCP is nonempty. Then for any given $p \in (0,1]$ the sequence $\{(x^k, y^k)\}$ is bounded. Moreover, if the solution set is nonempty and bounded, then for any given $p \in (0, \infty)$ the sequence $\{(x^k, y^k)\}$ is bounded.*

Proof. Let x^* be an arbitrary solution of the NCP. By Lemma 3.1, the inequality (15) holds. Dividing both sides of (15) by $(\theta_k)^p$, we have

$$\begin{aligned}&(1-\theta_k)(x^k)^T(x^k - x^*) + (\theta_k)^{1-p}(x^k - x^*)^T(b + v^k - f(x^*))\\ &+(1-\theta_k)\kappa|I_+| \min_{1\le j\le n} \pi_j^k \le (\theta_k)^{1-p}(1+\kappa)e^T(a+u^k).\end{aligned} \tag{21}$$

Since θ_k is monotonically decreasing, i.e., $\theta_k \le ... \le \theta_0 < 1$, we see that $1-\theta_k \ge 1-\theta_0 > 0$. Since (u^k, v^k) is bounded, the right-hand side of (21) is bounded when $p \le 1$. It follows that the sequence $\{x^k\}$ is bounded since otherwise the left-hand side of (21) is unbounded from above. Hence, by (14) and continuity of f, the sequence $\{y^k\}$ is also bounded. The first part of the result is proved.

We now prove that if the solution set is bounded then for any given $p \in (0,\infty)$ the sequence $\{x^k\}$ is bounded, and hence $\{y^k\}$ is also bounded. We show this fact by contradiction. Assume that $\{x^k\}$ is unbounded. Notice that $x^k > 0$ for all k. By Lemma 3.2, there exist an index j and a subsequence of $\{x^k\}$, denoted still by $\{x^k\}$, such that $x_j^k \to \infty$ and $f_j(x^k)$ is bounded from below. Since $x_j^k \to \infty$ and the right-hand side of (13) is bounded, it follows that $y_j^k \to 0$. By (14), we have

$$f_j(x^k) - \frac{y_j^k}{1-\theta_k} + \frac{\theta_k}{1-\theta_k}(b_j + v_j^k) = -(\theta_k)^p x_j^k. \tag{22}$$

Notice that $0 < 1-\theta_0 \le 1-\theta_k < 1$, $y_j^k \to 0$, and v^k is bounded. The left-hand side of (22) is bounded from below. This together with the fact $x_j^k \to \infty$ implies that θ_k tends to zero. Denote by $\bar{y}^k = y^k/(1-\theta_k)$. From (13) and (14), we have

$$\mathcal{F}_{(\theta_k)^p}(x^k, \bar{y}^k) = \begin{pmatrix} X^k\bar{y}^k \\ \bar{y}^k - (f(x^k) + (\theta_k)^p x^k)\end{pmatrix} = \begin{pmatrix} \theta_k(a+u^k)/(1-\theta_k) \\ \theta_k(b+v^k)/(1-\theta_k)\end{pmatrix}.$$

Since $\theta_k \to 0$, there exists a k_0 such that for all $k \ge k_0$ we have $(\theta_k)^p \le \bar{\mu}$ and

$$\mathcal{F}_{(\theta_k)^p}(x^k, \bar{y}^k) \in D_{\bar{\mu}} = [0, \bar{\mu}e] \times [-\bar{\mu}e, \bar{\mu}e],$$

where $\bar{\mu}$ is the scalar defined in Lemma 3.2. Thus

$$\{(x^k, \bar{y}^k)\}_{k\ge k_0} \subseteq \bigcup_{k\ge k_0} \mathcal{F}^{-1}_{(\theta_k)^p}(D_{\bar{\mu}}) \subseteq \bigcup_{\mu\in(0,\bar{\mu}]} \mathcal{F}_\mu^{-1}(D_{\bar{\mu}}).$$

By Lemma 3.2, the right hand-side of the above is bounded. This contradicts the unboundedness of the left hand-side. □

Before proving the next result, we first introduce the following condition.

Condition 3.1. *There exists a constant $C > 0$ such that*

$$\|[\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k)]^{-1}\| \le C \tag{23}$$

for all sufficiently large k.

Similar conditions were used in several work such as [4, 5, 6, 34]. Condition 3.1 can be stated in other versions. For instance, by using Burke and Xu-type assumption (see [5]), we can state the condition as follows: *Given* $\beta > 0$ *and* $\theta_0 \in (0,1)$, *there exists a* $C > 0$ *such that* $\|[\nabla_{(x,y)}\mathcal{H}(\hat{x},\hat{y},\hat{\theta})]^{-1}\| \le C$ *for all* $0 < \hat{\theta} \le \theta_0$ *and* $(\hat{x},\hat{y}) \in \mathcal{N}_\beta(\hat{\theta})$. It is evident that this Burke and Xu-type assumption implies Condition 3.1. We now consider the following assumption used in [31, 33]: *There exists a solution* x^* *that satisfies* $x^* + f(x^*) > 0$ *(i.e.,* x^* *is a strictly complementary solution),* $\nabla f(x^*)_{II}$ *is nonsingular, where* $I = \{i : x_i^* > 0\}$, *and* $\nabla f(x)$ *is Lipschitz continuous in the neighborhood of* x^*. We now point out that this condition implies Condition 3.1. In fact, under this assumption, it is easy to verify (see, for example, Tseng [31]) that the matrix

$$\nabla_{(x,y)}\mathcal{H}(x^*,y^*,0) = \begin{bmatrix} Y^* & X^* \\ -\nabla f(x^*) & E \end{bmatrix} \tag{24}$$

is nonsingular for a P_0 NCP, where $y^* = f(x^*)$. Therefore, Condition 3.1 holds in the local area of (x^*,y^*). (We also mention that the above assumption in [31, 33] actually implies that the solution is unique. In fact, the nonsingularity of (24) implies that (x^*,y^*) is a locally isolated solution (Proposition 2.5 in [25]), and this further implies that (x^*,y^*) is the unique solution for a P_0 NCP (see [17])).

The following result shows that the parameter θ_k can be reduced to zero. It should be pointed out that if the constant β is taken relatively small as in the case (i) of the next theorem, then Condition 3.1 is not required. Furthermore, in this case, the global convergence of the Algorithm 2.1 also does not require Condition 3.1 (see Theorem 3.3 in details). We only impose Condition 3.1 for the global convergence analysis with $\beta \ge \min_{1\le i\le n} a_i$ and for the local convergence analysis in Section 4.

Theorem 3.2. *Let* f *be a continuously differentiable* P_**-mapping. Assume that the solution set of the NCP is nonempty.*

(i) Let β *be taken such that* $0 < \beta < \min_{1\le i\le n} a_i$. *Then*

$$\lim_{k\to\infty} \theta_k = 0 \text{ and } \lim_{k\to\infty} \|\mathcal{H}(x^k,y^k,\theta_k)\|_\infty = 0. \tag{25}$$

(ii) Let β *be taken such that* $\beta \ge \min_{1\le i\le n} a_i$. *If Condition 3.1 is satisfied for* $\{(x^k,y^k,\theta_k)\}$, *then (25) remains valid.*

Proof. Since θ_k is monotonically decreasing, there exists a $\tilde{\theta} \ge 0$ such that $\theta_k \to \tilde{\theta}$. Note that $\|\mathcal{H}(x^k,y^k,\theta_k)\|_\infty \le \beta\theta_k$ for all k (see Lemma 2.2). It suffices to show that $\tilde{\theta} = 0$. If Step 3 is accepted infinitely many times, then it is evident that $\tilde{\theta} = 0$. We now consider the case that Step 3 is accepted only finite many times. That is, there exists a $\bar{k}$ such that for all $k > \bar{k}$ the sequence $\{(x^k,y^k,\theta_k)\}$ is generated by Step 4 through Step 6 of the algorithm. We show that $\tilde{\theta} = 0$ remains valid. Assume contrarily that $\tilde{\theta} > 0$. We now derive a contradiction. Notice that $\theta_{k+1} = (1-\gamma^k)\theta_k$ for any $k > \bar{k}$ and that $\theta_k \to \tilde{\theta} > 0$, we conclude that $\gamma^k \to 0$. Thus, from Step 6, it follows that $(x^{k+1},y^{k+1}) \notin \mathcal{N}_\beta((1-\frac{1}{\alpha_2}\gamma^k)\theta_k)$ for all sufficiently large k, that is,

$$\left\|\mathcal{H}(x^{k+1},y^{k+1},(1-\gamma^k/\alpha_2)\theta_k)\right\|_\infty > \beta(1-\gamma^k/\alpha_2)\theta_k. \tag{26}$$

If $p \le 1$, by Theorem 3.1 the sequence $\{(x^k,y^k)\}$ is bounded. We now consider the case of $p > 1$. Let x^* be an arbitrary solution of the NCP. By Lemma 3.1, we have

$$\begin{aligned}&(1-\theta_k)(\theta_k)^p(x^k)^T(x^k-x^*) + \theta_k(x^k-x^*)^T(b+v^k-f(x^*)) \\ &+(1-\theta_k)(\theta_k)^p\kappa|I_+|\min_{1\le j\le n}\pi_j^k \le \theta_k(1+\kappa)e^T(a+u^k).\end{aligned}$$

Since $\tilde{\theta} > 0$, it follows that $1 \ge (1-\theta_k)(\theta_k)^p \ge (1-\theta_0)(\tilde{\theta})^p > 0$. Notice that the right-hand side of the above inequality is bounded. We conclude that the sequence $\{x^k\}$ is bounded since otherwise the left-hand side is unbounded from above. This implies that $\{y^k\}$ is also bounded by (14) and the continuity of f. Hence, the sequence $\{(x^k, y^k)\}$ is bounded under the assumption of $\tilde{\theta} > 0$. Taking a subsequence if necessary, we may assume that (x^k, y^k) converges to $(\tilde{x}, \tilde{y})$. Clearly, $(\tilde{x}, \tilde{y}) \geq 0$. Since $\gamma^k \to 0$, taking the limit in (26) leads to $\|\mathcal{H}(\tilde{x}, \tilde{y}, \tilde{\theta})\|_\infty \geq \beta\tilde{\theta} > 0$. On the other hand, if follows from $\|\mathcal{H}(x^k, y^k, \theta_k)\|_\infty \leq \beta\theta_k$ that $\|\mathcal{H}(\tilde{x}, \tilde{y}, \tilde{\theta})\|_\infty \leq \beta\tilde{\theta}$. Therefore,

$$\|\mathcal{H}(\tilde{x}, \tilde{y}, \tilde{\theta})\|_\infty = \beta\tilde{\theta}.$$

Case (i): $0 < \beta < \min_{1\leq i\leq n} a_i$. The above equation implies that $\|\tilde{X}\tilde{y} - \tilde{\theta}a\|_\infty \leq \beta\tilde{\theta}$. Since $\beta < \min_{1\leq i\leq n} a_i$, this further implies that $\tilde{x}_i\tilde{y}_i \neq 0$ for every i. Thus, we have $(\tilde{x}, \tilde{y}) > 0$. Since $\tilde{\theta} \in (0, 1)$, by Lemma 2.1, the matrix $\nabla_{(x,y)}\mathcal{H}(\tilde{x}, \tilde{y}, \tilde{\theta})$ is nonsingular.

