Journal of Global Optimization (2005) 33: 477–510
DOI 10.1007/s10898-004-2708-5

A Superlinearly Convergent SSLE Algorithm for Optimization Problems with Linear Complementarity Constraints

JIAN-LING LI[1,2] and JIN-BAO JIAN[1]
[1]*College of Mathematics and Informatics Science, Guangxi University, 530004, Nanning, P.R. China*
[2]*Department of Mathematics, Shanghai University, Baoshan, Shanghai 200436, P.R. China (e-mail: {ljl123, jianjb}@gxu.edu.cn)*

Abstract. In this paper we study a special kind of optimization problems with linear complementarity constraints. First, by a generalized complementarity function and perturbed technique, the discussed problem is transformed into a family of general nonlinear optimization problems containing parameters. And then, using a special penalty function as a merit function, we establish a sequential systems of linear equations (SSLE) algorithm. Three systems of equations solved at each iteration have the same coefficients. Under some suitable conditions, the algorithm is proved to possess not only global convergence, but also strong and superlinear convergence. At the end of the paper, some preliminary numerical experiments are reported.

Key words: algorithm, complementarity constraints, global convergence, sequential systems of linear equations, superlinear convergence

1. Introduction

Optimization problems with complementarity constraints have wide applications in economy, engineering design, game theory and so on, so many scholars are interested in studying on this kind of problems and make great achievements (see [2, 4, 12, 13, 15]).

In this paper, we discuss a special kind of optimization problem in which the constraints are defined by a linear complementarity problem (LCP) described as follows:

$$\text{(LCP)}\quad \begin{array}{ll} \min & f(x,y) \\ \text{s.t.} & Ax\leqslant b, \\ & w = Nx + My + q, \\ & 0\leqslant w \perp y \geqslant 0, \end{array} \tag{1.1}$$

where $f\colon \Re^{n+m} \to \Re$ is continuously differentiable function, $A \in \Re^{p\times n}$, $N \in \Re^{m\times n}$, $M \in \Re^{m\times m}$, $b \in \Re^{p}$, $q \in \Re^{m}$.

Obviously, if one writes the complementarity condition $w\perp y$ as an inner product $w^{\mathrm{T}}y = 0$, then (LCP)(1.1) is equivalent to a standard smoothing nonlinear programming (SSNP). So, theoretically speaking, (LCP)(1.1) should be studied and solved with existing theory and methods for SSNP. Unfortunately, J.V. Outrata, M. Kocvara et al. pointed out in [16] that the weaker Mangasarian–Fromotitz constraint qualification did not hold at any feasible points of the equivalent SSNP. Therefore, it will be rather difficult to obtain the solution of (LCP)(1.1) by means of solving the equivalent SSNP directly.

As we know, the sequential quadratic programming (SQP) type algorithms and sequential system of linear equations (SSLE) type algorithms are effective methods for optimization problems with nonlinear constraints, many scholars have made a lot of research on them and made great achievements [3, 5–8, 17]. Recently, Fukushima, Luo and Pang proposed in [4] an SQP algorithm to (LCP)(1.1). The main idea of their algorithm is as follows: first, by Fischer–Burmeister complementarity function, (LCP)(1.1) is equivalently transformed into a general optimization problem (GP). And then solve the GP by means of SQP method. The initial point is demanded to satisfy the constraints $Ax \leqslant b, w = Nx + My + q$, and at each iteration, a quadratic programming subproblem need to be solved. So the computational amount is slight large. In addition, their algorithm possesses only global convergence.

In this paper, first by perturbed technique and a generalized complementarity function, we equivalently transform (LCP)(1.1) into a family of gereral optimization problems containing parameters. Then, motivated by the ideas from [6], we propose an SSLE algorithm to (LCP)(1.1). The proposed algorithm possesses a few important properties as follows: the initial point is only demanded to satisfy the constraint $w = Nx + My + q$; the algorithm uses a special penalty function as a merit function; the three systems of equations solved at each iteration have the same coefficients, so the computational amount is less than that of SQP algorithms. Under suitable conditions, the algorithm is proved to possess not only global convergence, but also strong and superlinear convergence.

The paper contains 7 sections. In Section 2, some known results are restated and the idea or formation of the algorithm is analysed. In Section 3, the algorithm is given and its feasibility is discussed. We prove respectively the global, strong convergence and superlinear convergence in Sections 4 and 5. Finally, some numerical experiments are presented in Section 6 and we conclude with some final remarks in Section 7.

2. Preliminaries

For convenience, we use the following notation throughout this paper:

$$X_0 = \{ z = (x, y, w) : w = Nx + My + q \},$$
$$X = \{z \in X_0 : Ax \leqslant b,\ 0 \leqslant y \perp w \geqslant 0\},$$

$$A^{\mathrm{T}} = (a_1^{\mathrm{T}}, \ldots, a_p^{\mathrm{T}}),\ b^{\mathrm{T}} = (b_1, \ldots, b_p), z = (x, y, w), s = (x, y),$$
$$t = (y, w), t_i = (y_i, w_i)$$
$$\mathrm{d}z = (\mathrm{d}x, \mathrm{d}y, \mathrm{d}w),\ \mathrm{d}s = (\mathrm{d}x, \mathrm{d}y),\ \mathrm{d}t = (\mathrm{d}y, \mathrm{d}w),\ \mathrm{d}t_i = (\mathrm{d}y_i, \mathrm{d}w_i),$$

$$L_1 = \{1, \ldots, p\},\ L_2 = \{1, \ldots, m\},\ \varphi(x) = \max\{0;\ a_j x - b_j, j \in L_1\}. \tag{2.1}$$

And denote directly $(x, y, w) = (x^{\mathrm{T}}, y^{\mathrm{T}}, w^{\mathrm{T}})^{\mathrm{T}}$.

Throughout this paper, we suppose that the following assumption holds:

(A1) (i) M is a P_0 matrix, that is, all principal minors of M are nonnegative;

(ii) For any $z \in X_0$, the vectors $\{a_j : j \in I(x)\}$ are linearly independent, where $I(x) = \{j \in L_1 : \varphi(x) = a_j x - b_j\}$.

Now we restate the definition of a stationary point of (LCP)(1.1) and a known result as follows.

DEFINITION 2.1. A feasible point $z^* = (x^*, y^*, w^*) \in X$ is said to be a stationary point of (LCP)(1.1) if

$$\mathrm{d}z = (\mathrm{d}s, \mathrm{d}w) = (\mathrm{d}x, \mathrm{d}y, \mathrm{d}w) \in T(z^*, X) \Longrightarrow \nabla f(x^*, y^*)^{\mathrm{T}} \mathrm{d}s \geqslant 0,$$

where $T(z^*, X)$ means the tangent cone of X at point z^*.

PROPOSITION 2.2 [15]. *Suppose that $z^* \in X$ satisfies the so-called nondegeneracy condition*:

$$(y_i^*, w_i^*) \neq (0, 0), i = 1, \ldots, m. \tag{2.2}$$

Then z^* is a stationary point of (1.1) *if and only if there exist multipliers $(\lambda^*, u^*, v^*) \in \Re^p \times \Re^m \times \Re^m$ such that*

$$\begin{pmatrix} \nabla f(x^*, y^*) \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ W^* \\ Y^* \end{pmatrix} v^* + \begin{pmatrix} N^{\mathrm{T}} \\ M^{\mathrm{T}} \\ -I \end{pmatrix} u^* + \begin{pmatrix} A^{\mathrm{T}} \\ 0 \\ 0 \end{pmatrix} \lambda^* = 0,$$
$$\lambda^* \geqslant 0, (Ax^* - b)^{\mathrm{T}} \lambda^* = 0. \tag{2.3}$$

Moreover, formula (2.3) is equivalent to the following conditions:

$$\nabla f(x^*, y^*) + \begin{pmatrix} N^{\mathrm{T}} Y^* \\ M^{\mathrm{T}} Y^* + W^* \end{pmatrix} v^* + \begin{pmatrix} A^{\mathrm{T}} \\ 0 \end{pmatrix} \lambda^* = 0,$$

$$\lambda^* \geqslant 0, (Ax^* - b)^{\mathrm{T}} \lambda^* = 0, \tag{2.4}$$

where diagonal matrices $W^* = \mathrm{diag}(w_i^*, i = 1, \ldots, m)$, $Y^* = \mathrm{diag}(y_i^*, i = 1, \ldots, m)$.

For a given parameter $\mu \geqslant 0$, we consider a perturbed problem associated with (LCP)(1.1) as follows:

$$\begin{aligned} &\min f(x, y) \\ &\text{ s.t. } Ax \leqslant b, \\ &\qquad w = Nx + My + q, \\ &\qquad y_i \geqslant 0,\ w_i \geqslant 0,\ y_i w_i = \frac{\mu}{2}, \quad i = 1, \ldots, m. \end{aligned} \tag{2.5}$$

A typical scheme for solving problem (2.5) may be described as follows: by a generalized complementarity function ϕ, the perturbed complementarity conditions "$y_i \geqslant 0, w \geqslant 0, y_i w_i = \frac{\mu}{2}$" are transformed into nonlinear equalities $\phi(y_i, w_i, \mu) = 0$, $i = 1, \ldots, m$, where the function $\phi : \Re^2 \times [0, +\infty] \longrightarrow \Re$ satisfies the following basic conditions:

(i) ϕ is continuously differentiable in $\{(a, b, \mu):(a, b, \mu) \neq (0, 0, 0)\}$;

(ii) $\phi(a, b, \mu) = 0 \Longleftrightarrow a \geqslant 0, b \geqslant 0, ab = \mu$;

(iii) For any $(a, b, \mu) \in \Re^2 \times (0, +\infty)$, it follows

$$\phi'_a(a, b, \mu)\phi'_b(a, b, \mu) > 0, \text{ where } \phi'_a(a, b, \mu) \triangleq \frac{\partial \phi(a, b, \mu)}{\partial a},$$
$$\phi'_b(a, b, \mu) \triangleq \frac{\partial \phi(a, b, \mu)}{\partial b}.$$

There are many constructions for ϕ satisfying the conditions above (see [1, 11, 10]), here we give some examples:

$$\phi(a, b, \mu) = a + b - \sqrt{(a - b)^2 + 4\mu}; \tag{2.6}$$

$$\phi(a, b, \mu) = a + b - \sqrt{a^2 + b^2 + 2\mu}; \tag{2.7}$$

$$\phi(a, b, \mu) = a + b - \sqrt{a^2 + b^2 + \lambda ab + (2 - \lambda)\mu},\ \lambda \in (-2, 2). \tag{2.8}$$

Without loss of generality, we choose the function ϕ defined by (2.8) as our complementarity function in this paper.

In order to analyze the derivative of ϕ, it is not difficult to verify that $a^2 + b^2 + \lambda ab + (2 - \lambda)\mu > 0$ for all $(a, b, \mu) \neq (0, 0, 0)$ and $\lambda \in (-2, 2)$, so we obtain

$$\nabla \phi(a, b, \mu) = \begin{pmatrix} \frac{\partial \phi(a,b,\mu)}{\partial a} \\ \frac{\partial \phi(a,b,\mu)}{\partial b} \end{pmatrix} = \begin{pmatrix} 1 - \frac{2a + \lambda b}{2\sqrt{a^2 + b^2 + \lambda ab + (2 - \lambda)\mu}} \\ 1 - \frac{2b + \lambda a}{2\sqrt{a^2 + b^2 + \lambda ab + (2 - \lambda)\mu}} \end{pmatrix}. \tag{2.9}$$

Now we define a vector value function $\Phi:\Re^{2m}\times[0,+\infty)\longrightarrow\Re^m$ as follows:

$$\Phi(t,\mu)=\Phi(y,w,\mu)=\begin{pmatrix}\phi(y_1,w_1,\mu)\\ \vdots\\ \phi(y_m,w_m,\mu)\end{pmatrix}. \tag{2.10}$$

Let $z^k=(x^k,y^k,w^k)\in\Re^n\times\Re^m\times\Re^m$ and $(y_i^k,w_i^k,\mu)\neq(0,0,0)$, then we have

$$\nabla_t\Phi(t^k,\mu_k)=\nabla\Phi(t^k,\mu_k)=(\Gamma_y^k\ \Gamma_w^k)^{\mathrm{T}}, \tag{2.11}$$

where

$$\begin{aligned}\Gamma_y^k&=\Gamma(y^k,w^k,\mu_k)=\mathrm{diag}(\gamma(y_i^k,w_i^k,\mu_k),i=1,\dots,m),\\ \Gamma_w^k&=\Gamma(w^k,y^k,\mu_k)=\mathrm{diag}(\gamma(w_i^k,y_i^k,\mu_k),i=1,\dots,m),\end{aligned} \tag{2.12}$$

$$\gamma(a,b,\mu)=1-\frac{2a+\lambda b}{2\sqrt{a^2+b^2+\lambda ab+(2-\lambda)\mu}},\quad \lambda\in(-2,2). \tag{2.13}$$

Therefore, problem (2.5) is transformed equivalently into the following standard nonlinear optimization problem:

$$(\mathrm{NLP}_\mu)\quad\begin{aligned}&\min f(x,y)\\ &\text{s.t. } Ax\leqslant b,\\ &\quad w=Nx+My+q,\\ &\quad \Phi(y,w,\mu)=0.\end{aligned} \tag{2.14}$$

The following result indicates a relation between (LCP)(1.1) and (NLP_μ)(2.14).