Case (ii): There exists a constant $C > 0$ such that $\|[\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k)]^{-1}\| \leq C$ for all sufficiently large k. This implies that the matrix $\nabla_{(x,y)}\mathcal{H}(\tilde{x}, \tilde{y}, \tilde{\theta})$ is nonsingular. Moreover, by Lemma 2.1, we have $\tilde{x} + \tilde{y} > 0$.

Both of the above two cases imply the nonsingularity of the matrix $\nabla_{(x,y)}\mathcal{H}(\tilde{x}, \tilde{y}, \tilde{\theta})$. Thus the following system has a unique solution, denoted by $(\Delta\tilde{x}, \Delta\tilde{y})$:

$$\mathcal{H}(\tilde{x}, \tilde{y}, \tilde{\theta}) + \nabla_{(x,y)}\mathcal{H}(\tilde{x}, \tilde{y}, \tilde{\theta})(\Delta x, \Delta y) = 0. \tag{27}$$

Then, $(\Delta\tilde{x}, \Delta\tilde{y})$ is a strictly descent direction of $\|\mathcal{H}(x, y, \theta)\|_\infty$ at $(\tilde{x}, \tilde{y}, \tilde{\theta})$. As a result, under one of the cases (i) and (ii), the line search stepsizes $\tilde{\lambda}$ in Step 5 and $\tilde{\gamma}$ in Step 6 are both bounded from below by a positive constant. In fact, in the first case, we have $(\tilde{x}, \tilde{y}) > 0$, and hence the value of $\tilde{t}'$, defined in Step 4, is positive. Thus, $(\tilde{\lambda}, \tilde{\gamma}) > 0$ according to Proposition 2.1. For the second case, we note that $\tilde{x} + \tilde{y} > 0$. This implies that for every component i, at least one of $\tilde{x}_i$ and $\tilde{y}_i$ must be positive. We now assume without loss of generality that $\tilde{x}_i = 0$. It follows from (27) that

$$\tilde{y}_i\Delta\tilde{x}_i = \tilde{x}_i\Delta\tilde{y}_i + \tilde{y}_i\Delta\tilde{x}_i = \tilde{\theta}a_i - \tilde{x}_i\tilde{y}_i = \tilde{\theta}a_i > 0,$$

which implies that $\Delta\tilde{x}_i > 0$. Thus the value of $\tilde{t}'$, defined in Step 4, is positive. By Proposition 2.1, we conclude that $\tilde{\lambda} > 0$ and $\tilde{\gamma} > 0$. Since $\mathcal{H}$ and $\nabla_{(x,y)}\mathcal{H}$ are continuous in the neighborhood of $(\tilde{x}, \tilde{y}, \tilde{\theta})$, it is easy to see that there exists a constant $\tilde{\alpha}$ such that $\gamma^k \geq \tilde{\alpha}\tilde{\gamma}$ for any iterate. This contradicts that $\gamma^k \to 0$. The proof is complete. $\Box$

Note that if we use Strategy 2.2 or Strategy 2.3 to construct the initial values of the algorithm, β can be fixed at any value in $(0, \infty)$. In particular, β can be fixed such that $0 < \beta < \min_{1\leq i\leq n} a_i$.

Combining Theorems 3.1 and 3.2, we can easily obtain the main global convergence result of this paper. We first catalog some concepts about the solution of an NCP. Let S^* denote the solution set of an NCP. Recall that an element $x^* \in S^*$ is said to be the least element solution if $x^* \leq x'$ for all $x' \in S^*$ (see, for instance, [7]). An element $x^* \in S^*$ is said to be the least 2-norm solution if $\|x^*\| \leq \|x'\|$ for all $x' \in S^*$. Clearly, the least element solution, if exists, coincides with the least 2-norm solution. The following concept is a generalization of the least 2-norm solution.

Definition 3.1. *An element x^* is said to be a scaled least 2-norm solution in S^* if there exists a nonnegative scalar $\delta' \geq 0$ such that for every element $x' \in S^*$ there exists a corresponding diagonal matrix M with one of its diagonal entries being $1 + \delta'$ and all other diagonal entries being 1's such that*

$$(x^*)^T M(x^* - x') \leq 0.$$

In another word, x^* is said to be a scaled least 2-norm solution if there exists a nonnegative scalar $\delta' \geq 0$ such that for every element $x' \in S^*$ there exists a corresponding index i such that

$$\delta' x_i^*(x_i^* - x_i') + (x^*)^T(x^* - x') \leq 0.$$

When $\delta' = 0$, the scaled least 2-norm solution is the least 2-norm solution (see the proof of Theorem 3.5). We also note that if the least element solution exists, the scaled least 2-norm solution coincides with this least element solution. We now state the main global convergence result of this paper. This result also characterizes the accumulation points of the iterative sequence.

Theorem 3.3. *Let f be a continuously differentiable P_*-mapping. Assume that the solution set of the NCP is nonempty.*

(i) Let $p \in (0, 1]$ and $0 < \beta < \min_{1\le i\le n} a_i$. Then the sequence $\{(x^k, y^k)\}$ generated by Algorithm 2.1 has at least one accumulation point, and every accumulation point of the sequence is a solution to the NCP. Furthermore, if $p < 1$ every accumulation point of $\{(x^k, y^k)\}$ is a scaled least 2-norm solution. In particular, if the least element solution exists, then for $p < 1$ the entire sequence $\{(x^k, y^k)\}$ converges to the least element solution.

(ii) Let $p \in (0, \infty)$ and $0 < \beta < \min_{1\le i\le n} a_i$. If the solution set of the NCP is bounded, then the sequence $\{(x^k, y^k)\}$ generated by Algorithm 2.1 has at least one accumulation point, and each accumulation point of this sequence is a solution to the NCP.

Proof. The boundedness of the sequence $\{(x^k, y^k)\}$ follows from Theorem 3.1. Thus $\{(x^k, y^k)\}$ has at least one accumulation point. Let $(\hat{x}, \hat{y})$ be an arbitrary accumulation point. By Theorem 3.2 and continuity of $\mathcal{H}$, we have $0 = \lim_{k\to\infty} \|\mathcal{H}(x^k, y^k, \theta_k)\| = \|\mathcal{H}(\hat{x}, \hat{y}, 0)\|$, which implies that $(\hat{x}, \hat{y})$ is a solution to the NCP. Thus, every accumulation point is a solution of the NCP. Furthermore, let $p < 1$ and x^* be an arbitrary solution to the NCP. Taking the limit in (21) and noting that $\theta_k \to 0$ and $p < 1$, we see that there exists an index i_0 (corresponding to $\hat{x}$ and x^*) such that

$$(\hat{x})^T(\hat{x} - x^*) + \kappa n \hat{x}_{i_0}(\hat{x}_{i_0} - x^*_{i_0}) \le 0. \tag{28}$$

By Definition 3.1, $\hat{x}$ is a scaled least 2-norm solution. In particular, if the least element solution exists, by setting x^* to be the least element solution, we see from (28) that $\hat{x}$ must equal to the least element solution. Since the least element solution is unique, the entire sequence $\{x^k\}$ is convergent, and so is $\{y^k\}$ by (14) and continuity of f. □

A feature of the result (i) above (and Theorem 3.6 below) is that no assumption is required for the global convergence other than the existence of a solution to the problem. This result, however, requires that β be taken relatively small. As a result, the neighborhood is possibly narrow.

From the viewpoint of numerical implementation, β should be taken relatively large such that the neighborhood is wide enough to permit a large stepsize at each iteration in order to achieve a fast convergence. Thus, the next result, an immediate consequence by combining Theorems 3.1 and 3.2 and by using the proof of Theorem 3.3, is concerned with the case of wide neighborhoods. However, such a result requires Condition 3.1, which essentially requires that the solution be strictly complementary. In fact, if (x^*, y^*, θ^*) is an accumulation point of the sequence (x^k, y^k, θ^k) generated by Algorithm 2.1, Condition 3.1 implies that the matrix $\nabla_{(x,y)}\mathcal{H}(x^*, y^*, \theta^*)$ is nonsingular, and hence by Lemma 2.1, (x^*, y^*) must be strictly complementary.

Theorem 3.4. *Let f be a continuously differentiable P_*-mapping. Assume that the solution set of the NCP is nonempty.*

(i) Let $p \in (0, 1]$ and $\beta \ge \min_{1\le i\le n} a_i$. If Condition 3.1 is satisfied, then all the results in part (i) of Theorem 3.3 hold.

(ii) Let $p \in (0, \infty)$ and $\beta \ge \min_{1\le i\le n} a_i$. If the solution set of the NCP is bounded and Condition 3.1 is satisfied, then the same results as in part (ii) of Theorem 3.3 hold.