PROPOSITION 2.3. *Suppose that the nondegeneracy condition* (2.2) *holds, then* $z^*=(x^*,y^*,w^*)$, *corresponding multipliers* $\omega^*=(\lambda^*,u^*,v^*)$, *is a stationary point of* (LCP)(1.1) *if and only if* $(z^*,\widetilde{\omega})$ *is a KKT pair of* (NLP_μ)(2.14) *for* $\mu=0$ *with multipliers* $\widetilde{\omega}^*=(\lambda^*,u^*,\widetilde{v}^*)$, *where*

$$\widetilde{v}_i^*=\begin{cases} w_i^*v_i^*/\gamma(y_i^*,w_i^*,0)=\frac{2}{2-\lambda}w_i^*v_i^*, & \text{if } i\in I_y(z^*)\stackrel{\text{def}}{=}\{i\in L_2: y_i^*=0\};\\ y_i^*v_i^*/\gamma(w_i^*,y_i^*,0)=\frac{2}{2-\lambda}y_i^*v_i^*, & \text{if } i\in I_w(z^*)\stackrel{\text{def}}{=}\{i\in L_2: w_i^*=0\}.\end{cases} \tag{2.15}$$

Proof. Let $z^*=(x^*,y^*,w^*)$ be a stationary point of (LCP)(1.1), by Proposition 2.2, we know that (2.3) holds, so

$$\begin{pmatrix} \nabla f(x^*, y^*) \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \Gamma_y^* \\ \Gamma_w^* \end{pmatrix} \widetilde{v}^* + \begin{pmatrix} N^{\mathrm{T}} \\ M^{\mathrm{T}} \\ -I \end{pmatrix} u^* + \begin{pmatrix} A^{\mathrm{T}} \\ 0 \\ 0 \end{pmatrix} \lambda^* = 0,$$
$$\lambda^* \geqslant 0, (Ax^* - b)^{\mathrm{T}} \lambda^* = 0. \tag{2.16}$$

where $\Gamma_y^* = \mathrm{diag}(\gamma(y_i^*, w_i^*, 0), i = 1, \ldots, m)$, $\Gamma_w^* = \mathrm{diag}(\gamma(w_i^*, y_i^*, 0)$, $i = 1, \ldots, m)$, $\widetilde{v}^*$ is defined by (2.15). In view of (2.11), (2.16) is reduced into the following system:

$$\begin{pmatrix} \nabla f(x^*, y^*) \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \Phi_y(y^*, w^*, 0) \\ \Phi_w(y^*, w^*, 0) \end{pmatrix} \widetilde{v}^* + \begin{pmatrix} N^{\mathrm{T}} \\ M^{\mathrm{T}} \\ -I \end{pmatrix} u^* + \begin{pmatrix} A^{\mathrm{T}} \\ 0 \\ 0 \end{pmatrix} \lambda^* = 0,$$
$$\lambda^* \geqslant 0, (Ax^* - b)^{\mathrm{T}} \lambda^* = 0,$$

which shows that $(z^*, \widetilde{\omega}^* = (\lambda^*, u^*, \widetilde{v}^*))$ is a KKT pair of (NLP$_0$)(2.14).

Conversely, the proof above holds, so the proof is completed. □

For an approximate solution $z^k = (x^k, y^k, w^k)$ of (NLP$_{\mu_k}$)(2.14), in order to produce an improving direction, we consider the following system of linear equations

$$\begin{pmatrix} B_k & U_k^{\mathrm{T}} \\ U_k & 0 \end{pmatrix} \begin{pmatrix} \mathrm{d}z \\ \widetilde{\omega} \end{pmatrix} = - \begin{pmatrix} \nabla f(s^k) \\ 0_{m \times 1} \\ A_{J_k} x^k - b_{J_k} \\ Nx^k + My^k + q - w^k \\ \Phi(t^k, \mu_k) \end{pmatrix}, \tag{2.17}$$

where $B_k \in \Re^{(n+2m)\times(n+2m)}, \widetilde{\omega} = (\lambda, u, v) \in \Re^p \times \Re^m \times \Re^m$, the index set J_k is a subset of L_1 which is produced by some method and satisfies some conditions, and

$$A_{J_k} = (a_j, j \in J_k), b_{J_k} = (b_j, j \in J_k), U_k \stackrel{\mathrm{def}}{=} U(z^k, \mu_k) = \begin{pmatrix} A_{J_k} & 0 & 0 \\ N & M & -I \\ 0 & \Gamma_y^k & \Gamma_w^k \end{pmatrix}.$$

Generally speaking, in order to make the SSLE algorithm possess fast convergence, the matrix B_k must be an approximation of the Hesse matrix of Lagrange function of (NLP$_{\mu_k}$)(2.14), so we consider the Lagrange function of (NLP$_{\mu_k}$)(2.14):

$$L(z, \widetilde{\omega}, \mu) = f(x, y) + \lambda^{\mathrm{T}}(Ax - b) + u^{\mathrm{T}}(Nx + My + q - w) + v^{\mathrm{T}}\Phi(y, w, \mu).$$

It is obvious that

$$H(z,v,\mu) \overset{\text{def}}{=} \nabla^2_{zz} L(z,\widetilde{\omega},\mu) = \begin{pmatrix} \nabla^2_{xx} f(s) & \nabla^2_{xy} f(s) & 0_{n\times m} \\ \nabla^2_{yx} f(s) & \nabla^2_{yy} f(s) + \operatorname{diag}\left(v_i \frac{\partial^2\phi(t_i,\mu)}{\partial a^2}\right) & \operatorname{diag}\left(v_i \frac{\partial^2\phi(t_i,\mu)}{\partial a\partial b}\right) \\ 0_{m\times n} & \operatorname{diag}\left(v_i \frac{\partial^2\phi(t_i,\mu)}{\partial b\partial a}\right) & \operatorname{diag}\left(v_i \frac{\partial^2\phi(t_i,\mu)}{\partial b^2}\right) \end{pmatrix}, \tag{2.18}$$

moreover, for the stationary point pair $(z^*,\widetilde{\omega}^*)$ of LCP(1.1) satisfying (2.3) and the corresponding $\widetilde{v}^*$ defined by (2.15), by further computations, we obtain

$$H(z^*,v^*,0) = \begin{pmatrix} \nabla^2_{xx} f(s^*) & \nabla^2_{xy} f(s^*) & 0_{n\times m} \\ \nabla^2_{yx} f(s^*) & \nabla^2_{yy} f(s^*) + \operatorname{diag}\left(v_i^* \frac{\partial^2\phi(t_i^*,0)}{\partial a^2}\right) & 0_{m\times m} \\ 0_{m\times n} & 0_{m\times m} & \operatorname{diag}\left(v_i^* \frac{\partial^2\phi(t_i^*,0)}{\partial b^2}\right) \end{pmatrix}, \tag{2.19}$$

$$H(z^*,\widetilde{v}^*,0) = \begin{pmatrix} \nabla^2_{xx} f(s^*) & \nabla^2_{xy} f(s^*) & 0_{n\times m} \\ \nabla^2_{yx} f(s^*) & \nabla^2_{yy} f(s^*) + \operatorname{diag}\left(\widetilde{v}_i^* \frac{\partial^2\phi(t_i^*,0)}{\partial a^2}\right) & 0_{m\times m} \\ 0_{m\times n} & 0_{m\times m} & \operatorname{diag}\left(\widetilde{v}_i^* \frac{\partial^2\phi(t_i^*,0)}{\partial b^2}\right) \end{pmatrix}, \tag{2.20}$$

where

$$s^* = (x^*,y^*),\ \ t_i^* = (y_i^*,w_i^*).$$

Based on the above relation (2.19) or (2.20), it is reasonable to choose B_k as the following form in this paper:

$$B_k = \begin{pmatrix} C_k & 0 \\ 0 & D_k \end{pmatrix},\ \ C_k \in \Re^{(n+m)\times(n+m)},\ \ D_k \in \Re^{m\times m}. \tag{2.21}$$

Therefore, if $z^k = (x^k,y^k,w^k)$ satisfies $w^k = Nx^k + My^k + q$, then (2.17) is transformed equivalently into the following system of linear equations with smaller scale:

$$(\mathrm{SLE}_1)(z^k,\mu_k) \quad \begin{pmatrix} H_k & G_k^{\mathrm{T}} \\ G_k & 0 \end{pmatrix}\begin{pmatrix} ds \\ \pi \end{pmatrix} = -\begin{pmatrix} \nabla f(s^k) \\ h(z^k,\mu_k) \end{pmatrix}, \tag{2.22}$$

where

$$H_k = C_k + (NM)^{\mathrm{T}} D_k (NM), \tag{2.23}$$

$$G_k = \begin{pmatrix} A_{J_k} & 0 \\ \Gamma_w^k N & \Gamma_y^k + \Gamma_w^k M \end{pmatrix}, \; h(z^k, \mu_k) = \begin{pmatrix} A_{J_k} x^k - b_{J_k} \\ \Phi(t^k, \mu_k) \end{pmatrix}. \tag{2.24}$$

It is not difficult to prove that the following lemma holds.

LEMMA 2.4. $0 < \gamma(a, b, \mu), \gamma(b, a, \mu) < 2$ *for any* $\mu > 0$.

LEMMA 2.5. *Suppose that assumption* $A1(i)$ *holds*, $\mu_k > 0$, *the index set* J_k *ensures that* A_{J_k} *is full of row rank and* H_k *is positive definite. Then for any* $z^k \in X_0$, *i.e.* $w^k = Nx^k + My^k + q$,

(i) *The coefficient matrix of* $(\mathrm{SLE}_1)(2.22)$ *is nonsingular, furthermore,* $(\mathrm{SLE}_1)(2.22)$ *has a unique solution.*

(ii) *The matrix* $G_k H_k^{-1} G_k^{\mathrm{T}}$ *is symmetric positive definite.*

Proof. (i) It follows from Lemma 2.4 that the diagonal matrices Γ_y^k, Γ_w^k are symmetric positive definite. On the other hand, in view of the fact that M is a P_0 matrix, it is not difficult to deduce from Lemma 4.1 in [14] that $\Gamma_y^k + \Gamma_w^k M$ is nonsingular. Therefore, since A_{J_k} is full of row rank, we know that G_k is also full of row rank; And since H_k is positive definite, so the coefficient matrix of $(\mathrm{SLE}_1)(2.22)$ $\begin{pmatrix} H_k & G_k^{\mathrm{T}} \\ G_k & 0 \end{pmatrix}$ is nonsingular, furthermore, $(\mathrm{SLE}_1)(2.22)$ has a unique solution.

(ii) $\forall x \neq 0$, $x^{\mathrm{T}} G_k H_k^{-1} G_k^{\mathrm{T}} x = (G_k^{\mathrm{T}} x)^{\mathrm{T}} H_k^{-1} (G_k^{\mathrm{T}} x)$, considering that G_k is full of row rank, and H_k is definite, we see $(G_k^{\mathrm{T}} x)^{\mathrm{T}} H_k^{-1} (G_k^{\mathrm{T}} x) > 0$, so $G_k H_k^{-1} G_k^{\mathrm{T}}$ is positive definite. □

As we know, generally, the solution $\mathrm{d}s^k$ of $(\mathrm{SLE}_1)(2.22)$ cannot avoid the Maratos effect and get superlinear convergence, so it needs a modification. The modificative method in this paper is to solve two systems of equations which have the same coefficient as $(\mathrm{SLE}_1)(2.22)$(The details are given in the following algorithm).

3. Algorithm

In this paper, we use a special penalty function as follows as the merit function:

$$\begin{aligned} \theta(z, \alpha, \mu) &= f(x, y) + \alpha\varphi(x) + \alpha \sum_{i=1}^{m} |\phi(y_i, w_i, \mu)| \\ &= f(s) + \alpha\varphi(x) + \alpha\|\Phi(t, \mu)\|_1, \end{aligned} \tag{3.1}$$

where $\alpha > 0$ is a penalty parameter. This penalty function doesn't penalize the linear equality constraints, and use the penalty term $\varphi(x)$ defined by (2.1) for the inequality $Ax \leqslant b$ instead of $\sum_{j=1}^{p} \max\{0; a_j x - b_j\}$.

ALGORITHM A

Step 0 (*Initialization*). Choose a parameter $\widehat{\varepsilon} > 0$ sufficiently small, parameters $\lambda \in (-2, 2)$, $\alpha_{-1}, \delta > 0$, $\varepsilon_{-1} > 0, 0 < \beta, \sigma < 1$, choose a sequence $\{\mu_k\}_{k=0}^{\infty}$ such that

$$\mu_k > 0, \mu_{k+1} < \mu_k, \lim_{k \to \infty} \mu_k = 0, \lim_{k \to +\infty} \frac{\mu_{k+1}}{\mu_k^{\gamma}} = \bar{\eta} \in (0, 1), \gamma \in (1, 2). \tag{3.2}$$

Choose an initial point $z^0 = (x^0, y^0, w^0) \in X_o$, a symmetric matrix B_0 with the form of (2.21) such that H_0 given by (2.23) is positive definite, let $k := 0$.

Step 1 (*Pivoting*). Compute an index set $J_k(x^k, \varepsilon_k)$:

(i) Let $i = 0, \varepsilon_{k,i} = \varepsilon_{k-1}$;

(ii) Compute $\bar{J}_{k,i} = \{j \in L_1 : 0 \leqslant \varphi(x^k) - (a_j x^k - b_j) \leqslant \varepsilon_{k,i}\}$.
If $\det(A_{\bar{J}} A_{\bar{J}}^{\mathrm{T}}) \geqslant \varepsilon_{k,i}$, where $A_{\bar{J}} = (a_j, j \in \bar{J}_{k,i})$, then let $J_k(x^k, \varepsilon_k) = \bar{J}_{k,i}$, $\varepsilon_k = \varepsilon_{k,i}$, go to Step 2. Otherwise, let $i := i + 1$, $\varepsilon_{k,i} = \frac{1}{2}\varepsilon_{k,i-1}$, go back to (ii).

Step 2 Compute $\Phi(y^k, w^k, \mu_k)$. If $\Phi(y^k, w^k, \mu_k) \neq 0$, or $\Phi(y^k, w^k, \mu_k) = 0$ and $\mu_k < \widehat{\varepsilon}$, then go to Step 3. Otherwise choose $\mu_k' \in (\mu_{k+1}, \mu_k)$ and let $\mu_k = \mu_k'$, compute $\Phi(y^k, w^k, \mu_k)$ again (From Lemma 3.1 below, we know $\Phi(y^k, w^k, \mu_k) \neq 0$ at this time).

Step 3 (*Generate a search direction*)

(i) Denote $J_k = J_k(x^k, \varepsilon_k)$, solve the system of linear equations $(\mathrm{SLE}_1)(z^k, \mu_k)$(i.e. (2.22)).
Let its solution be $\mathrm{d}s_0^k$, $\pi_0^k \overset{\text{def}}{=} \begin{pmatrix} \lambda_{0,J_k}^k \\ v_0^k \end{pmatrix}$, where $\lambda_{0,J_k}^k = (\lambda_{0,j}^k, j \in J_k)$, $v_0^k = (v_{0,i}^k, i \in L_2)$.
Let

$$\mathrm{d}w_0^k = N dx_0^k + M dy_0^k, \mathrm{d}z_0^k = (\mathrm{d}s_0^k, \mathrm{d}w_0^k). \tag{3.3}$$

If $\mathrm{d}s_0^k = 0$, $\lambda_{0,j}^k \geqslant 0$, $\forall j \in J_k$, then z^k is a KKT point of (NLP_{μ_k})(2.14), in view of $\mu_k < \widehat{\varepsilon}$, we can conclude that z^k is an acceptable approximate stationary point of (LCP)(1.1), stop. Otherwise go to (ii).

(ii) Solve the following system of linear equations:

$$(\mathrm{SLE}_2)(z^k, \mu_k) \begin{pmatrix} H_k & G_k^{\mathrm{T}} \\ G_k & 0 \end{pmatrix} \begin{pmatrix} \mathrm{d}s \\ \pi \end{pmatrix} = -\begin{pmatrix} \nabla f(s^k) \\ h(z^k, \mu_k) - \eta^k \end{pmatrix}, \tag{3.4}$$

where

$$\eta_j^k = \begin{cases} (\lambda_{0,j}^k)^3, & \text{if } \lambda_{0,j}^k \leqslant 0 \text{ and } j \in J_k; \\ 0, & \text{otherwise}, \end{cases} \tag{3.5}$$

let its solution be $\mathrm{d}s^k$, $\pi^k \stackrel{\text{def}}{=} \begin{pmatrix} \lambda_{J_k}^k \\ v^k \end{pmatrix}$. Let

$$\mathrm{d}w^k = N\mathrm{d}x^k + M\mathrm{d}y^k,\ \mathrm{d}z^k = (\mathrm{d}s^k, \mathrm{d}w^k). \tag{3.6}$$

(iii) Solve the following system of linear equations:

$$(\mathrm{SLE}_3)(z^k, \mu_k)\begin{pmatrix} H_k & G_k^{\mathrm{T}} \\ G_k & 0 \end{pmatrix}\begin{pmatrix} \mathrm{d}s \\ \pi \end{pmatrix} = -\begin{pmatrix} \nabla f(s^k) \\ h(z^k + \mathrm{d}z^k, \mu_k) + h(z_k, \mu_k) \end{pmatrix}, \tag{3.7}$$

let its solution be $\widehat{\mathrm{d}s}^k$, $\widehat{\pi}^k \stackrel{\text{def}}{=} \begin{pmatrix} \widehat{\lambda}_{J_k}^k \\ \widehat{v}^k \end{pmatrix}$. Let

$$\widehat{\mathrm{d}w}^k = N\widehat{\mathrm{d}x}^k + M\widehat{\mathrm{d}y}^k, \widehat{\mathrm{d}z}^k = (\widehat{\mathrm{d}s}^k, \widehat{\mathrm{d}w}^k). \tag{3.8}$$

If $\|\widehat{\mathrm{d}z}^k - \mathrm{d}z^k\| > \|\mathrm{d}z^k\|$, then let $\widehat{\mathrm{d}z}^k = \mathrm{d}z^k$.

Step 4 (*Update the penalty parameter*). Denote $\lambda_0^k \stackrel{\text{def}}{=} (\lambda_{0,j}^k, j \in L_1)$, $\lambda^k \stackrel{\text{def}}{=} (\lambda_j^k, j \in L_1)$, where

$$\lambda_{0,j}^k = \begin{cases} \lambda_{0,j}^k, & j \in J_k; \\ 0, & j \notin J_k, \end{cases} \quad \lambda_j^k = \begin{cases} \lambda_j^k, & j \in J_k; \\ 0, & j \notin J_k. \end{cases} \tag{3.9}$$

Denote

$$\xi_k = \max\{\|\lambda_0^k\|_1, \|\lambda^k\|_1, \|v_0^k\|_1, \|v^k\|_1, \|2\lambda_0^k - \lambda^k\|_1, \|2v_0^k - v^k\|_1\}, \tag{3.10}$$

where $\|p\|_1 \stackrel{\text{def}}{=} \sum_{i=1}^r |p_i|$ for $p = (p_1, p_2, \ldots, p_r) \in \Re^r$. The updating rule for α is as follows:

$$\alpha_k = \begin{cases} \alpha_{k-1}, & \text{if } \alpha_{k-1} \geqslant \xi_k + \delta; \\ \max\{\xi_k + \delta, \alpha_{k-1} + 2\delta\}, & \text{otherwise}. \end{cases} \tag{3.11}$$

Step 5 (*Curve search*). Compute the step size τ_k, which is the first number τ of the sequence $\{1, \beta, \beta^2, \ldots\}$ satisfying

$$\theta(z^k + \tau\mathrm{d}z^k + \tau^2(\widehat{\mathrm{d}z}^k - \mathrm{d}z^k), \alpha_k, \mu_k) \leqslant \theta(z^k, \alpha_k, \mu_k) + \sigma\tau\psi(z^k, \mathrm{d}z^k, \alpha_k, \mu_k), \tag{3.12}$$

where

$$\psi(z^k, \mathrm{d}z^k, \alpha_k, \mu_k) = \nabla f(s^k)^{\mathrm{T}}\mathrm{d}s^k - \alpha_k\varphi(x^k) - \alpha_k\|\Phi(t^k, \mu_k)\|_1 + \frac{1}{2}(\mathrm{d}s_0^k)^{\mathrm{T}}H_k\mathrm{d}s_0^k. \tag{3.13}$$

Step 6. Produce a new iteration point by $z^{k+1} = z^k + \tau_k \mathrm{d}z^k + \tau^2(\widehat{\mathrm{d}z}^k - \mathrm{d}z^k)$, and update B_k by some given method to yield a new matrix B_{k+1} with the form of (2.21) such that the matrix H_{k+1} defined by (2.23) is symmetric positive definite, let $k := k+1$ and go back Step 1.

In the remainder of this section, we will analyze the feasibility and some properties of algorithm A by a few lemmas.

LEMMA 3.1. *Suppose that assumption* (A1) *holds, then the following statements are true*:

(i) *Pivoting in step 2 is terminated after a finite number of computations.*
(ii) *If* $\Phi(y_i^k, w_i^k, \mu_k) = 0$, *then* $\Phi(y_i^k, w_i^k, \mu'_k) \neq 0$ for $\mu'_k < \mu_k$.

Proof. (i) The proof is similar to that of Lemma 1.1 in [6] and omitted. (ii) Suppose by contradiction that $\Phi(y^k, w^k, \mu_k) = 0$, $\Phi(y^k, w^k, \mu'_k) = 0$, then it follows immediately from the properties of Φ that $\mu_k = y_i^k w_i^k$, $\mu'_k = y_i^k w_i^k$, $i = 1, 2, \ldots,$ therefore $\mu_k = \mu'_k$, contradicting $\mu'_k < \mu_k$. This completes the proof. □

From (SLE$_1$)(2.22) and (SLE$_2$)(3.4), it is not difficult to prove that the following conclusions hold true by straightforward calculation.

LEMMA 3.2. *Suppose that assumption* (A1) *holds*, *then*

$$\begin{aligned} \mathrm{d}s_0^k &= -P_k \nabla f(s^k) - F_k h(z^k, \mu_k), \\ \pi_0^k &= -F_k^{\mathrm{T}} \nabla f(s^k) + (G_k H_k^{-1} G_k^{\mathrm{T}})^{-1} h(z^k, \mu_k), \\ \mathrm{d}s^k &= \mathrm{d}s_0^k + F_k \eta_k, \quad \pi^k = \pi_0^k - (G_k H_k^{-1} G_k^{\mathrm{T}})^{-1} \eta_k, \end{aligned} \tag{3.14}$$

where

$$P_k = H_k^{-1} - H_k^{-1} G_k^{\mathrm{T}} (G_k H_k^{-1} G_k^{\mathrm{T}})^{-1} G_k H_k^{-1}, \quad F_k = H_k^{-1} G_k^{\mathrm{T}} (G_k H_k^{-1} G_k^{\mathrm{T}})^{-1}. \tag{3.15}$$

LEMMA 3.3. *Suppose that assumption* (A1) *holds and* H_k *is symmetric positive definite*, *then the following statements are true*:

(i) (SLE$_2$)(3.4) *and* (SLE$_3$)(3.7) *have a unique solution, respectively.*
(ii) *For any* k, $\psi(z^k, \mathrm{d}z^k, \alpha_k, \mu_k) < 0$.
(iii) *For any* k, *there exists a* $\tau > 0$ *satisfying the curve search* (3.12), *that is, Step 5 in algorithm A will be terminated after a finite number of computations, so algorithm A is well defined.*
(iv) *The sequence* $\{z^k\}$ *generated by algorithm A satisfies* $w^k = Nx^k + My^k + q$ for all k.