The existence of the least element solution is not always assured. However, the least 2-norm solution always exists provided that the solution set is nonempty. The next result shows when the sequence $\{(x^k, y^k)\}$ converges to the least 2-norm solution.

Theorem 3.5. *Let f be a continuously differentiable monotone mapping. Assume that the solution set of the NCP is nonempty.*

(i) Let $p \in (0,1)$ and $0 < \beta < \min_{1\le i\le n} a_i$. Then the sequence $\{(x^k, y^k)\}$ converges to $(\hat{x}, \hat{y})$ where $\hat{y} = f(\hat{x})$ and $\hat{x}$ is the least 2-norm solution.

(ii) Let $p \in (0,1)$ and $\beta \ge \min_{1\le i\le n} a_i$. If Condition 3.1 is satisfied, then the result of (i) above remains valid.

Proof. Let $(\hat{x}, \hat{y})$ be an accumulation point of $\{(x^k, y^k)\}$. Since each monotone mapping is a P_*-mapping with $\kappa = 0$, by Theorems 3.3 and 3.4, $\hat{x}$ is a solution to the NCP. Setting $\kappa = 0$ in (28), we have

$$(\hat{x})^T(\hat{x} - x^*) \le 0 \text{ for all } x^* \in S^*. \tag{29}$$

We now show that $\hat{x}$ must be unique. Assume that there exists another solution x' such that

$$(x')^T(x' - x^*) \le 0 \text{ for all } x^* \in S^*. \tag{30}$$

Setting $x^* = x'$ in (29) and $x^* = \hat{x}$ in (30), respectively, and adding two inequalities, we have that $(\hat{x} - x')^T(\hat{x} - x') \le 0$ which indicates that $\hat{x} = x'$. Therefore the accumulation point of iterates is unique, and hence $\{(x^k, y^k)\}$ is convergent. It follows from (29) that $\|\hat{x}\|^2 \le (\hat{x})^T x^* \le \|\hat{x}\|\|x^*\|$ for all $x^* \in S^*$. Therefore, $\hat{x}$ is the least 2-norm solution. □

The results established so far do not cover the case where the solution set is unbounded and p is taken from $(1, \infty)$. For completeness, the remainder of this section is devoted to the study of this case.

By a closer inspection of the previous proofs, we conclude that b can take any vector in R^n provided that the initial conditions $(x^0, y^0) > 0$ and $(x^0, y^0) \in \mathcal{N}_\beta(\theta_0)$ are satisfied. In what follows, we restrict the vector b to be in R^n_{++}, to show the global convergence for the above-mentioned case. We first recall the concept of the maximally complementary solution that has been widely used in the literature. Denote

$$I = \{i : x_i^* > 0, \text{ for some } x^* \in S^*\}, \tag{31}$$

$$J = \{j : f_j(x^*) > 0, \text{ for some } x^* \in S^*\}, \tag{32}$$

$$O = \{k : x_k^* = f_k(x^*) = 0, \text{ for all } x^* \in S^*\}. \tag{33}$$

A solution x^* is said to be the maximally complementary solution of an NCP if $x_i^* > 0$ for all $i \in I$, $f_j(x^*) > 0$ for all $j \in J$, and $x_k^* = f_k(x^*) = 0$ for all $k \in O$. When $O = \emptyset$, i.e., $x^* + f(x^*) > 0$, x^* is called a strictly complementary solution. We now prove the following result.

Theorem 3.6. *Let f be a continuously differentiable P_*-mapping. Suppose that the solution set of the NCP is nonempty. Let p, b, β be chosen such that $p > 1$, $b \in R^n_{++}$, and $0 < \beta < \min_{1\le i\le n} \min\{a_i, b_i\}$, then the sequence $\{(x^k, y^k)\}$ generated by Algorithm 2.1 is bounded, and every accumulation point of the sequence is a maximally complementary solution of the NCP.*

Proof. We first show the boundedness of $\{(x^k, y^k)\}$. By the choice of b and β, it follows that $b + v^k \ge b - \beta e > 0$, and thus,

$$\begin{aligned}
\pi_j^k &\ge -x_j^k x_j^* + (\theta_k)^{1-p}(x_j^k - x_j^*)(b_j + v_j^k - f_j(x^*))/(1-\theta_k) \\
&\ge -x_j^k x_j^* - (\theta_k)^{1-p}(x_j^k f_j(x^*) + x_j^*(b_j + v_j^k))/(1-\theta_k) \\
&\ge -(x^k)^T x^* - (\theta_k)^{1-p}((x^k)^T f(x^*) + (b + v^k)^T x^*)/(1-\theta_k).
\end{aligned}$$

This together with (15) implies that

$$\begin{aligned}
(x^*)^T y^k + f(x^*)^T x^k \le\ & \theta_k(1+\kappa)e^T(a + u^k) - (1-\theta_k)(\theta_k)^p \kappa |I_+|[-(x^k)^T x^* \\
& -(\theta_k)^{1-p}((x^k)^T f(x^*) + (b + v^k)^T x^*)/(1-\theta_k)] \\
& +(1-\theta_k)(\theta_k)^p (x^k)^T x^* - \theta_k(x^k - x^*)^T(b + v^k) + \theta_k(x^k)^T f(x^*).
\end{aligned}$$

Thus,

$$\begin{array}{rcl}(x^*)^T y^k + f(x^*)^T x^k & \leq & \theta_k(1+\kappa)e^T(a+u^k) \\ & & +(1-\theta_k)(\theta_k)^p(1+\kappa|I_+|)(x^k)^T x^* + \theta_k(1+\kappa|I_+|)(x^k)^T f(x^*) \\ & & -\theta_k(x^k)^T(b+v^k) + \theta_k(1+\kappa|I_+|)(b+v^k)^T x^*.\end{array}$$

That is,

$$\begin{array}{rl} & (x^*)^T y^k + [1-\theta_k(1+\kappa|I_+|)]f(x^*)^T x^k \\ \leq & \theta_k(1+\kappa)e^T(a+u^k) - \theta_k(x^k)^T[b+v^k \\ & -(1-\theta_k)(\theta_k)^{p-1}(1+\kappa|I_+|)x^*] + \theta_k(1+\kappa|I_+|)(b+v^k)^T x^*. \end{array} \tag{34}$$

Since $\theta_k \to 0$ (by Theorem 3.2) and $(x^k, y^k) > 0$, the left-hand side of the above inequality is nonnegative for all sufficiently large k. Thus, for all sufficiently large k we have

$$(x^k)^T[b+v^k-(1-\theta_k)(\theta_k)^{p-1}(1+\kappa|I_+|)x^*] \leq (1+\kappa)e^T(a+u^k) + (1+\kappa|I_+|)(b+v^k)^T x^*.$$

Since $p > 1$, $\theta_k \to 0$, and $b+\beta e \geq b+v^k \geq b-\beta e > 0$, for all large k we have that

$$(b+v^k) - (1-\theta_k)(\theta_k)^{p-1}(1+\kappa|I_+|)x^* \geq (b-\beta e)/2 > 0.$$

Thus, for all sufficiently large k, we have

$$(x^k)^T(b-\beta e)/2 \leq (1+\kappa)e^T(a+\beta e) + (1+\kappa|I_+|)(b+\beta e)^T x^*$$

which implies that the sequence $\{x^k\}$ is bounded, and so is $\{y^k\}$ by (14) and continuity of f. Since $\theta_k \to 0$ and $\|\mathcal{H}(x^k, y^k, \theta_k)\|_\infty \to 0$, every accumulation point of $\{(x^k, y^k)\}$ is a solution to the NCP.

We now prove that each accumulation point of $\{x^k\}$ is a maximally complementary solution. Since $(1-\theta_k)(\theta_k)^p \leq \theta_k$ and $b+v^k > 0$, the inequality (34) can be further written as

$$\begin{array}{l}(x^*)^T y^k + (1-\theta_k(1+\kappa|I_+|))f(x^*)^T x^k \\ \leq \theta_k\left[(1+\kappa)e^T(a+u^k) + (1+\kappa|I_+|)(x^*)^T x^k + (1+\kappa|I_+|)(b+v^k)^T x^*\right].\end{array}$$

Let I, J be defined by (31) and (32). Then $x_i^* = 0$ for all $i \notin I$ and $f_j(x^*) = 0$ for all $j \notin J$. The above inequality can be written as

$$\begin{array}{l}(x^*)_I^T y_I^k + (1-\theta_k(1+\kappa|I_+|))f_J(x^*)^T x_J^k \\ \leq \theta_k\left[(1+\kappa)e^T(a+u^k) + (1+\kappa|I_+|)(x_I^*)^T x_I^k + (1+\kappa|I_+|)(b+v^k)_I^T x_I^*\right]\end{array}$$

i.e.,

$$\begin{array}{l}(x^*)_I^T(X_I^k)^{-1}X_I^k y_I^k + (1-\theta_k(1+\kappa|I_+|))f_J(x^*)^T(Y_J^k)^{-1}Y_J^k x_J^k \\ \leq \theta_k\left[(1+\kappa)e^T(a+u^k) + (1+\kappa|I_+|)(x_I^*)^T x_I^k + (1+\kappa|I_+|)(b+v^k)_I^T x_I^*\right].\end{array}$$

By (13), the above inequality can be written as

$$\begin{array}{l}(x^*)_I^T(X_I^k)^{-1}(a_I+u_I^k) + (1-\theta_k(1+\kappa|I_+|))f_J(x^*)^T(Y_J^k)^{-1}(a_J+u_J^k) \\ \leq (1+\kappa)e^T(a+u^k) + (1+\kappa|I_+|)(x_I^*)^T x_I^k + (1+\kappa|I_+|)(b+v^k)_I^T x_I^*.\end{array}$$

Thus, we have

$$(x_i^*/x_i^k)(a_i+u_i^k) \leq (1+\kappa)e^T(a+u^k) + (1+\kappa|I_+|)(x_I^*)^T x_I^k + (1+\kappa|I_+|)(b+v^k)_I^T x_I^* \tag{35}$$

for all $i \in I$, and

$$
\begin{aligned}
&(1-\theta_k(1+\kappa|I_+|))(f_j(x^*)/y_j^k)(a_j+u_j^k)\\
&\leq (1+\kappa)e^T(a+u^k)+(1+\kappa|I_+|)(x_I^*)^T x_I^k+(1+\kappa|I_+|)(b+v^k)_I^T x_I^* \qquad (36)
\end{aligned}
$$

for all $j \in J$. The inequalities (35) and (36) hold for any solution x^* of the NCP. Since $\{(x^k, y^k)\}$ is bounded, it has at least one accumulation point. Assume that $(\hat{x}, \hat{y})$, where $\hat{y} = f(\hat{x})$, is an accumulation point of this sequence. To show $(\hat{x}, \hat{y})$ to be a maximally complementary solution, it is sufficient to show that $\hat{x}_I > 0$ and $\hat{y}_J > 0$. By the choice of u^k, we have that $a + \beta e \geq a + u^k \geq a - \beta e > 0$. Notice that the right-hand sides of (35) and (36) are bounded. Since (35) holds for any solution x^*, and since for each $i \in I$, there exists an $x^* \in S^*$ such that $x_i^* > 0$. It follows that $\hat{x}_I > 0$ for all $i \in I$. Otherwise, if $\hat{x}_i = 0$ for some $i \in I$, by the definition of I, there must exist an x^* such that $x_i^* > 0$. Then $x_i^*/x_i^k \to \infty$, and hence the left-hand side of the inequality (35) is unbounded, contradicting to the boundedness of the right-hand side. Thus, $\hat{x}_I > 0$. Similarly, it follows from (36) that $\hat{y}_j = f_j(\hat{x}) > 0$ for all $j \in J$. □

The requirement of $b > 0$ and $0 < \beta < \min_{1\leq i\leq n} \min\{a_i, b_i\}$ can be satisfied if we use Strategy 2.3 to obtain the initial points and parameters for our algorithm (see Remark 3.1).

Since $\|\mathcal{H}(x^k, y^k, \theta_k)\|_\infty/\theta_k \leq \beta$, the sequence $\mathcal{H}(x^k, y^k, \theta_k)/\theta_k$ is possibly convergent. If so, we claim that a convergence result stronger than Theorem 3.6 can be obtained for monotone linear complementarity problems.

Theorem 3.7. *Let $f = Mx + q$ where $q \in R^n$ and M is a positive semi-definite matrix. Assume that the solution set of the linear complementarity problem is nonempty. Let p, b, β be chosen such that $p > 1, b \in R^n_{++}$ and $0 < \beta < \min_{1\leq i\leq n} \min\{a_i, b_i\}$. If $\lim_{k\to\infty} \mathcal{H}(x^k, y^k, \theta_k)/\theta_k = (\hat{u}, \hat{v})$, i.e., $\lim_{k\to\infty}(u^k, v^k) = (\hat{u}, \hat{v})$, then the entire sequence $\{(x^k, y^k)\}$ converges to $(\hat{x}, \hat{y})$, where $\hat{y} = f(\hat{x})$ and $\hat{x}$ is a maximally complementary solution.*

Proof. By Theorem 3.6, the sequence (x^k, y^k) is bounded, and each accumulation point is the maximally complementary solution of the problem. It is sufficient to show that the accumulation point of (x^k, y^k) is unique. Let I, J, O be defined by (31)-(33). It is well-known that the partition I, J, O is unique. Consider the following affine set

$$\tilde{S} = \{(x, y) \in R^{2n} : \ x_{J\cup O} = 0, \ y_{I\cup O} = 0, \ Mx - y + q = 0\},$$

which is the smallest affine set containing the solution set. By assumption, $\tilde{S} \neq \emptyset$. Let $(\tilde{x}, \tilde{y})$ be an arbitrary point in $\tilde{S}$. Notice that $\tilde{x}^T\tilde{y} = 0$ and $\tilde{y} = f(\tilde{x}) = M\tilde{x} + q$. We have

$$
\begin{aligned}
\tilde{x}^T y^k + \tilde{y}^T x^k &= -(x^k-\tilde{x})^T(y^k-\tilde{y}) + (x^k)^T y^k\\
&= -(x^k-\tilde{x})^T[(1-\theta_k)(f(x^k)+(\theta_k)^p x^k)+\theta_k(b+v^k)-f(\tilde{x})]+(x^k)^T y^k\\
&= -(1-\theta_k)(x^k-\tilde{x})^T(f(x^k)-f(\tilde{x}))\\
&\quad -(x^k-\tilde{x})^T[(1-\theta_k)(\theta_k)^p x^k+\theta_k(b+v^k-f(\tilde{x}))]+(x^k)^T y^k\\
&\leq -(x^k-\tilde{x})^T[(1-\theta_k)(\theta_k)^p x^k+\theta_k(b+v^k-f(\tilde{x}))]+\theta_k e^T(a+u^k).
\end{aligned}
$$

The last inequality follows from the monotonicity of f and (13). Since $\tilde{x}_{J\cup O} = 0$ and $\tilde{y}_{I\cup O} = 0$, by the same proof of Theorem 3.5, we have

$$
\begin{aligned}
&\tilde{x}_I^T(X_I^k)^{-1}(a_I+u_I^k)+\tilde{y}_J^T(Y_J^k)^{-1}(a_J+u_J^k)\\
&\leq -(x^k-\tilde{x})^T[(1-\theta_k)(\theta_k)^{p-1}x^k+b+v^k-f(\tilde{x})]+e^T(a+u^k).
\end{aligned}
$$

Let $(\hat{x}, \hat{y})$ be an arbitrary accumulation point of the sequence $\{(x^k, y^k)\}$. As we have pointed out, $(\hat{x}, \hat{y})$ is the maximally complementary solution. Noting that $\theta_k \to 0$ and $p > 1$, from the above inequality we have

$$\begin{aligned}\tilde{x}_I^T \hat{X}_I^{-1}(a_I+\hat{u}_I)+\tilde{y}_J^T\hat{Y}_J^{-1}(a_J+\hat{u}_J) &\leq -(\hat{x}-\tilde{x})^T(b+\hat{v}-f(\tilde{x}))+e^T(a+\hat{u})\\ &= -(\hat{x}_I-\tilde{x}_I)^T(b_I+\hat{v}_I)+e^T(a+\hat{u}). \end{aligned} \tag{37}$$

The last equation follows from the fact that $\hat{x}^T f(\tilde{x}) = \tilde{x}^T f(\tilde{x}) = 0$. The above inequality holds for all $(\tilde{x},\tilde{y}) \in \tilde{S}$. Let $(\bar{x},\bar{y})$ be an arbitrary solution of the problem. Notice that

$$(\bar{x}+\lambda(\hat{x}-\bar{x}), \bar{y}+\lambda(\hat{y}-\bar{y})) \in \tilde{S} \text{ for any } \lambda \in R.$$

Setting $(\tilde{x},\tilde{y}) = (\bar{x}+\lambda(\hat{x}-\bar{x}), \bar{y}+\lambda(\hat{y}-\bar{y}))$ in (37), we have

$$\begin{aligned}&\lambda\left[(\hat{x}_I-\bar{x}_I)^T\hat{X}_I^{-1}(a_I+\hat{u}_I)+(\hat{y}_J-\bar{y}_J)^T\hat{Y}_J^{-1}(a_J+\hat{u}_J)-(\hat{x}_I-\bar{x}_I)^T(b_I+\hat{v}_I)\right]\\ &\leq -\bar{x}_I^T\hat{X}_I^{-1}(a_I+\hat{u}_I)-\bar{y}_J^T\hat{Y}_J^{-1}(a_J+\hat{u}_J)-(\hat{x}_I-\bar{x}_I)^T(b_I+\hat{v}_I)+e^T(a+\hat{u}).\end{aligned}$$

Since the above inequality holds for any $\lambda \in R$, it follows that

$$(\hat{x}_I-\bar{x}_I)^T\hat{X}_I^{-1}(a_I+\hat{u}_I)+(\hat{y}_J-\bar{y}_J)^T\hat{Y}_J^{-1}(a_J+\hat{u}_J) = (\hat{x}_I-\bar{x}_I)^T(b_I+\hat{v}_I).$$

That is,

$$\sum_{i\in I}(a_i+\hat{u}_i)(1-\bar{x}_i/\hat{x}_i)+\sum_{j\in J}(a_j+\hat{u}_j)(1-\bar{y}_j/\hat{y}_j) = (\hat{x}_I-\bar{x}_I)^T(b_I+\hat{v}_I).$$

Since $1-t \leq -\log t$ for all $t>0$, we see from the above that

$$-\sum_{i\in I}(a_i+\hat{u}_i)\log(\bar{x}_i/\hat{x}_i)-\sum_{j\in J}(a_j+\hat{u}_j)\log(\bar{y}_j/\hat{y}_j) \geq (\hat{x}_I-\bar{x}_I)^T(b_I+\hat{v}_I).$$

i.e.,

$$\begin{aligned}&\sum_{i\in I}(a_i+\hat{u}_i)\log\bar{x}_i+\sum_{j\in J}(a_j+\hat{u}_j)\log\bar{y}_j-\bar{x}_I^T(b_I+\hat{v}_I)\\ &\leq \sum_{i\in I}(a_i+\hat{u}_i)\log\hat{x}_i+\sum_{j\in J}(a_j+\hat{u}_j)\log\hat{y}_j-\hat{x}_I^T(b_I+\hat{v}_I).\end{aligned} \tag{38}$$

Notice that above inequality holds for any solution $(\bar{x},\bar{y})$. Since the solution set S^* is convex and the function

$$\phi(x,y) = \sum_{i\in I}(a_i+\hat{u}_i)\log x_i+\sum_{j\in J}(a_j+\hat{u}_j)\log y_i - x_I^T(b_I+\hat{v}_I)$$

is a strict concave function, the inequality (38) implies that the accumulation point $(\hat{x},\hat{y})$ of the sequence $\{(x^k,y^k)\}$ is the optimal solution of the strict concave program $\max\{\phi(x,y) : (x,y) \in S^*\}$. Since a strict concave program has at most one solution, $(\hat{x},\hat{y})$ must be unique. The proof is complete. □

4. Local superlinear convergence

The global convergence of Algorithm 2.1 has been proved in the previous section. Now, we further show that the algorithm is also locally superlinearly convergent under some condition. We make use of the following assumption.