Proof. (i) Since the coefficients of $(SLE_2)(3.4)$ and $(SLE_3)(3.7)$ are the same as the one of $(SLE_1)(2.22)$, part (i) follows immediately from Lemma 2.5.
(ii) From (3.13) and (3.14), we obtain

$$\begin{aligned}
&\psi(z^k, \mathrm{d}z^k, \alpha_k, \mu_k) \\
&= \nabla f(s^k)^{\mathrm{T}}(\mathrm{d}s_0^k + F_k\eta_k) - \alpha_k\varphi(x^k) - \alpha_k\|\Phi(t^k,\mu_k)\|_1 + \frac{1}{2}(\mathrm{d}s_0^k)^{\mathrm{T}}H_k\mathrm{d}s_0^k \\
&= \nabla f(s^k)^{\mathrm{T}}\mathrm{d}s_0^k + (F_k^{\mathrm{T}}\nabla f(s^k))^{\mathrm{T}}\eta_k - \alpha_k\varphi(x^k) - \alpha_k\|\Phi(t^k,\mu_k)\|_1 + \frac{1}{2}(\mathrm{d}s_0^k)^{\mathrm{T}}H_k\mathrm{d}s_0^k.
\end{aligned} \tag{3.16}$$

Again, from $(SLE_1)(2.22)$, one gets

$$\nabla f(s^k)^{\mathrm{T}}\mathrm{d}s_0^k = -(\mathrm{d}s_0^k)^{\mathrm{T}}H_k\mathrm{d}s_0^k - (\lambda_{0,J_k}^k)^{\mathrm{T}}(A_{J_k},0)\mathrm{d}s_0^k - (v_0^k)^{\mathrm{T}}(\Gamma_w^kN\Gamma_y^k + \Gamma_w^kM)\mathrm{d}s_0^k. \tag{3.17}$$

On the other hand, we can deduce from (3.14) and (2.24) that

$$F_k^{\mathrm{T}}\nabla f(s^k) = (G_kH_k^{-1}G_k^{\mathrm{T}})^{-1}\begin{pmatrix} A_{J_k}x^k - b_{J_k} \\ \Phi(t_i^k,\mu_k)\end{pmatrix} - \pi_0^k. \tag{3.18}$$

Substituting (3.17) and (3.18) into (3.16), and taking into account (3.14) and (3.5), we have

$$\begin{aligned}
\psi(z^k, \mathrm{d}z^k, \alpha_k, \mu_k) = & -\frac{1}{2}(\mathrm{d}s_0^k)^{\mathrm{T}}H_k\mathrm{d}s_0^k + \sum_{j\in J_k}(2\lambda_{0,j}^k - \lambda_j^k)(a_jx^k - b_j) \\
& + \sum_{i=1}^{m}(2v_{0,i}^k - v_i^k)\phi(t_i^k,\mu_k) - \alpha_k\varphi(x^k) - \alpha_k\|\Phi(t^k,\mu_k)\|_1 \\
& - \sum_{j\in J_k,\lambda_{0,j}^k<0}(\lambda_{0,j}^k)^4 \leqslant -\frac{1}{2}(\mathrm{d}s_0^k)^{\mathrm{T}}H_k\mathrm{d}s_0^k \\
& - \left(\alpha_k\varphi(x^k) - \sum_{j\in J_k}(2\lambda_{0,j}^k - \lambda_j^k)(a_jx^k - b_j)\right) \\
& + \sum_{i=1}^{m}(|2v_{0,i}^k - v_i^k| - \alpha_k)|\phi(t_i^k,\mu_k)| - \sum_{j\in J_k,\lambda_{0,j}^k<0}(\lambda_{0,j}^k)^4.
\end{aligned}$$

So it follows immediately from (3.11) and (2.1) that

$$\psi(z^k, \mathrm{d}z^k, \mu_k) \leqslant -\frac{1}{2}(\mathrm{d}s_0^k)^{\mathrm{T}}H_k\mathrm{d}s_0^k - \sum_{j\in J_k,\lambda_{0,j}^k<0}(\lambda_{0,j}^k)^4 < 0, \forall k. \tag{3.19}$$

(iii) For the sake of convenience, denote

$$T_1 = f(s^k + \tau \mathrm{d}s^k + \tau^2(\widehat{\mathrm{d}s}^k - \mathrm{d}s^k)) - f(s^k),$$

$$T_2 = \varphi(x^k + \tau \mathrm{d}x^k + \tau^2(\widehat{\mathrm{d}x}^k - \mathrm{d}x^k)) - \varphi(x^k),$$

$$T_3 = \sum_{i=1}^{m}(|\phi(t_i^k + \tau \mathrm{d}t^k + \tau^2(\widehat{\mathrm{d}t}^k - \mathrm{d}t^k), \mu_k)| - |\phi(t_i^k, \mu_k)|),$$

so we know that from (3.1)

$$\theta(z^k + \tau \mathrm{d}z^k + \tau^2(\widehat{\mathrm{d}z}^k - \mathrm{d}z^k), \alpha_k, \mu_k) - \theta(z^k, \alpha_k, \mu_k) = T_1 + \alpha_k T_2 + \alpha_k T_3. \tag{3.20}$$

Notice that $I(x^k + \tau \mathrm{d}x^k + \tau^2(\widehat{\mathrm{d}x}^k - \mathrm{d}x^k)) \subseteq I(x^k)$ for all sufficiently small τ, so

$$\begin{aligned} T_2 &= \max\{0; a_j(x^k + \tau \mathrm{d}x^k + \tau^2(\widehat{\mathrm{d}x}^k - \mathrm{d}x^k)) - b_j, j \in I(x^k + \tau \mathrm{d}x^k \\ &\quad + \tau^2(\widehat{\mathrm{d}x}^k - \mathrm{d}x^k))\} - \max\{0; a_j x^k - b_j, j \in I(x^k)\} \\ &\leqslant \max\{0;\ a_j(x^k + \tau \mathrm{d}x^k + \tau^2(\widehat{\mathrm{d}x}^k - \mathrm{d}x^k)) - b_j, j \in I(x^k)\} \\ &\quad - \max\{0;\ a_j x^k - b_j,\ j \in I(x^k)\} \\ &= \max\{0; (a_j x^k - b_j) + \tau a_j \mathrm{d}x^k, j \in I(x^k)\} \\ &\quad - \max\{0; a_j x^k - b_j, j \in I(x^k)\} + o(\tau). \end{aligned} \tag{3.21}$$

Since

$$\widetilde{g}(x^k, \tau \mathrm{d}x^k) \stackrel{\text{def}}{=} \max\{0; (a_j x^k - b_j) + \tau a_j \mathrm{d}x^k, j \in I(x^k)\}$$

is convex with respect to τ, so

$$\begin{aligned} \widetilde{g}(x^k, \tau \mathrm{d}x^k) &= \widetilde{g}(x^k, (1-\tau)0 + \tau \mathrm{d}x^k) \leqslant (1-\tau)\widetilde{g}(x^k, 0) + \tau \widetilde{g}(x^k, \mathrm{d}x^k) \\ &\leqslant (1-\tau)\max\{0; a_j x^k - b_j, j \in I(x^k)\} \\ &\quad + \tau \max\{0; a_j x^k - b_j + a_j \mathrm{d}x^k, j \in I(x^k)\}, \end{aligned}$$

which together with (3.21) gives

$$\begin{aligned} T_2 &= -\tau \max\{0; a_j x^k - b_j, j \in I(x^k)\} \\ &\quad + \tau \max\{0; a_j x^k - b_j + a_j \mathrm{d}x^k, j \in I(x^k)\} + o(\tau) \\ &= -\tau \varphi(x^k) + \tau \max\{0; a_j x^k - b_j + a_j \mathrm{d}x^k, j \in I(x^k)\} + o(\tau). \end{aligned}$$

Notice that $j \in I(x^k) \subseteq J_k$, it follows from $(\mathrm{SLE}_2)(3.4)$ that $a_j x^k - b_j + a_j dx^k = \eta_k \leqslant 0$, therefore we obtain

$$T_2 = -\tau\varphi(x^k) + o(\tau).$$

On the other hand, arising from Taylor expansion, one has

$$T_1 = f(s^k + \tau \mathrm{d}s^k + \tau^2(\widehat{\mathrm{d}s}^k - \mathrm{d}s^k)) - f(s^k) = \tau \bigtriangledown f(s^k)^{\mathrm{T}} \mathrm{d}s^k + o(\tau),$$

$$\begin{aligned} T_3 &= \sum_{i=1}^{m} (|\phi(t_i^k, \mu_k) + \tau \bigtriangledown \phi(t_i^k, \mu_k) \mathrm{d}t^k + o(\tau)| - |\phi(t_i^k, \mu_k)|) \\ &= \sum_{i=1}^{m} (|\phi(t_i^k, \mu_k) - \tau\phi(t_i^k, \mu_k)| - |\phi(t_i^k, \mu_k)|) + o(\tau) \\ &= \sum_{i=1}^{m} (-\tau)|\phi(t_i^k,\ \mu_k)| + o(\tau). \end{aligned}$$

Therefore

$$\begin{aligned} T_1 + \alpha_k T_2 + \alpha_k T_3 &\leqslant \tau(\nabla f(s^k)^{\mathrm{T}} \mathrm{d}s^k - \alpha_k \varphi(x^k) - \alpha_k \sum_{i=1}^{m} |\phi(t_i^k, \mu_k)|) + o(\tau) \\ &\leqslant \tau\psi(z^k, \mathrm{d}z^k, \alpha_k, \mu_k) + o(\tau), \end{aligned} \tag{3.22}$$

which together with (3.20) and $\psi(z^k, \mathrm{d}z^k, \alpha_k, \mu_k) < 0$ shows that there exists a $\tau_k > 0$ such that

$$\theta(z^k + \tau_k \mathrm{d}z^k + \tau_k^2(\widehat{\mathrm{d}z}^k - \mathrm{d}z^k), \alpha_k, \mu_k) \leqslant \theta(z^k, \alpha_k, \mu_k) + \sigma\tau_k\psi(z^k, \mathrm{d}z^k, \alpha_k, \mu_k).$$

This indicates that algorithm A is well defined.

Lastly, the proof of part(iv) is easy by induction method, and omitted.

□

4. Global Convergence and Strong Convergence

In this section, we first analyze and verify the global convergence of algorithm A, and then further discuss and prove the strong convergence of algorithm A under some additional assumptions.

For the sake of convenience, let z^* be an accumulation point of sequence $\{z^k\}$, then there exists a subsequence $\mathcal{K} \subseteq \{1, 2, 3, \ldots\}$ such that

$$\lim_{k \in \mathcal{K}} z^k = z^* = (x^*, y^*, w^*),\ s^* = (x^*, y^*), \quad t^* = (y^*, w^*), \quad J_k \equiv J. \tag{4.1}$$

LEMMA 4.1. *There exists a constant $\bar{\varepsilon} > 0$ such that $\varepsilon_k \geqslant \bar{\varepsilon}$ for all k.*

Proof. Suppose, by contradiction, that there exists an infinite subset $\mathcal{K}_1$ such that $\varepsilon_k \xrightarrow{\mathcal{K}_1} 0$. In view of the fact that $\{\varepsilon_k\}$ is monotonically decreasing, we know that $\lim_{x\to\infty} \varepsilon_k = 0$. Without loss of generality, let $\varepsilon_k = \varepsilon_{k,i_k}$, $\widehat{J}_k = J_{k,i_k-1}$, so we have

$$\det\,(A_{\widehat{J}_k} A^{\mathrm{T}}_{\widehat{J}_k}) < 2\varepsilon_{k,i_k}, \quad 0 < \varphi(x^k) - (a_j x^k - b_j) < 2\varepsilon_{k,i_k}, \quad j \in \widehat{J}_k. \tag{4.2}$$

Note that $L_1 = \{1, 2, \ldots, p\}$ is a finite set, so we can suppose, without loss of generality, that there exists an infinite subset $\mathcal{K}_2 \subseteq \mathcal{K}_1$ such that $\widehat{J}_k \equiv J,\ \forall k \in \mathcal{K}_2$. Define

$$\varphi(x^*) = \max\{0; a_j x^* - b_j,\ j \in J\}, \quad A^* = \{a_j : j \in J\},$$

Passing to the limit $k \to \infty (k \in \mathcal{K})$ in (4.2), we have

$$\det\,(A^*_J A^{*^{\mathrm{T}}}_J) = 0, \quad \varphi(x^*) = a_j x^* - b_j,\ j \in J.$$

From the second formula above, we know $J \subseteq I(x^*)$, contradicting assumption (A1). This completes the proof. □

In order to verify the global convergence of algorithm A, additional assumptions as follows are necessary:
(A2) There exist constants $c_2 \geqslant c_1 > 0$ such that

$$c_1\|s\|^2 \leqslant s^{\mathrm{T}} H_k s \leqslant c_2 \|s\|^2, \quad \forall s \in \Re^{n+m}, \quad \forall k = 0, 1, 2, \ldots.$$

(A3) The point sequence $\{z^k\}$ produced by algorithm A is bounded, and every accumulation point $z^* = (x^*, y^*, w^*)$ of $\{z^k\}$ satisfies the following conditions:

(i) The (lower level) nondegeneracy condition (2.2) holds;
(ii) The submatrix $M_{J^*J^*}$ is nondegenerate, i.e., all of its principal minors are nonsingular, where the index set $J^* = \{i : w_i^* = 0\}$.

From the closeness of X_0 and $z^k \in X_0$, it is obvious that $z^* \in X_0$, i.e. $w^* = Nx^* + My^* + q$. In view of (A2), (A3)(i), (3.2) and (2.24), we have

$$\lim_{k\in\mathcal{K}} H_k = H^*, \Gamma_y^k \xrightarrow{\mathcal{K}} \Gamma_y^* \stackrel{\text{def}}{=} \Gamma_y(y^*, w^*, 0), \quad \Gamma_w^k \xrightarrow{\mathcal{K}} \Gamma_w^* \stackrel{\text{def}}{=} \Gamma_w(w^*, y^*, 0),$$

$$G_k \xrightarrow{\mathcal{K}} G^* \stackrel{\text{def}}{=} \begin{pmatrix} A_J & 0 \\ \Gamma_w^* N & \Gamma_y^* + \Gamma_w^* M \end{pmatrix}. \tag{4.3}$$

If define an index set $J_0 = \{i \in L_2 : (\Gamma_y^*)_{ii} = \gamma(y_i^*, w_i^*, 0) = 0\}$, then from (2.13) and $\lambda \in (-2, 2)$, we have $J_0 \subseteq J^*$.

PROPOSITION 4.2 *Suppose that assumptions* (A1)−(A3) *hold, then*

(i) *The limit matrix* $\Gamma_y^* + \Gamma_w^* M$ *of the sequence* $\{(\Gamma_y^k + \Gamma_w^k M), k \in \mathcal{K}\}$ *of matrices is nonsingular.*

(ii) *The limit matrix* G^* *of* $\{G_k\}$ *is full of row rank, the matrix* $\begin{pmatrix} H^* & G^{*\mathrm{T}} \\ G^* & 0 \end{pmatrix}$ *is nonsingular.*

Proof. (i) We know that the principal minors $M_{J_0J_0}$ of $M_{J^*J^*}$ are nonsingular from (A3)(ii), so result (i) is true from Proposition 3.2 in [4].

(ii) From assumption (A1) and result (i), we know that A_J and $\Gamma_y^* + \Gamma_w^* M$ are full of row rank, so is G^*. Similar to the proof of Lemma 2.5, we can prove that $\begin{pmatrix} H^* & G^{*\mathrm{T}} \\ G^* & 0 \end{pmatrix}$ is nonsingular. □

LEMMA 4.3. *Suppose that assumptions* (A1)–(A3) *hold, then*

(i) *There exists a constant* $c > 0$ *such that* $\|(\Gamma_y^k + \Gamma_w^k M)^{-1}\| \leqslant c$, $\left\| \begin{pmatrix} H_k & G_k^{\mathrm{T}} \\ G_k & 0 \end{pmatrix} \right\| \leqslant c$, for all $k \in \mathcal{K}$.

(ii) *The sequences* $\{\mathrm{d}z_0^k, k \in \mathcal{K}\}$, $\{\mathrm{d}z^k, k \in \mathcal{K}\}$, $\{\widehat{\mathrm{d}z}^k, k \in \mathcal{K}\}$, $\{\lambda_0^k, k \in \mathcal{K}\}$, $\{\lambda^k, k \in \mathcal{K}\}$, $\{v_0^k, k \in \mathcal{K}\}$ and $\{v^k, k \in \mathcal{K}\}$ are all bounded.

(iii) *There exists a positive integer* k_0 *such that* $\alpha_k = \alpha_{k_0} \equiv \alpha, \forall\, k \geqslant k_0$.

Proof. (i) Result (i) follows immediately from Proposition 4.2.

(ii) By (SLE_1)(2.22), we obtain

$$\begin{pmatrix} \mathrm{d}s_0^k \\ \pi_0^k \end{pmatrix} = -\begin{pmatrix} H_k & G_k^{\mathrm{T}} \\ G_k & 0 \end{pmatrix}^{-1} \begin{pmatrix} \nabla f(s^k) \\ h(z^k, \mu_k) \end{pmatrix}. \tag{4.4}$$

Since $\nabla f(s^k) \xrightarrow{\mathcal{K}} \nabla f(s^*)$, $h(z^k, \mu_k) \xrightarrow{\mathcal{K}} h(z^*, 0)$, so there exists a constant $\tilde{c} > 0$ such that

$$\left\| \begin{pmatrix} \nabla f(s^k) \\ h(z^k, \mu_k) \end{pmatrix} \right\| \leqslant \tilde{c}$$

which together with result (i) gives that $\{\mathrm{d}s_0^k, k \in \mathcal{K}\}$ and $\{\pi_0^k, k \in \mathcal{K}\}$ are bounded, furthermore, we obtain from (3.3) that $\{\mathrm{d}z_0^k, k \in \mathcal{K}\}$, $\{\lambda_0^k, k \in \mathcal{K}\}$ and $\{v_0^k, k \in \mathcal{K}\}$ are also bounded.

Similarly, we can verify that $\{\mathrm{d}z^k, k \in \mathcal{K}\}$, $\{\lambda^k, k \in \mathcal{K}\}$, $\{v^k, k \in \mathcal{K}\}$ are bounded, furthermore, we obtain $\{\widehat{\mathrm{d}z}^k,\ k \in \mathcal{K}\}$ is bounded from Step 3 in algorithm A.

(iii) Suppose by contradiction that this conclusion is not true, then, in view of the updating formula (3.11), there exists an infinite subset $\{k_i\}$ such that

$$\alpha_{k_i-1} < \xi_{k_i} + \delta,\ \alpha_{k_i} = \max\{\xi_{k_i} + \delta, \alpha_{k_i-1} + 2\delta\} \geqslant \alpha_{k_i-1} + 2\delta, \forall i.$$

On the other hand, from the updating formula (3.11), we know that the whole sequence $\{\alpha_k\}$ is monotonically nondecreasing, so $\alpha_{k_i-1} \geqslant \alpha_{k_{(i-1)}}$ since $k_i - 1 \geqslant k_{(i-1)}$. Therefore, one has

$$\alpha_{k_{(i-1)}} \leqslant \alpha_{k_i-1} < \xi_{k_i} + \delta, \ \ \alpha_{k_i} \geqslant \alpha_{k_i-1} + 2\delta, \ \ \forall i.$$

The second inequalities above show that $\lim_{i\to\infty} \alpha_{k_i} = +\infty$. But, from the first equalities above and result (ii), we know that

$$\lim_{i\to\infty} \alpha_{k_{(i-1)}} \leqslant \sup_i \{\xi_{k_i}\} + \delta < +\infty,$$

this contradicts $\lim_{i\to\infty} \alpha_{k_i} = +\infty$. So result (iii) is true. □

Based on Lemma 4.3, in the remainder of this paper we suppose, without loss of generality, that $\alpha_k \equiv \alpha$ for all k.

LEMMA 4.4. *For any* $\delta_1 > \delta_2 > 0$ *and* $(a, b) \in \Re^2$, *the following inequality holds*:

$$|\phi(a, b, \delta_2)| \leqslant |\phi(a, b, \delta_1)| + \frac{\sqrt{2-\lambda}\,\delta_1}{2\sqrt{\delta_2}}. \tag{4.5}$$

Proof. Using the mean value theorem, we know that there exists $\bar{\delta} \in (\delta_2, \delta_1)$ such that

$$\begin{aligned} |\phi(a, b, \delta_2)| &= |\phi(a, b, \delta_1) + \phi'_\mu(a, b, \bar{\delta})(\delta_2 - \delta_1)|, \\ &= |\phi(a, b, \delta_1)| + \frac{(2-\lambda)(\delta_1 - \delta_2)}{2\sqrt{a^2 + b^2 + \lambda ab + (2-\lambda)\bar{\delta}}} \\ &\leqslant |\phi(a, b, \delta_1)| + \frac{(2-\lambda)(\delta_1 - \delta_2)}{2\sqrt{(a + \frac{\lambda b}{2})^2 + \frac{(4-\lambda^2)b^2}{4} + (2-\lambda)\bar{\delta}}} \\ &\leqslant |\phi(a, b, \delta_1)| + \frac{\sqrt{2-\lambda}\,\delta_1}{2\sqrt{\delta_2}}. \end{aligned}$$

□

LEMMA 4.5. *Suppose that sequences* $\{v_k\}$ *and* $\{\gamma_k\}$ *of scalars satisfy*

$$v_k \geqslant 0, \sum_{k=1}^{\infty} v_k < \infty, \ \ \gamma_{k+1} \leqslant \gamma_k + v_k, k = 1, 2, \ldots. \tag{4.6}$$

Then (i) The sequence $\{\gamma_k\}$ *is bounded from above, i.e.* $\overline{\lim}_{k\to\infty} \gamma_k < +\infty$.

(ii) *The entire sequence* $\{\gamma_k\}$ *is convergent (including a finite limit or an infinite limit).*

Proof. We first have from (4.6)

$$\gamma_{k+1} \leqslant \gamma_k + v_k \leqslant \gamma_{k-1} + v_{k-1} + v_k \leqslant \cdots \leqslant \gamma_1 + \sum_{i=1}^{k} v_i,$$

which together with $\sum_{k=1}^{\infty} v_k < \infty$ shows that result (i) holds true.