Condition 4.1. *(i) The NCP has a unique solution x^*, and (ii) $\nabla f(x)$ is Lipschitz continuous in the neighborhood of x^*.*

The following result is an immediate consequence of Theorem 3.4.

Corollary 4.1. *Let f be a continuously differentiable P_*-mapping. Let $\{(x^k, y^k)\}$ be generated by Algorithm 2.1. Assume that Condition 3.1 and Condition 4.1 are satisfied. Then the sequence $\{(x^k, y^k)\}$ converges to (x^*, y^*), the unique solution of the problem.*

The following result shows that the stepsize $\hat{\lambda}_k$ in Step 3 converges to 1.

Lemma 4.1. *Under the same conditions in Corollary 4.1, if $\hat{\lambda}_k$ is given as in Step 3 of Algorithm 2.1, then* $\lim_{k\to\infty} \hat{\lambda}_k = 1$.

Proof. According to Theorem 3.2 and Corollary 4.1, the sequence $\{(x^k, y^k, \theta_k)\}$ converges to $(x^*, y^*, 0)$. Thus, by continuity, we have $\mathcal{H}(x^k, y^k, 0) \to \mathcal{H}(x^*, y^*, 0) = 0$. By (7) and (23), we have

$$\|(\Delta\hat{x}^k, \Delta\hat{y}^k)\| \le \|[\nabla\mathcal{H}(x^k, y^k, \theta_k)]^{-1}\|\|\mathcal{H}(x^k, y^k, 0)\| \to 0.$$

Since $\{(x^k, y^k, \theta_k)\} \to (x^*, y^*, 0)$, Condition 3.1 implies that $\nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)$ is nonsingular. From (24), this indicates that $x^* + y^* > 0$. Thus, (x^*, y^*) is a strictly complementary solution. Let $I = \{i : x_i^* > 0\}$ and $J = \{j : y_j^* > 0\}$. Then by strict complementarity, we have $I \cup J = \{1, ..., n\}$. Since $x_I^* > 0, y_J^* > 0$ and $(\Delta\hat{x}^k, \Delta\hat{y}^k) \to 0$, there exists a k' such that for all $k \ge k'$ we have

$$x_I^k + \lambda\Delta\hat{x}_I^k > 0, \;\; y_J^k + \lambda\Delta\hat{y}_J^k > 0 \text{ for every } \lambda \in (0, 1] \tag{39}$$

On the other hand, by (7), we have

$$Y^k\Delta\hat{x}^k + X^k\Delta\hat{y}^k = -X^k y^k.$$

Multiplying both sides by $(X^kY^k)^{-1}$, we have

$$(X^k)^{-1}\Delta\hat{x}^k + (Y^k)^{-1}\Delta\hat{y}^k = -e.$$

Therefore,

$$(X_J^k)^{-1}\Delta\hat{x}_J^k = -e_J - (Y_J^k)^{-1}\Delta\hat{y}_J^k, \;\; (Y_I^k)^{-1}\Delta\hat{y}_I^k = -e_I - (X_I^k)^{-1}\Delta\hat{x}_I^k.$$

Let $\lambda \in (0, 1)$ be an arbitrary scalar. We thus have

$$\begin{aligned} x_J^k + \lambda\Delta\hat{x}_J^k &= X_J^k(X_J^k)^{-1}(x_J^k + \lambda\Delta\hat{x}_J^k) = X_J^k(e_J + \lambda(X_J^k)^{-1}\Delta\hat{x}_J^k) \\ &= X_J^k[(1-\lambda)e_J - \lambda(Y_J^k)^{-1}\Delta\hat{y}_J^k]. \end{aligned} \tag{40}$$

Similarly,

$$y_I^k + \lambda\Delta\hat{y}_I^k = Y_I^k[(1-\lambda)e_I - \lambda(X_I^k)^{-1}\Delta\hat{x}_I^k]. \tag{41}$$

If $(1-\theta_k)t_k \ge 1$ where t_k is given as in Step 2 of Algorithm 2.1, by the definition of $\hat{\lambda}_k$ we have that $\hat{\lambda}_k = 1$. We now consider the case of $(1-\theta_k)t_k < 1$. Since $(\Delta\hat{x}^k, \Delta\hat{y}^k) \to 0$ and $(x_I^k, y_J^k) \to (x_I^*, y_J^*) > 0$, we have that $(X_I^k)^{-1}\Delta\hat{x}_I^k \to 0$ and $(Y_J^k)^{-1}\Delta\hat{y}_J^k \to 0$. It follows that for any given $\alpha \in (0, 1)$, there exists a k'' such that for all $k > k''$ we have

$$(1-\lambda)e_I - \lambda(X_I^k)^{-1}\Delta\hat{x}_I^k > 0, \;\; (1-\lambda)e_J - \lambda(Y_J^k)^{-1}\Delta\hat{y}_J^k > 0,$$

for all $\lambda \in (0, \alpha]$. Therefore, for every $k > k''$, it follows from (40) and (41) that

$$x_J^k + \lambda\Delta\hat{x}_J^k > 0, \;\; y_I^k + \lambda\Delta\hat{y}_I^k > 0 \text{ for all } \lambda \in (0, \alpha]. \tag{42}$$

Combining (39) and (42), we have $t_k \ge \alpha$. Therefore, $\hat{\lambda}_k = \min\{1, (1-\theta_k)t_k\} = (1-\theta_k)t_k \ge (1-\theta_k)\alpha$. In the summary, for all sufficiently large k, we have either $\hat{\lambda}_k = 1$ or $1 > \hat{\lambda}_k \ge (1-\theta_k)\alpha$. Since $\theta_k \to 0$ and α is an arbitrary scalar less than 1, we conclude that $\hat{\lambda}_k \to 1$ as $k \to \infty$. □

We are now ready to prove the superlinear convergence when $p \geq 1$. The superlinear convergence when $p < 1$ is not known.

Theorem 4.1. *Assume that f is a continuously differentiable P_*-mapping satisfying Condition 4.1. Let p and β be chosen such that $p \geq 1$ and $\beta > 1 + w^*$ where w^* is a constant given by*

$$w^* = \|a\|_\infty + \|b\|_\infty + \|f(x^*)\|_\infty + \|x^*\|_\infty. \tag{43}$$

Let the sequence (x^k, y^k) be generated by Algorithm 2.1. If Condition 3.1 is satisfied, then the sequence (x^k, y^k) is locally superlinearly convergent, i.e., $\lim_{k\to\infty} \|x^{k+1} - x^*\|/\|x^k - x^*\| = 0$.

Proof. By Corollary 4.1, we have that $\{(x^k, y^k, \theta_k)\} \to (x^*, y^*, 0)$. This together with $p \geq 1$ implies that there exists a constant $c > 0$ such that $\|\nabla f(x^k) - (1-\theta_k)(\theta_k)^{p-1} I\| \leq c$ for all k. Therefore

$$\begin{aligned}
&\|\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k) - \nabla_{(x,y)}\mathcal{H}(x^k, y^k, 0)\| \\
&\leq \left\| \begin{pmatrix} Y^k & X^k \\ -(1-\theta_k)(\nabla f(x^k) + (\theta_k)^p E) & E \end{pmatrix} - \begin{pmatrix} Y^k & X^k \\ -\nabla f(x^k) & E \end{pmatrix} \right\| \\
&\leq \theta_k \|\nabla f(x^k) - (1-\theta_k)(\theta_k)^{p-1} E\| \\
&\leq c\theta_k.
\end{aligned}$$

By Lipschitz continuity of $\nabla f(x)$, for all sufficiently large k we have

$$\begin{aligned}
\|\nabla_{(x,y)}\mathcal{H}(x^k, y^k, 0) - \nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)\| &\leq \left\| \begin{pmatrix} Y^k & X^k \\ -\nabla f(x^k) & E \end{pmatrix} - \begin{pmatrix} Y^* & X^* \\ -\nabla f(x^*) & E \end{pmatrix} \right\| \\
&\leq \|Y^k - Y^*\| + \|X^k - X^*\| + \|\nabla f(x^k) - \nabla f(x^*)\| \\
&\leq c'\|y^k - y^*\| + c''\|x^k - x^*\|,
\end{aligned}$$

where c' and c'' are two positive constants. Thus, for all sufficiently large k we have

$$\begin{aligned}
&\|\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k) - \nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)\| \\
&\leq \|\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k) - \nabla_{(x,y)}\mathcal{H}(x^k, y^k, 0)\| + \|\nabla_{(x,y)}\mathcal{H}(x^k, y^k, 0) - \nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)\| \\
&\leq c\theta_k + c'\|y^k - y^*\| + c''\|x^k - x^*\|.
\end{aligned} \tag{44}$$

On the other hand, denote

$$\tau_k = \frac{\|\mathcal{H}(x^k, y^k, 0) - \mathcal{H}(x^*, y^*, 0) + \nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)((x^k, y^k) - (x^*, y^*))\|}{\|(x^k, y^k) - (x^*, y^*)\|}. \tag{45}$$