In order to prove result (ii), let

$$\bar{a} = \overline{\lim_{k\to\infty}} \gamma_k = \lim_{k\in\overline{K}} \gamma_k, \underline{a} = \underline{\lim_{k\to\infty}} \gamma_k = \lim_{k\in\underline{K}} \gamma_k,$$

and then it is sufficient to verify $\bar{a} \leqslant \underline{a}$. For convenience, we assume that both $\bar{a}$ and $\underline{a}$ are finite real numbers. Let $\varepsilon > 0$ be any given real number and small enough, we know that, from (4.3) and $\sum_{k=1}^{\infty} \gamma_k < \infty$, there exist positive integers $\overline{N} \in \overline{K}$, $\underline{N} \in \underline{K}$ and N such that

$$-\varepsilon < \gamma_k - \bar{a} < \varepsilon, k \in \overline{K}, \forall k > \overline{N}, \tag{4.7}$$

$$-\varepsilon < \gamma_k - \underline{a} < \varepsilon, k \in \underline{K}, \forall k > \underline{N}, \tag{4.8}$$

$$\sum_{i=0}^{t-1} v_{k+i} < \varepsilon, \forall k > N, t = 1, 2, \ldots$$

On the other hand, for any given $k \in \underline{K}$ and $k > \max\{\overline{N}, \underline{N}, N\}$, since $\overline{K}$ is an infinite subset of $\{1, 2, \ldots\}$, there exists at least an integer t such that $k + t \in \overline{K}$, moreover we have from (4.6)

$$\gamma_{k+t} \leqslant \gamma_{k+t-1} + v_{k+t-1} \leqslant \cdots \leqslant \gamma_k + \sum_{i=0}^{t-1} v_{k+i},$$

which together with (4.7) and (4.8) gives

$$\bar{a} - \varepsilon \leqslant \underline{a} + \varepsilon + \varepsilon,$$

So we can conclude that $\bar{a} \leqslant \underline{a}$ since $\varepsilon > 0$ is sufficiently small. The proof is completed. □

LEMMA 4.6. *Suppose that assumptions* (A1)–(A3) *hold, then the sequences* $\{\theta(z^k, \alpha, \mu_k)\}$ *and* $\{\theta(z^{k+1}, \alpha, \mu_k)\}$ *are both convergent and have the same limit.*

Proof. In view of $\lim_{k\to\infty} \frac{\mu_{k+1}}{\mu_k^{\gamma}} = \bar{\eta}$ in (3.2), we have $\sqrt{2}\sqrt{\mu_{k+1}} > \sqrt{\bar{\eta}\mu_k^{\gamma}}$, furthermore, by Lemma 4.4, we have

$$|\phi(t_i^{k+1}, \mu_{k+1})| \leqslant |\phi(t_i^{k+1}, \mu_k)| + \frac{\sqrt{2-\lambda}\mu_k}{2\sqrt{\mu_{k+1}}} \leqslant |\phi(t_i^{k+1}, \mu_k)| + \hat{\delta}\mu_k^{1-\frac{\gamma}{2}},$$

where $\widehat{\delta} = \frac{\sqrt{2-\lambda}}{\sqrt{\bar{\eta}}}$. Combining this inequality with (3.1), we have

$$\theta(z^{k+1}, \alpha, \mu_{k+1}) \leqslant \theta(z^{k+1}, \alpha, \mu_k) + m\alpha\hat{\delta}\mu_k^{1-\frac{\gamma}{2}}. \tag{4.9}$$

In view of the step size search (3.12) and Lemma 3.3(ii), we obtain

$$\theta(z^{k+1}, \alpha, \mu_k) \leqslant \theta(z^k, \alpha, \mu_k) + \sigma\tau\psi(z^k, \mathrm{d}z^k, \alpha, \mu_k) < \theta(z^k, \alpha, \mu_k), \tag{4.10}$$

and which together with (4.9) gives

$$\theta(z^{k+1}, \alpha, \mu_{k+1}) \leq \theta(z^k, \alpha, \mu_k) + m\alpha\widehat{\delta}\mu_k^{1-\frac{\gamma}{2}}. \tag{4.11}$$

On the other hand, from $\lim_{k\to\infty} \frac{\mu_{k+1}}{\mu_k^{\gamma}} = \bar{\eta}$, we know that $\sum_{k=0}^{\infty} \mu_k^{1-\frac{\gamma}{2}}$ is convergent, so $\{\theta(z^k, \alpha, \mu_k)\}$ is convergent from Lemma 4.4 and the inequality above. Again, by (4.9) and (4.10), we get

$$\theta(z^{k+1}, \alpha, \mu_{k+1}) - m\alpha\hat{\delta}\mu_k^{1-\frac{\gamma}{2}} \leqslant \theta(z^{k+1}, \alpha, \mu_k) \leqslant \theta(z^k, \alpha, \mu_k).$$

Passing to the limit $k \to \infty$, we have that $\lim_{k\to\infty} \theta(z^{k+1}, \alpha, \mu_k) = \lim_{k\to\infty} \theta(z^k, \alpha, \mu_k)$.The proof is completed. □

LEMMA 4.7. *Suppose that assumptions* (A1)–(A3) *hold, then*

$$\psi(z^k, \mathrm{d}z^k, \alpha, \mu_k) = 0, \quad k \in \mathcal{K}.$$

Proof. Suppose, by contradiction, that there exists a constant $\bar{c} > 0$ such that

$$|\psi(z^k, \mathrm{d}z^k, \alpha, \mu_k)| \geqslant \bar{c}, \quad \forall\ k \in \mathcal{K},$$

which together with Lemma 3.3(ii) gives

$$\psi(z^k, \mathrm{d}z^k, \alpha, \mu_k) \leqslant -\bar{c}, \quad \forall\ k \in \mathcal{K}. \tag{4.12}$$

Next we will show that there exists $\underline{\tau} > 0$, such that $\tau_k \geqslant \underline{\tau}, \forall\ k \in \mathcal{K}$. From (3.20) and (3.22), we have

$$\theta(z^k + \tau \mathrm{d}z^k + \tau^2(\widehat{\mathrm{d}z}^k - \mathrm{d}z^k), \alpha, \mu_k) - \theta(z^k, \alpha, \mu_k) \leqslant \tau\psi(z^k, \mathrm{d}z^k, \alpha, \mu_k) + o(\tau),$$

Which together with the boundedness of $\{\mathrm{d}z^k, k \in \mathcal{K}\}$ and $\{\widehat{\mathrm{d}z}^k - \mathrm{d}z^k, k \in \mathcal{K}\}$ shows that there exists a constant $\bar{\tau} > 0$ such that

$$\theta(z^k + \tau \mathrm{d}z^k + \tau^2(\widehat{\mathrm{d}z}^k - \mathrm{d}z^k), \alpha, \mu_k) \leqslant \theta(z^k, \alpha, \mu_k) + \sigma\tau\psi(z^k, \mathrm{d}z^k, \alpha, \mu_k), \tag{4.13}$$

for all $k \in \mathcal{K}$ and $\forall \tau \in [0, \bar{\tau}]$. Therefore according to the method of computating the step size τ_k, there exists $\underline{\tau} > 0$ such that $\tau_k \geqslant \underline{\tau}$, $\forall k \in \mathcal{K}$ so by (4.12), we can rewrite (4.13) as

$$\theta(z^{k+1}, \alpha, \mu_k) \leqslant \theta(z^k, \alpha, \mu_k) - \sigma \underline{\tau} \bar{c}.$$

Passing to the limit $k \to \infty$ $(k \in \mathcal{K})$ and using Lemma 4.6, we obtain $-\sigma \underline{\tau} \bar{c} \geqslant 0$, which contradicts "$\sigma > 0$, $\underline{\tau} > 0$, $\bar{c} > 0$", so $\lim_{\mathcal{K}} \psi(z^k, \mathrm{d}z^k, \alpha, \mu_k) = 0$. □

Now we are ready to state and prove the global convergence theorem for algorithm A.

THEOREM 4.8. *Suppose that the stated assumption* (A1)–(A3) *hold, then each accumulation point of the sequence* $\{z^k\}$ *generated by algorithm A is a stationary point of* (*LCP*)(1.1).

Proof. Let z^* be a given accumulation point of $\{z^k\}$ and subsequence $\mathcal{K}$ ensure (4.1) holds. By (3.19) and assumption (A2), we obtain

$$\psi(z^k, \mathrm{d}z^k, \alpha, \mu_k) \leqslant -\frac{1}{2} c_1 \|\mathrm{d}s_0^k\|^2 - \sum_{j \in J, \lambda_{0,j}^k < 0} (\lambda_{0,j}^k)^4 < 0,$$

From Lemma 4.7, we have $\lim_{\mathcal{K}} \psi(z^k, \mathrm{d}z^k, \alpha, \mu_k) = 0$, so

$$\|\mathrm{d}s_0^k\| \to 0, \sum_{j \in J, \lambda_{0,j}^k < 0} (\lambda_{0,j}^k)^4 \to 0, (k \xrightarrow{\mathcal{K}} \infty) \tag{4.14}$$

Notice that $\{\lambda_{0,j}^k, k \in \mathcal{K}\}$ and $\{v_{0,i}^k,\ k \in \mathcal{K}\}$ are bounded from Lemma 4.3(ii), without loss of generality, we suppose that $\lambda_{0,j}^k \xrightarrow{\mathcal{K}} \tilde{\lambda}_j^*$, $v_{0,i}^k \xrightarrow{\mathcal{K}} \tilde{v}_i^*$. Hence we can verify by (4.14) that $\tilde{\lambda}_j^* \geqslant 0$, $\forall j \in J$. If not, then there exists a $\widetilde{\lambda}_t^* < 0$, $t \in J$, so $\lambda_{0,t}^k < 0$, $t \in J$, furthermore,

$$\sum_{j \in J, \lambda_{0,j}^k < 0} (\lambda_{0,j}^k)^4 \geqslant (\lambda_{0,t}^k)^4 \longrightarrow (\tilde{\lambda}_t^*)^4 > 0,$$

which is in contradiction with (4.14). So passing to $k \xrightarrow{\mathcal{K}} \infty$ in (SLE_1)(2.22), we obtain

$$\begin{aligned} &\nabla f(s^*) + \begin{pmatrix} N^{\mathrm{T}} \Gamma_w^* \\ M^{\mathrm{T}} \Gamma_w^* + \Gamma_y^* \end{pmatrix} \tilde{v}^* + \begin{pmatrix} A_J^{\mathrm{T}} \\ 0 \end{pmatrix} (\widetilde{\lambda}^*)_J = 0, \\ &(\widetilde{\lambda}^*)_J \geqslant 0, \quad A_J x^* - b_J = 0, \quad \Phi(y^*, w^*, 0) = 0, \end{aligned} \tag{4.15}$$

In view of $I(x^*) \subseteq J$, it follows from above (4.15) that $\varphi(x^*) = 0$, furthermore, $Ax^* - b \leqslant 0$. Passing to the limit $k \to \infty$ in $w^k = Nx^k + My^k + q$, we get $w^* = Nx^* + My^* + q$, so $z^* = (x^*, y^*, w^*) \in X$. Now we define

$$v_i^* = \begin{cases} \frac{2-\lambda}{2y_i^*}\widetilde{v}_i^*, & \text{if } i \in I_w(z^*) \overset{\text{def}}{=} \{i \in L_2 :\ w_i^* = 0\} \\ \frac{2-\lambda}{2w_i^*}\widetilde{v}_i^*, & \text{if } i \in I_y(z^*) \overset{\text{def}}{=} \{i \in L_2 :\ y_i^* = 0\}. \end{cases} \tag{4.16}$$
$$\lambda_j^* = \widetilde{\lambda}_j^*, j \in J, \lambda_j^* = 0,\ j \notin J.$$

Hence we obtain from (4.15)

$$\nabla f(s^*) + \begin{pmatrix} N^{\mathrm{T}} Y^* \\ M^{\mathrm{T}} Y^* + W^* \end{pmatrix} v^* + \begin{pmatrix} A^{\mathrm{T}} \\ 0 \end{pmatrix} \lambda^* = 0,$$
$$\lambda^* \geqslant 0, \quad (Ax^* - b)^{\mathrm{T}} \lambda^* = 0,$$

where $Y^* = \mathrm{diag}(y_i^*, i = 1, 2, \ldots, m)$, $W^* = \mathrm{diag}(w_i^*, i = 1, 2, \ldots, m)$. So the theorem holds from the equivalent condition (2.4) of a stationary point. □

Denote

$$I_y^* = I_y(z^*) \overset{\text{def}}{=} \{i \in L_2 :\ y_i^* = 0\}, \quad I_w^* = I_w(z^*) \overset{\text{def}}{=} \{i \in L_2 : w_i^* = 0\},$$

$$L_{1+}^* = \{j \in L_1 : \lambda_j^* > 0\}, \quad I^* = I(x^*), \quad A_{L_{1+}^*} = (a_j, j \in L_{1+}^*).$$

In order to prove the strong and superlinear convergence, the following assumptions are necessary:

(A4) Function f is twice continuously differentiable.