By continuous differentiability of $\mathcal{H}$, it follows that $\lim_{k\to\infty} \tau_k = 0$. Therefore, by (23), (44) and (45), for all sufficiently large k we have that

$$\begin{aligned}
&\|(\hat{x}^{k+1}, \hat{y}^{k+1}) - (x^*, y^*)\| \\
= \;& \|(x^k, y^k) + \hat{\lambda}_k(\Delta\hat{x}^k, \Delta\hat{y}^k) - (x^*, y^*)\| \\
= \;& \|(x^k, y^k) - (x^*, y^*) + \hat{\lambda}_k[\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k)]^{-1}\mathcal{H}(x^k, y^k, 0)\| \\
= \;& \|[\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k)]^{-1}\{[\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k) - \nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)]((x^k, y^k) \\
& -(x^*, y^*)) + \hat{\lambda}_k[\mathcal{H}(x^k, y^k, 0) - \mathcal{H}(x^*, y^*, 0) + \nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)((x^k, y^k) \\
& -(x^*, y^*))] + (1-\hat{\lambda}_k)\nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)((x^k, y^k) - (x^*, y^*))\}\| \\
\leq \;& \|[\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k)]^{-1}\|\{\|[\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k) - \nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)]((x^k, y^k) \\
& -(x^*, y^*))\| + \hat{\lambda}_k\|\mathcal{H}(x^k, y^k, 0) - \mathcal{H}(x^*, y^*, 0) + \nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)((x^k, y^k)
\end{aligned}$$

$$
\begin{aligned}
& -(x^*, y^*))\| + (1-\hat{\lambda}_k)\|\nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)((x^k, y^k) - (x^*, y^*))\|\} \\
\leq \; & C(c\theta_k + c'\|y^k - y^*\| + c''\|x^k - x^*\|)\|(x^k, y^k) - (x^*, y^*)\| + \hat{\lambda}_k C\tau_k\|(x^k, y^k) - (x^*, y^*)\| \\
& + C(1-\hat{\lambda}_k)\|\nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)\|\|(x^k, y^k) - (x^*, y^*)\| \\
= \; & \phi_k\|(x^k, y^k) - (x^*, y^*)\|,
\end{aligned} \tag{46}
$$

where

$$
\phi_k = C(c\theta_k + c'\|y^k - y^*\| + c''\|x^k - x^*\|) + \hat{\lambda}_k C\tau_k + C(1-\hat{\lambda}_k)\|\nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)\|.
$$

Since $(x^k, y^k) \to (x^*, y^*), \theta_k \to 0$, $\tau_k \to 0$, and $\hat{\lambda}_k \to 1$ (by Lemma 4.1), it follows that $\phi_k \to 0$ as $k \to \infty$. Therefore, to show the local superlinear convergence of the algorithm, it is sufficient to prove that the algorithm eventually takes only Step 2 and Step 3, that is, to show that there exists a k' such that for all $k > k'$ the iterate (x^k, y^k) is produced by Step 3. Since $(x^k, y^k, \theta_k) \to (x^*, y^*, 0)$ and $\nabla_{(x,y)}\mathcal{H}(x^*, y^*, 0)$ is nonsingular, we have

$$
\lim_{k\to\infty} \|E - \nabla_{(x,y)}\mathcal{H}(x^k, y^k, 0)[\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k)]^{-1}\| = 0,
$$

and

$$
\|(\Delta\hat{x}^k, \Delta\hat{y}^k)\| = \|[\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k)]^{-1}\mathcal{H}(x^k, y^k, 0)\| = O(\|\mathcal{H}(x^k, y^k, 0)\|),
$$

which also implies that $(\Delta\hat{x}^k, \Delta\hat{y}^k) \to 0$, and hence $(\hat{x}^{k+1}, \hat{y}^{k+1}) \to (x^*, y^*)$. Notice that $\hat{\lambda}_k \to 1$ (by Lemma 4.1). Combining the above two relations yields

$$
\begin{aligned}
& \|\mathcal{H}(\hat{x}^{k+1}, \hat{y}^{k+1}, 0)\| \\
& = \|\mathcal{H}(x^k, y^k, 0) + \hat{\lambda}_k\nabla_{(x,y)}\mathcal{H}(x^k, y^k, 0)(\Delta\hat{x}^k, \Delta\hat{y}^k)\| + o(\hat{\lambda}_k\|(\Delta\hat{x}^k, \Delta\hat{y}^k)\|) \\
& = \|[I - \hat{\lambda}_k\nabla_{(x,y)}\mathcal{H}(x^k, y^k, 0)\nabla_{(x,y)}\mathcal{H}(x^k, y^k, \theta_k)^{-1}]\mathcal{H}(x^k, y^k, 0)\| + o(\hat{\lambda}_k\|(\Delta\hat{x}^k, \Delta\hat{y}^k)\|) \\
& = o(\|\mathcal{H}(x^k, y^k, 0)\|) + o(\hat{\lambda}_k\|(\Delta\hat{x}^k, \Delta\hat{y}^k)\|) \\
& = o(\|\mathcal{H}(x^k, y^k, 0)\|).
\end{aligned} \tag{47}
$$

Therefore, for any given $\tau < \eta/(\beta + w^*)$, where w^* is given by (43) and η is the constant given as in Algorithm 2.1, there exists a k_0 such that for all $k > k_0$ we have

$$
\|\mathcal{H}(\hat{x}^{k+1}, \hat{y}^{k+1}, 0)\|_\infty \leq \tau\|\mathcal{H}(x^k, y^k, 0)\|_\infty. \tag{48}
$$

Denote

$$
\begin{aligned}
w^k &= \|a\|_\infty + \|b\|_\infty + \|f(x^k)\|_\infty + \|x^k\|_\infty, \\
\hat{w}^{k+1} &= \|a\|_\infty + \|b\|_\infty + \|f(\hat{x}^{k+1})\|_\infty + \|\hat{x}^{k+1}\|_\infty.
\end{aligned}
$$

Clearly, $w^k \to w^*$ and $\hat{w}^{k+1} \to w^*$, there is an index $\bar{k}_0$ such that for all $k > \bar{k}_0$,

$$
\tau < \eta/(\beta + w^k), \; 1 + \hat{w}^{k+1} < \beta. \tag{49}
$$

If for all $k > \max\{k_0, \bar{k}_0\}$ the iterate (x^k, y^k) is generated by Step 3, from (46) we have the desired superlinear convergence result. Otherwise, there exists an $l > \max\{k_0, \bar{k}_0\}$ such that the iterate (x^l, y^l) is produced by Step 5. We prove that after this step all succeeding iterates, i.e., $\{(x^k, y^k)\}$ where $k > l$, are generated by Step 3. Indeed, since $p \geq 1$ and $\theta_l \leq 1$, it is easy to see that

$$
\|\mathcal{H}(x^l, y^l, 0) - \mathcal{H}(x^l, y^l, \theta_l)\|_\infty \leq \theta_l w^l,
$$

$$
\|\mathcal{H}(\hat{x}^{l+1}, \hat{y}^{l+1}, \hat{\theta}_{l+1}) - \mathcal{H}(\hat{x}^{l+1}, \hat{y}^{l+1}, 0)\|_\infty \leq \hat{\theta}_{l+1}\hat{\omega}^{l+1}.
$$

Notice that $\|\mathcal{H}(x^l, y^l, \theta_l)\|_\infty \le \beta\theta_l$. It follows that

$$\|\mathcal{H}(x^l, y^l, 0)\|_\infty \le \|\mathcal{H}(x^l, y^l, 0) - \mathcal{H}(x^l, y^l, \theta_l)\|_\infty + \|\mathcal{H}(x^l, y^l, \theta_l)\|_\infty \le (\omega^l + \beta)\theta_l.$$

Since $l > \max\{k_0, \bar{k}_0\}$, by (48) and (49), we have

$$\hat{\theta}_{l+1} = \|\mathcal{H}(\hat{x}^{l+1}, \hat{y}^{l+1}, 0)\|_\infty \le \tau\|\mathcal{H}(x^l, y^l, 0)\|_\infty \le \tau(\omega^l + \beta)\theta_l < \eta\theta_l.$$

On the other hand, we have

$$\begin{aligned} \|\mathcal{H}(\hat{x}^{l+1}, \hat{y}^{l+1}, \hat{\theta}_{l+1})\|_\infty &\le \|\mathcal{H}(\hat{x}^{l+1}, \hat{y}^{l+1}, \hat{\theta}_{l+1}) - \mathcal{H}(\hat{x}^{l+1}, \hat{y}^{l+1}, 0)\|_\infty + \|\mathcal{H}(\hat{x}^{l+1}, \hat{y}^{l+1}, 0)\|_\infty \\ &\le \hat{\theta}_{l+1}\hat{\omega}^{l+1} + \hat{\theta}_{l+1} \\ &< \beta\hat{\theta}_{l+1}. \end{aligned} \tag{50}$$

The last inequality follows from (49). Therefore, by the construction of the algorithm, the iterate (x^{l+1}, y^{l+1}) is generated by Step 3. Thus,

$$(x^{l+1}, y^{l+1}) = (\hat{x}^{l+1}, \hat{y}^{l+1}),\ \theta_{l+1} = \hat{\theta}_{l+1} = \|\mathcal{H}(\hat{x}^{l+1}, \hat{y}^{l+1}, 0)\|_\infty.$$

We now show that the next iterate (x^{l+2}, y^{l+2}) is also generated by Step 3. Indeed, by (48) and noting that $\tau < \eta$, we have

$$\hat{\theta}_{l+2} = \|\mathcal{H}(\hat{x}^{l+2}, \hat{y}^{l+2}, 0)\|_\infty \le \tau\|\mathcal{H}(x^{l+1}, y^{l+1}, 0)\|_\infty = \tau\theta_{l+1} < \eta\theta_{l+1}.$$

By a proof similar to (50), we still have $\|\mathcal{H}(\hat{x}^{l+2}, \hat{y}^{l+2}, \hat{\theta}_{l+2})\|_\infty < \beta\hat{\theta}_{l+2}$. Thus, (x^{l+2}, y^{l+2}) is also generated by Step 3. Repeating the proof, we conclude that Algorithm 2.1 eventually takes only Steps 2 and 3. Since $\lim_{k\to\infty} \phi_k = 0$, it follows from (46) that $\lim_{k\to\infty} \|x^{k+1} - x^*\|/\|x^k - x^*\| = 0$. □

5. Preliminary numerical experiments

Algorithm 2.1 was implemented and tested on some linear and nonlinear complementarity problems that are not necessarily P_* problems. We first introduce these test problems.

(P1) *Walrasian equilibrium model* (WEM) [22]. This is a 4-variable problem depending on three parameters (α, b_2, b_3). We use two sets of constants, i.e., $(\alpha, b_2, b_3) = (0.75, 1, 0.5)$ and $(0.75, 1, 2)$, and denote by $\mathrm{WEM}_{0.5}$ and WEM_2 the two cases, respectively.

(P2) *Nash-cournot production problem* (NCPP)[22]. We denote by NCPP_5 and NCPP_{10} the 5 and 10-variable problems, respectively.

(P3) *Invariant capital stock model* (ICSM) [13]. This is an NCP formulated from an invariant capital stock model described in [13]. We solve this example with $(n, m, l) = (10, 2, 2)$, $q = (0.8, 0.8)^T$, and $v(u) = (u_1 + 2.5u_2)^{0.2}(2.5u_3 + u_4)^{0.2}(2u_5 + 3u_6)^{0.2}$. The data for A, B and C can be found in [13]. We run our algorithm for two different values of the discount factor, i.e., $\alpha = 0.7$ and 0.9. We denote the two cases by $\mathrm{ICSM}_{0.7}$ and $\mathrm{ICSM}_{0.9}$, respectively.

(P4) *Waston's fourth problems* (WFP) [32]. This is an NCP representing the KKT conditions for a convex programming problem.

(P5) *Kojima-Shindo problems* (KSP) [22].

(P6) *The generalized von Thünen model* (GTM) [22]. This is an NCP with 106 variables, which pertains to an agriculture economy with 20 farming regions, 4 commodities and 3 consumers.

(P7) *Fathi's example* [12]. This is an LCP with the vector $q = -e$, and the positive definite matrix M_1 given in (51).

$$M_1 = \begin{pmatrix} 1 & 2 & 2 & \dots & 2 \\ 2 & 5 & 6 & \dots & 6 \\ 2 & 6 & 9 & \dots & 10 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 2 & 6 & 9 & \dots & 4(n-1)+1 \end{pmatrix},\ M_2 = \begin{pmatrix} 1 & 2 & 2 & \dots & 2 \\ 0 & 1 & 2 & \dots & 2 \\ 0 & 0 & 1 & \dots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}. \tag{51}$$

(P8) *Murty's example* [21]. This is an LCP with $q = -e$ and the matrix M_2 given in (51).

(P9) *Ahn's example* [1]. The matrix M_3 is given in (52) and the vector $q = -e$.

(P10) *Waston's third problem* (WTP) [32]. This is a 10-variable LCP. The matrix is given in [32], and q is the vector with -1 in the 8th coordinate and zeros elsewhere.

(P11) *P_* problem.* The matrix M_4 is a P_*-matrix given in (52), and $q = (1, -2, 0, 0)^T$. This LCP has no strictly feasible point and the solution set of the problem is unbounded. Clearly, this problem is not a monotone LCP.

$$M_3 = \begin{pmatrix} 4 & -2 & 0 & 0 & 0 & ... & 0 \\ 1 & 4 & -2 & 0 & 0 & ... & 0 \\ 0 & 1 & 4 & -2 & 0 & ... & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & ... & ... & ... & 0 & 1 & 4 \end{pmatrix}, \ M_4 = \begin{pmatrix} 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \\ -2 & -1 & 0 & 0 \\ 4 & 8 & 0 & 0 \end{pmatrix}. \tag{52}$$

Table 1: Numerical results for $p = 0.9$

Problem	DIM	R^0	IT5	IT3	NF	CPU (sec.)	R^k	θ_k
$WEM_{0.5}$	4	1.5	6	3	24	0.00	2.2e-16	8.7e-09
WEM_2	4	2.25	5	5	29	0.00	2.2e-16	6.4e-12
$NCPP_5$	5	428.162	7	4	26	0.00	7.1e-16	2.1e-14
$NCPP_{10}$	10	253.713	8	5	30	0.00	9.8e-16	1.1e-14
$ICSP_{0.7}$	14	10.2	31	1	95	0.00	8.9e-16	1.7e-17
$ICSP_{0.9}$	14	10.2	12	3	44	0.00	2.2e-16	3.1e-12
WFP	5	88106.8	31	1	95	0.00	3.8e-16	2.2e-19
WTP	10	4	30	4	1750	0.01	3.4e-22	6.7e-20
KSP	4	13	9	3	31	0.00	8.9e-16	4.5e-09
GTM ($\varepsilon = 10^{-5}$)	106	52.475	64	0	1068	1.22	6.2e-06	5.8e-08
P_*	4	11	4	3	24	0.00	0.0	0.0
Murty	100	197	7	3	25	0.13	6.6e-16	3.7e-11
Murty	200	397	8	3	28	1.60	4.4e-16	9.4e-12
Murty	400	797	9	4	32	21.94	0.00	1.1e-15
Ahn	100	3	5	3	19	0.12	2.2e-16	7.8e-12
Ahn	200	3	5	3	19	1.05	2.2e-16	7.8e-12
Ahn	400	3	5	3	19	13.04	2.2e-16	7.8e-12
Fathi	100	19997	11	3	37	0.22	0.0	1.3e-15
Fathi	200	79997	13	2	42	1.98	3.0e-22	3.8e-16
Fathi	400	319997	14	3	46	28.01	2.3e-16	4.2e-15

Algorithm 2.1 is coded in Fortran 90 and run on a DEL Alpha workstation. For a given tolerance ε, the following aspects are recorded to examine the numerical effectiveness of the algorithm: Dimension of the problem (DIM), the executing time (in second) of the algorithm (CPU), the total number of iterates generated by Step 5 (IT5), the total number of iterates generated by Step 3 (IT3), and the number of evaluations of functions (NF). R^0 stands for the initial residual $\|\mathcal{H}(x^0, y^0, 0)\|_\infty$, and $R^k = \|\mathcal{H}(x^k, y^k, 0)\|_\infty$ denotes the final residual if the algorithm terminates at x^k. The final θ_k is also recorded. The total number of iterations is omitted here since it is the summation of IT5 and IT3.

In general, a wide neighborhood enables the algorithm to take a large stepsize at each iteration to ensure a rapid convergence. It is evident that in Algorithm 2.1 the width of the neighborhood around the regularized central path is determined by the value of β. Therefore, a large β should be taken in practical applications. This is also inspired by the analysis in Section 4 (see Theorem 4.1). In our experiments, we use Strategy 2.1 to obtain the initial points and parameters. This strategy permits the algorithm to start from arbitrary vectors $(x^0, y^0) > 0$. The following initial points and parameters are used in our code for all test problems: $a = b = e$, $x^0 = y^0 = e$, $\theta_0 = 0.9$, $\sigma = 0.001$, $\alpha_1 = \alpha_2 = 0.9$, $\eta = 0.99$, $p = 0.9$, and $\beta = \|\mathcal{H}(x^0, y^0, \theta_0)\|_\infty/\theta_0 + 100$. The termination criterion is $\|\mathcal{H}(x, y, 0)\|_\infty < \varepsilon$, where ε is a given tolerance. In our code, $\varepsilon = 1e-15$ is used for all test problems except for the generalized von Thünen model. From our numerical experience, when p is relatively large, e.g., $p \geq 0.8$, the algorithm converges quickly and the performance of algorithm has no remarkable difference except for von Thünen model. Numerical results for $p = 0.9$ are summarized in Table 1, and for $p = 1.8$ are summarized in Table 2. We note that when $p = 0.9$ is taken, it is difficult for the algorithm to achieve an accuracy higher than $1e-7$ for the generalized von Thünen model. The result with $\varepsilon = 1e-5$ is given in Table 1. However, if $p = 1.8$ is taken, the algorithm can reach an approximate solution with $\varepsilon = 1e-14$.

Table 2: Numerical results for $p = 1.8$

Problem	DIM	R^0	IT5	IT3	NF	CPU (sec.)	R^k	θ_k
$WEM_{0.5}$	4	1.5	6	3	22	0.00	2.1e-16	1.1e-13
WEM_2	4	2.25	4	4	17	0.00	2.1e-17	1.3e-13
$NCPP_5$	5	428.162	6	3	22	0.00	6.9e-16	2.1e-09
$NCPP_{10}$	10	253.713	9	4	32	0.00	9.6e-16	1.2e-15
$ICSP_{0.7}$	14	10.2	34	1	104	0.00	5.6e-16	1.7e-17
$ICSP_{0.9}$	14	10.2	12	3	43	0.00	2.1e-16	1.9e-14
WFP	5	88106.8	31	1	95	0.00	3.8e-16	2.2e-19
WTP	10	4	28	3	1702	0.01	4.5e-23	5.2e-21
KSP	4	13	11	4	93	0.00	1.4e-23	7.5e-13
GTM ($\varepsilon = 10^{-14}$)	106	52.475	57	0	254	1.04	7.1e-15	5.6e-17
P_*	4	11	4	3	24	0.00	0.0	0.0
Murty	100	197	7	3	25	0.16	4.4e-16	3.9e-11
Murty	200	397	8	3	28	1.45	6.6e-16	9.6e-12
Murty	400	797	9	3	31	30.38	8.8e-16	1.9e-12
Ahn	100	3	5	3	19	0.12	2.2e-16	9.0e-12
Ahn	200	3	5	3	19	1.05	2.2e-16	9.1e-12
Ahn	400	3	5	3	19	14.08	2.2e-16	9.1e-12
Fathi	100	19997	12	2	39	0.22	2.1e-23	2.9e-16
Fathi	200	79997	14	2	51	2.12	2.3e-24	3.3e-17
Fathi	400	319997	14	3	46	43.56	2.2e-16	4.0e-15

It is interesting to see how the convergence speed of Algorithm 2.1 changes if the parameter p varies in $(0, \infty)$. As an example, we consider the problem (P11) which is a P_* problem. For different p, the performance of the algorithm is demonstrated in Table 3. We see that the algorithm becomes very slow when p is too small.