(A5) (i) There exists an accumulation point z^* of $\{z^k\}$, so from Theorem 4.8, there exist multipliers (λ^*, v^*) such that the stationary point pair (z^*, λ^*, v^*) satisfying (2.3), suppose that second order sufficient conditions as follows hold true:

$$(\mathrm{d}s)^{\mathrm{T}} \nabla^2 f(x^*, y^*) \mathrm{d}s > 0, \quad \forall \mathrm{d}s \in \Omega,$$

where

$$\Omega \overset{\text{def}}{=} \{0 \neq \mathrm{d}s \in \Re^{n+m} : N_{I_w^*} \mathrm{d}x + (M_{I_w^* I_w^*}) \mathrm{d}y_{I_w^*} = 0, A_{L_{1+}^*} \mathrm{d}x = 0, \mathrm{d}y_{I_y^*} = 0\}. \tag{4.17}$$

(ii) (Upper level) Strict complementary conditions hold, i.e. $\lambda_j^* > 0 (\forall j \in I(x^*))$.

By (2.20), (4.3), (2.11), (2.12) and Proposition 2.3, it is easy to prove the following result holds.

LEMMA 4.9. *Assumption* (A5)(i) *is equivalent to the following statement*:

$$\mathrm{d}z^{\mathrm{T}} H(z^*, \tilde{v}^*, 0) \mathrm{d}z > 0, \quad \forall \mathrm{d}z \in \Omega, \tag{4.18}$$

where $\Omega^+ \overset{\text{def}}{=} \{\mathrm{d}z \neq 0 : A_{L_{1+}^*} \mathrm{d}x = 0, (N \quad M) \mathrm{d}s = \mathrm{d}w, \begin{pmatrix} \Gamma_y^* & \Gamma_w^* \end{pmatrix} \mathrm{d}t = 0\}$.

LEMMA 4.10. *Suppose that assumptions* (A1)–(A5) *hold, then*
(i) $\lim_{k\to\infty} dz_0^k = \lim_{k\to\infty} dz^k = 0$.
(ii) $\lim_{k\to\infty} \|z^{k+1} - z^k\| = 0$.

Proof. From (4.14) and assumption (A5), we know for any infinite subsequence $\mathcal{K} \subseteq \{1, 2, 3, \ldots\}$, that there exists subsequence $\mathcal{K}' \subseteq \mathcal{K}$ such that $\lim_{k\in\mathcal{K}'} ds_0^k = 0$, so $\lim_{k\to\infty} ds_0^k = 0$. On the other hand, $\lim_{k\to\infty} \eta_k = 0$ follows from (3.5) and $\tilde{\lambda}_j^* \geqslant 0$, $\forall j \in J$, hence we have $\lim_{k\to\infty} ds^k = 0$ from $ds^k = ds_0^k + B_k\eta_k$ in (3.15), combining (3.3) and (3.6), we have

$$\lim_{k\to\infty} dz_0^k = \lim_{k\to\infty} (ds_0^k, dw_0^k) = 0, \quad \lim_{k\to\infty} dz^k = \lim_{k\to\infty} (ds^k, dw^k) = 0.$$

Furthermore, since

$$\begin{aligned}\|z^{k+1} - z^k\| =& \|\tau_k dz^k + \tau_k^2(\widehat{dz}^k - dz^k)\| \leqslant \tau_k\|dz^k\| + \tau_k^2\|\widehat{dz}^k - dz^k\| \\ &\leqslant \tau_k\|dz^k\| + \tau_k^2\|dz^k\| = \tau_k(1+\tau_k)\|dz^k\|\end{aligned}$$

Therefore $\lim_{k\to\infty} \|z^{k+1} - z^k\| = 0$. □

THEOREM 4.11. *Suppose that assumptions* (A1)–(A5) *hold, z^* is an accumulation point of sequence $\{z^k\}$ produced by algorithm A, then $\lim_{k\to\infty} z^k = z^*$, i.e. algorithm A is strongly convergent.*

Proof. From Theorem 4.8, Proposition 2.3 and the given assumptions, we know that $(z^*, \tilde{\omega}^*)$ (where $\tilde{\omega}^* = (\lambda^*, u^*, \bar{v}^* = \tilde{v}^*)$) is a KKT pair of (NLP$_\mu$) (2.14) for $\mu = 0$. On the other hand, by using the given assumptions and Proposition 4.2, it is not difficult to see that the second order sufficient conditions and the linear independence constraint qualification (LICQ) hold for (NLP$_\mu$) (2.14) with $\mu = 0$ at $(z^*, \tilde{\omega}^*)$. Thus we can conclude z^* is an isolated KKT point of (NLP$_\mu$) (2.14) for $\mu = 0$ (see Theorem 1.2.5 in [16]). Furthermore, we know from Theorem 4.8 that z^* is an isolated accumulation point of $\{z^k\}$. Therefore, combining this conclusion with Lemma 4.10(ii) and Theorem 1.1.5 in [9], we obtain $\lim_{k\to\infty} z^k = z^*$.
□

5. Superlinear Convergence

In this section, we will analyze and verify the superlinear convergence of algorithm A. For this goal, we first give some lemmas.

LEMMA 5.1 [6]. *Suppose that assumptions* (A1)–(A5) *hold, then*

$J_k \equiv I(x^*) = I^*$, *for k large enough.*

Based on Lemma 5.1, for convenience, let $I^* = I(x^*)$ throughout the remainder of the paper.

LEMMA 5.2. *Suppose that assumptions* (A1)–(A5) *hold, then* $\lim_{k\to\infty} \lambda_0^k = \lambda^*$, $\lim_{k\to\infty} v_0^k = \tilde{v}^*$, *and* (λ^*, v^*) (*where* v^* *and* $\tilde{v}^*$ *satisfy* (2.15)) *are the multipliers associated with stationary point* z^* *of* (LCP)(1.1) *satisfying* (2.3), *and* $(\lambda^*, u^*, \tilde{v}^*) = (\lambda^*, Y^* v^*, \tilde{v}^*)$ *are KKT multipliers associated with KKT point* z^* *of* (NLP_0) (2.14).

Proof. From Lemma 5.1 and (SLE_1)(2.22), we have

$$H_k \mathrm{d}s_0^k + \nabla f(s^k) + G_k^T \begin{pmatrix} \lambda_{0,I^*}^k \\ v_0^k \end{pmatrix} = 0,$$

which together with Proposition 4.2, Lemma 4.10 and Theorem 4.11 gives

$$\begin{pmatrix} \lambda_{0,I^*}^k \\ v_0^k \end{pmatrix} = -(G_k G_k^T)^{-1} G_k (H_k \mathrm{d}s_0^k + \nabla f(s^k)).$$

Passing to the limit $k \to \infty$, and notice that $\lim_{k\to\infty} \mathrm{d}s_0^k = 0$, we get

$$\begin{pmatrix} \lambda_{0,I^*}^k \\ v_0^k \end{pmatrix} \longrightarrow -(G_* G_*^T)^{-1} G_* \nabla f(s^*) \stackrel{\mathrm{def}}{=} \begin{pmatrix} \lambda_{I^*}^* \\ \tilde{v}^* \end{pmatrix}.$$

On the other hand, from the strict complementarity condition, we have $\lambda_j^* = 0$ for $j \in L_1 \backslash I^*$, so if we set $\lambda_{0,L_1\backslash I(x^*)}^k = 0$, then

$$\begin{pmatrix} \lambda_0^k \\ v_0^k \end{pmatrix} \to \begin{pmatrix} \lambda^* \\ \tilde{v}^* \end{pmatrix}, \text{i.e.} \lim_{k\to\infty} \lambda_0^k = \lambda^*, \lim_{k\to\infty} v_0^k = \tilde{v}^*.$$

furthermore, the remaining conclusions follow from Theorem 4.8 and Theorem 4.11. □

By Lemma 5.1, assumption A(ii), (3.5), Lemma 5.2 and Step 3(ii), we see that the following result is true:

LEMMA 5.3. *Suppose that assumptions* (A1)–(A5) *hold, then* $\eta_k \equiv 0$ *for all sufficiently large* k, *moreover,* $\mathrm{d}z_0^k = \mathrm{dz}^{\mathrm{k}}$.

LEMMA 5.4. *Suppose that assumption* (A1)–(A5) *hold, then*

$$\|\widehat{\mathrm{d}z}^k - \mathrm{d}z^k\| = O(\|\mathrm{d}z_0^k\|^2).$$

Proof. By Lemma 5.3, we have $\mathrm{d}s^k = \mathrm{d}s_0^k$ for k enough large, so from (SLE_1)(2.22) and (SLE_3)(3.7), we obtain

$$\begin{pmatrix} H_k & G_k^T \\ G_k & 0 \end{pmatrix} \begin{pmatrix} \widehat{ds}^k - ds_0^k \\ \widehat{\pi}^k - \pi_0^k \end{pmatrix} = -\begin{pmatrix} 0 \\ h(z^k + dz_0^k, \mu_k) \end{pmatrix}. \tag{5.1}$$

Again, by (SLE_1)(2.22), we have

$$A_{J_k}(x^k + dx_0^k) - b_{J_k} = A_{J_k}x^k - b_{J_k} + A_{J_k}dx_0^k = 0,$$

and using Taylor expansion and (SLE_1)(2.22), one knows

$$\begin{aligned}\Phi(t^k + dt_0^k, \mu_k) &= \Phi(t^k, \mu_k) + \nabla\Phi(t^k, \mu_k)dt_0^k + O(\|dt_0^k\|^2) \\ &= O(\|dt_0^k\|^2) = O(\|ds_0^k\|^2) = O(\|dz_0^k\|^2).\end{aligned}$$

So (5.1) together with (2.24) gives

$$\begin{pmatrix} H_k & G_k^T \\ G_k & 0 \end{pmatrix} \begin{pmatrix} \widehat{ds}^k - ds_0^k \\ \hat{\pi}^k - \pi_0^k \end{pmatrix} = \begin{pmatrix} 0 \\ O(\|dz_0^k\|^2) \end{pmatrix}.$$

Thus

$$\begin{pmatrix} \widehat{ds}^k - ds_0^k \\ \widehat{\pi}^k - \pi_0^k \end{pmatrix} = \begin{pmatrix} H_k & G_k^T \\ G_k & 0 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ O(\|dz_0^k\|^2) \end{pmatrix}$$

follows from Lemma 2.5. Hence $\|\widehat{ds}^k - ds_0^k\| = O(\|dz_0^k\|^2)$, furthermore,

$$\|\widehat{dz}^k - dz_0^k\| = O(\|dz_0^k\|^2). \qquad \square$$

In order to prove superlinear convergence of algorithm A, the following assumption is necessary:

(A6) $\quad \|(H(z^k, v^k, \mu_k) - B_k)dz^k\| = o(\|dz^k\|)$

THEOREM 5.5. *Suppose that assumptions* (A1)–(A6) *hold, then* $\tau_k \equiv 1$ *for all sufficiently large* k.