It is also of interest to see how the behavior of the algorithm changes as β varies. To ensure the initial condition of the algorithm, the least value of β is given by $\beta_0 = \|\mathcal{H}(x^0, y^0, \theta_0)\|_\infty/\theta_0$. Let $p = 0.9$ and $\varepsilon = 1e-15$. For different β, the numerical results on the P_* problem are given in Table 4. We see that the convergence speed is indeed improved as β increases, but the convergence speed does not improve any more after β becomes sufficiently large.

Table 3: Performance of the algorithm for different p

p	0.4	0.45	0.5	0.6	0.7	0.9	1.2	1.8	2
ITS	750	132	39	8	7	7	7	7	7
NF	2998	526	154	28	24	24	24	24	26
CPU	0.07	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4: Performance of the algorithm for different β

β	β_0	β_0+1	β_0 +5	β_0 +8	β_0 +10	$\geq \beta_0 + 20$
ITS	126	79	33	24	9	8
NF	502	314	126	90	30	26
CPU	0.14	0.08	0.03	0.02	0.01	0.00

Finally, we make some comments on the reduction of θ_k. To assure a fast convergence, the iterate should be eventually generated by Step 3 since the search direction generated in Step 2 is actually an approximate Newton direction. Thus, a large value of $\eta \in (0,1)$ should be chosen in order to generate the new iterates by Step 3. That is why $\eta = 0.99$ is taken in our code. But this does not imply a slow decrease of θ_k. In the proof in Theorem 4.1, we have proved that the algorithm gradually phases out the Steps 4, 5 and 6, and eventually repeats Steps 2 and 3. Thus, by the construction of Step 3, $\theta_k = \|\mathcal{H}(x^k, y^k, 0)\|_\infty$ for all sufficiently large k. Therefore, under conditions of Theorem 4.1, it follows from (47) that $\lim_{k\to\infty} \theta_{k+1}/\theta_k = 0$, i.e., the sequence θ_k converges to zero superlinearly. Fast decrease of θ_k indeed can be seen from our numerical results (see, Table 1 and Table 2).

6. Final remarks. In this paper, a new path-following algorithm for NCPs is presented, which is based on a new concept of the regularized central path. This method is globally convergent for any P_* NCP with a nonempty solution set. The boundedness assumption on the solution set, or equivalently the strict feasibility condition is not required in our algorithm. Under certain assumptions, a fast local convergence of the algorithm can also be achieved. Tikhonov regularization plays a role in our method, which enables the algorithm to tackle ill-posed or unstable complementarity problems. Moreover, since this algorithm tracks approximately the so-called regularized central path which is proved to exist under a mild condition for general P_0 problems (see [39]), the algorithm proposed in this paper can be extended to nonlinear P_0 problems without any difficulty.

Acknowledgments: We are grateful to the three anonymous referees and the associate editor for their constructive comments.

References

[1] B.H. Ahn, *Iterative methods for linear complementarity problem with uperbounds and lower-bounds,* Math. Programming, 26 (1983), pp. 265–315.

[2] E.D. Andersen and Y. Ye, *On a homogeneous algorithm for the monotone complementarity problem,* Math. Programming, 84 (1999), pp. 375-399.

[3] M. Anitescu, G. Lesaja and F.A. Potra, *An infeasible-interior-point predictor-corrector algorithm for the P_*-geometric LCP*, Appl. Math. Optim., 36 (1997), pp. 203–228.

[4] J.V. Burke and S. Xu, *The global linear convergence of a non-interior path-following algorithm for linear complementarity problems,* Math. Oper. Res., 23 (1998), pp. 719–734.

[5] J.V. Burke and S. Xu, *A non-interior predictor-corrector path following algorithm for the monotone linear complementarity problem,* Math. Programming, 87 (2000), pp. 113–130.

[6] B. Chen and X. Chen, *A global and local superlinear continuation method for P_0 and R_0 and monotone NCP,* SIAM J. Optim., 9 (1999), pp. 624–645.

[7] R.W. Cottle, J.S. Pang and R.E. Stone, *The Linear Complementarity Problem*, Academic Press, Boston, 1992.

[8] R.W. Cottle, J.S. Pang and V. Venkateswaran, *Sufficient matrices and the linear complementarity problem,* Linear Algebra Appl., 114/115 (1989), pp. 231–249.

[9] F. Facchinei, *Structural and stability properties of P_0 nonlinear complementarity problems,* Math. Oper. Res., 23 (1998), pp. 735–745.

[10] F. Facchinei and C. Kanzow, *Beyond monotonicity in regularization methods for nonlinear complementarity problems,* SIAM J. Control Optim., 37 (1999), pp. 1150–1161.

[11] F. Facchinei and J.S. Pang, *Finite-dimensional variational inequalities and complementarity problems,* Volumes 1 and 2, Springer-Verlag, New York, 2003.

[12] Y. Fathi, *Computational complexity of LCPs associated with positive definite matrices,* Math. Programming, 17 (1979), pp. 335– 344.

[13] T. Hansen and T.C. Koopmans, *On the definition and computation of a capital stock invariant under optimization,* J. Economic Theory, 5 (1972), 487–523.

[14] Z. Huang, J. Han and Z. Chen, *Predictor-corrector smoothing Newton method, based on a new smoothing function, for solving the nonlinear complementarity problem with a P_0 function,* J. Optim. Theory Appl., 117 (2003), 39-68.

[15] G. Isac, *Tikhonov's regularization and the complementarity problem in Hilbert space,* J. Math. Anal. Appl., 174 (1991), pp. 53–66.

[16] B. Jansen, C. Roos, T. Terlaky and A. Yoshise, *Polynomiality of primal dual affine scaling algorithm for nonlinear complementarity problems,* Math. Programming, 78 (1997), pp. 315–345.

[17] C. Jones and M.S. Gowda, *On the connectedness of solution sets in linear complementarity problems*, Linear Algebra Appl., 272 (1998), pp. 33-44.

[18] M. Kojima, N. Megiddo, and T. Noma, *Homotopy continuation methods for nonlinear complementarity problems,* Math. Oper. Res., 16 (1991), pp. 754–774.

[19] M. Kojima, N. Megiddo, T. Noma and A. Yoshise, *A unified approach to interior point algorithms for linear complementarity problems,* Lecture Notes in Computer Sciences, 538, Springer-Verlag, New York, 1991.

[20] M. Kojima, M. Mizuno, and T. Noma, *Limiting behavior of trajectories generated by a continuation method for monotone complementarity problems,* Math. Oper. Res., 43 (1990), pp. 662-675.

[21] K.G. Murty, *Linear Complementarity, Linear and Nonlinear Programming,* Sigma Ser. Appl. Math. 3, Heldermann-Verlag, Berlin, 1988.

[22] J.S. Pang and S.A. Gabriel, *NE/SQP: A robust algorithm for the nonlinear complementarity problem,* Math. Programming, 60 (1993), pp. 295–338.

[23] F.A. Potra and R. Sheng, *Predictor-corrector algorithm for solving $P_*(\kappa)$-matrix LCP from arbitrary positive starting points,* Math. Programming, 76 (1997), pp. 223–244.

[24] F.A. Potra and R. Sheng, *A large-step infeasible-interior-point method for the P_*matrix LCP,* SIAM J. Optim., 7 (1997), pp. 318-335.

[25] L. Qi, *Convergence analysis of some algorithms for solving nonsmooth equations*, Math. Oper. Res., 18 (1993), pp. 227–243.

[26] G. Ravindran and M.S. Gowda, *Regularization of P_0-functions in box variational inequality problems,* SIAM J. Optim., 11 (2000), pp. 748–760.

[27] J. Stoer, M. Wechs, *Infeasible-interior-point paths for sufficient linear complementarity problems and their analyticity,* Math. Programming, 83 (1998), pp. 407–423.

[28] J. Stoer, M. Wechs, and S. Mizuno, *High order infeasible-interior-point methods for solving sufficient linear complementarity problems.* Math. Oper. Res., 23 (1998), pp. 832–862.

[29] R. Sznajder and M.S. Gowda, *On the limiting behavior of the trajectory of regularized solutions of P_0 complementarity problems,* Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, M. Fukushima and L. Qi (Eds), Kluwer Academic Publishers, pp. 371-379, 1998.

[30] P. Tseng, *Error Bounds for Regularized Complementarity Problems,* Report, Department of Mathematics, University of Washington, Seattle, 1998.

[31] P. Tseng, *An infeasible path-following method for monotone complementarity problems,* SIAM J. Optim., 7 (1997), pp. 386–402.

[32] L.T. Watson, *Solving the nonlinear complementarity problem by a homotopy method,* SIAM J. Control Optim., 17 (1979), pp. 36–46.

[33] S. Wright and D. Ralph, *Superlinear infeasible-interior-point algorithm for monotone complementarity problems*, Math. Oper. Res., 21 (1996), pp. 815–838.

[34] S. Xu and J.V. Burke, *A polynomial time interior-point path-following algorithm for LCP based on Chen-Harker-Kanzow smoothing techniques,* Math. Programming, 86 (1999), pp. 91–103.

[35] Y. Ye, *On homogeneous and self-dual algorithms for LCP*, Math. Programming, 76 (1997), pp. 211-222.

[36] Y. Ye, *Interior Point Algorithms Theory and Analysis,* John Wiley and Sons, Chichester, UK, 1997.

[37] Y.B. Zhao and G. Isac, *Properties of a multi-valued mapping associated with some nonmonotone complementarity problems,* SIAM J. Control Optim., 39 (2000), pp. 571–593.

[38] Y.B. Zhao and D. Li, *Strict feasibility conditions of complementarity problems,* J. Optim. Theory Appl., 107 (2000), pp. 641–664.

[39] Y.B. Zhao and D. Li, *On a new homotopy continuation trajectory for complementarity problems,* Math. Oper. Res., 26 (2001), pp. 119–146.

[40] Y.B. Zhao and D. Li, *Existence and limiting behavior of a non-interior-point trajectory for complementarity problems without strictly feasible condition,* SIAM J. Control Optim., 40 (2001), pp. 898-924.