Proof. From (3.12), it is sufficient to verify the following inequality holds:

$$\theta(z^k + \widehat{dz}^k, \alpha, \mu_k) - \theta(z^k, \alpha, \mu_k) \leqslant \sigma\psi(z^k, dz^k, \alpha, \mu_k)$$

Using (3.1) and Taylor expansion, we have

$$\begin{aligned}&\theta(z^k + \widehat{dz}^k, \alpha, \mu_k) - \theta(z^k, \alpha, \mu_k) \\ &= \nabla f(s^k)^T\widehat{ds}^k + \frac{1}{2}(\widehat{ds}^k)^T\nabla^2 f(s^k)\widehat{ds}^k + o(\|\widehat{ds^k}\|^2) - \alpha\varphi(x^k) - \alpha\|\Phi(t^k, \mu_k)\|_1 \\ &+ \alpha\varphi(x^k + \widehat{dx}^k) + \alpha\|\Phi(t^k + \widehat{dt}^k, \mu_k)\|_1.\end{aligned} \tag{5.2}$$

In view of $\nabla f(s^k)^T\widehat{ds}^k = \nabla f(s^k)^T ds^k + \nabla f(s^k)^T(\widehat{ds}^k - ds^k)$, it follows from Lemma 5.4 that

$$\frac{1}{2}(\widehat{\mathrm{d}s^k})^T\nabla^2 f(s^k)\widehat{\mathrm{d}s^k} = \frac{1}{2}(\mathrm{d}s^k)^T\nabla^2 f(s^k)\mathrm{d}s^k + o(\|\mathrm{d}z_0^k\|^2)$$

since

$$o(\|\widehat{\mathrm{d}s^k}\|^2) = o(\|\mathrm{d}s^k + \widehat{\mathrm{d}s^k} - \mathrm{d}s^k\|^2) \leqslant o(\|\mathrm{d}s^k\|^2) + o(\|\widehat{\mathrm{d}s^k} - \mathrm{d}s^k\|^2) = o(\|\mathrm{d}z_0^k\|^2),$$

so by above results and (3.13), (5.2) is rewritten as

$$\begin{aligned}&\theta(z^k + \widehat{\mathrm{d}z^k}, \alpha, \mu_k) - \theta(z^k, \alpha, \mu_k)\\ &\quad\leqslant \psi(z^k, \mathrm{d}z^k, \alpha, \mu_k) - \frac{1}{2}(\mathrm{d}s_0^k)^T H_k \mathrm{d}s_0^k + \nabla f(s^k)^T(\widehat{\mathrm{d}s^k} - \mathrm{d}s^k) + \frac{1}{2}(\mathrm{d}s_0^k)^T\nabla^2 f(s^k)\mathrm{d}s_0^k\\ &\quad + \alpha\max\{0; a_j(x^k + \widehat{\mathrm{d}x^k}) - b_j, j\in L_1\} + \alpha\|\Phi(t^k + \widehat{\mathrm{d}t^k}, \mu_k)\|_1 + o(\|\mathrm{d}z_0^k\|^2).\end{aligned} \tag{5.3}$$

For $j \notin J_k = I^*$, $a_j x^k - b_j \leqslant 0$ follows from $a_j x^* - b_j < 0$. Since

$$\|\widehat{\mathrm{d}z^k}\| = \|\mathrm{d}z^k + \widehat{\mathrm{d}z^k} - \mathrm{d}z^k\| \leqslant \|\mathrm{d}z^k\| + \|\widehat{\mathrm{d}z^k} - \mathrm{d}z^k\|,$$

so from Lemma 4.10 and Lemma 5.4, we have $\lim_{k\to\infty}\|\widehat{\mathrm{d}z^k}\| = 0$, moreover, $\lim\limits_{k\to\infty}\|\widehat{\mathrm{d}x^k}\| = 0$, therefore

$$a_j(x^k + \widehat{\mathrm{d}x^k}) - b_j = a_j x^k - b_j + a_j\widehat{\mathrm{d}x^k} \leqslant 0,\ j \notin J_k. \tag{5.4}$$

For $j \in J_k$, by (SLE_2)(3.4), (SLE_3)(3.7) and Lemma 5.3, we obtain

$$\begin{aligned}&a_j(\widehat{\mathrm{d}x^k} - \mathrm{d}x^k) = -(a_j(x^k + \mathrm{d}x^k) - b_j),\\ &\nabla\phi(t_i^k, \mu_k)^T(\widehat{\mathrm{d}t^k} - \mathrm{d}t^k) = -\phi(t_i^k + \mathrm{d}t^k, \mu_k).\end{aligned} \tag{5.5}$$

Hence

$$a_j(x^k + \widehat{\mathrm{d}x^k}) - b_j = a_j(x^k + \mathrm{d}x^k) - b_j + a_j(\widehat{\mathrm{d}x^k} - \mathrm{d}x^k) = 0, j \in J_k. \tag{5.6}$$

In view of (5.4) and (5.6), we have

$$\max\{0; a_j(x^k + \widehat{\mathrm{d}x^k}) - b_j,\ j \in L_1\} = 0. \tag{5.7}$$

Using Taylor expansion, we have

$$\begin{aligned}\phi(t_i^k + \widehat{\mathrm{d}t^k}, \mu_k) &= \phi(t_i^k + \mathrm{d}t^k, \mu_k) + \nabla\phi(t^k + \mathrm{d}t^k, \mu_k)^{\mathrm{T}}(\widehat{\mathrm{d}t^k} - \mathrm{d}t^k)\\ &\quad + O(\|\widehat{\mathrm{d}t^k} - \mathrm{d}t^k\|^2)\\ &= \phi(t_i^k + \mathrm{d}t^k, \mu_k) + (\nabla\phi(t_i^k, \mu_k) + O(\|\mathrm{d}t^k\|))^{\mathrm{T}}(\widehat{\mathrm{d}t^k} - \mathrm{d}t^k) + O(\|\widehat{\mathrm{d}t^k} - \mathrm{d}t^k\|^2)\\ &= \phi(t_i^k + \mathrm{d}t^k, \mu_k) + \nabla\phi(t_i^k, \mu_k)^{\mathrm{T}}(\widehat{\mathrm{d}t^k} - \mathrm{d}t^k) + O(\|\mathrm{d}t^k\|)(\widehat{\mathrm{d}t^k} - \mathrm{d}t^k)\\ &\quad + O(\|\widehat{\mathrm{d}t^k} - \mathrm{d}t^k\|^2)\\ &= O(\|\mathrm{d}t^k\|)(\widehat{\mathrm{d}t^k} - \mathrm{d}t^k) + O(\|\widehat{\mathrm{d}t^k} - \mathrm{d}t^k\|^2) = o(\|\mathrm{d}z_0^k\|^2).\end{aligned} \tag{5.8}$$

On the other hand, for $j \in J_k$,

$$a_j(x^k + \widehat{\mathrm{d}x^k}) - b_j = a_j x^k - b_j + a_j \mathrm{d}x^k + a_j(\widehat{\mathrm{d}x^k} - dx^k) = a_j(\widehat{\mathrm{d}x^k} - \mathrm{d}x^k),$$

which together with (5.6) gives

$$a_j(\widehat{\mathrm{d}x^k} - \mathrm{d}x^k) = 0. \tag{5.9}$$

Denote $\phi_i(t, \mu) = \phi(t_i, \mu)$, then by Taylor expansion, we have

$$\begin{aligned}
&\phi_i(t^k + \widehat{\mathrm{d}t}^k, \mu_k) \\
&= \phi_i(t^k, \mu_k) + \nabla\phi_i(t^k, \mu_k)\widehat{\mathrm{d}t}^k + \frac{1}{2}(\widehat{\mathrm{d}t}^k)^T \nabla^2 \phi_i(t^k, \mu_k)\widehat{\mathrm{d}t}^k + o(\|\widehat{\mathrm{d}t}^k\|^2) \\
&= \phi_i(t^k, \mu_k) + \nabla\phi_i(t^k, \mu_k)dt^k + \nabla\phi_i(t^k, \mu_k)(\widehat{\mathrm{d}t}^k - dt^k) \\
&\quad + \frac{1}{2}(\widehat{\mathrm{d}t}^k)^T \nabla^2 \phi_i(t^k, \mu_k)\widehat{\mathrm{d}t}^k + o(\|\widehat{\mathrm{d}t}^k\|^2) \\
&= \nabla\phi_i(t^k, \mu_k)(\widehat{\mathrm{d}t}^k - \mathrm{d}t^k) + \frac{1}{2}(\mathrm{d}t_0^k)^T \nabla^2 \phi_i(t^k, \mu_k)\mathrm{d}t_0^k + o(\|\mathrm{d}_0^k\|^2),
\end{aligned}$$

which together with (5.8) shows that

$$\nabla\phi_i(t^k, \mu_k)(\widehat{\mathrm{d}t}^k - \mathrm{d}t^k) + \frac{1}{2}(\mathrm{d}t_0^k)^T \nabla^2 \phi_i(t^k, \mu_k)\mathrm{d}t_0^k = o(\|\mathrm{d}z_0^k\|^2). \tag{5.10}$$

By (SLE_1)(2.22), (5.9), (5.10) and Lemma 5.4, we have

$$\begin{aligned}
&\nabla f(s^k)^T(\widehat{\mathrm{d}s}^k - \mathrm{d}s^k) \\
&= -(\mathrm{d}s_0^k)^T H_k(\widehat{\mathrm{d}s}^k - \mathrm{d}s^k) - \sum_{j \in J_k} \lambda_{0,j}^k a_j(\widehat{\mathrm{d}x^k} - \mathrm{d}x^k) \\
&\quad - \sum_{i=1}^m v_{0,i}^k \nabla\phi_i(t^k, \mu_k)(\widehat{\mathrm{d}t}^k - \mathrm{d}t^k) \\
&= \frac{1}{2}\sum_{i=1}^m v_{0,i}^k (\mathrm{d}t_0^k)^T \nabla^2 \phi_i(t^k, \mu_k)\mathrm{d}t_0^k + o(\|\mathrm{d}z_0^k\|^2).
\end{aligned} \tag{5.11}$$

Substituting (5.7), (5.8) and (5.11) into (5.3), we get

$$\begin{aligned}
&\theta(z^k + \widehat{\mathrm{d}z}^k, \alpha, \mu_k) - \theta(z^k, \alpha, \mu_k) \leqslant \psi(z^k, \mathrm{d}z^k, \alpha, \mu_k) \\
&\quad + \frac{1}{2}\sum_{i=1}^m v_{0,i}^k (\mathrm{d}t_0^k)^T \nabla^2 \phi_i(t^k, \mu_k)\mathrm{d}t_0^k + \frac{1}{2}(\mathrm{d}s^k)^T \nabla^2 f(s^k)\mathrm{d}s^k - \frac{1}{2}(\mathrm{d}s_0^k)^T H_k \mathrm{d}s_0^k + o(\|\mathrm{d}z_0^k\|^2) \\
&= \psi(z^k, \mathrm{d}z^k, \alpha, \mu_k) + \frac{1}{2}\sum_{i=1}^m v_{0,i}^k (\mathrm{d}t_0^k)^T \nabla^2 \phi_i(t^k, \mu_k)\mathrm{d}t_0^k + \frac{1}{2}(\mathrm{d}s_0^k)^T \nabla^2 f(s^k)\mathrm{d}s_0^k \\
&\quad - \frac{1}{2}(\mathrm{d}s_0^k)^T H_k \mathrm{d}s_0^k + o(\|\mathrm{d}z_0^k\|^2).
\end{aligned}$$

In view of $dw^k = Ndx^k + Mdy^k$, (2.21), (2.23) and (2.18), we have

$$(\mathrm{d}s_0^k)^{\mathrm{T}}\nabla^2 f(s^k)\mathrm{d}s_0^k + \sum_{i=1}^{m} v_{0,i}^k(\mathrm{d}t_0^k)^T\nabla^2\phi_i(t^k,\mu_k)\mathrm{d}t_0^k - (\mathrm{d}s_0^k)^T H_k \mathrm{d}s_0^k = (\mathrm{d}z_0^k)^T(H(z^k,v_0^k,\mu_k) - B_k)\mathrm{d}z_0^k,$$

combining with the assumption (A6), the following result holds true:

$$\begin{aligned}\theta(z^k+\widehat{\mathrm{d}z}^k,\alpha,\mu_k)-\theta(z^k,\alpha,\mu_k) &\leqslant \psi(z^k,\mathrm{d}z^k,\mu_k) \\ &\quad + \frac{1}{2}(\mathrm{d}z_0^k)^T(H(z^k,v_0^k,\mu_k)-B_k)\mathrm{d}z_0^k + o(\|\mathrm{d}z_0^k\|^2) \\ &= \psi(z^k,\mathrm{d}z^k,\mu_k) + o(\|\mathrm{d}z_0^k\|^2) < \sigma\psi(z^k,\mathrm{d}z^k,\mu_k).\end{aligned}$$

This completes the proof □

Although the step size of the algorithm A identically equals 1 near the solution, (NLP_{μ_k}) is a series of approaching problems containing parameters, the convergent speed has something to do with the properties of $\{\mu_k\}$ and the superlinear convergence can not be obtained directly by means of the existing results. So in order to verify the superlinear convergence of algorithm A, denote

$$R_* = \begin{pmatrix} A_{I^*}^{\mathrm{T}} & N^{\mathrm{T}} & 0 \\ 0 & M^{\mathrm{T}} & \Gamma_y^* \\ 0 & -E_m & \Gamma_w^* \end{pmatrix}, \quad R(z,\mu_k) = \begin{pmatrix} A_{J_k}^{\mathrm{T}} & N^{\mathrm{T}} & 0 \\ 0 & M^{\mathrm{T}} & \Gamma(y,w,\mu_k) \\ 0 & -E_m & \Gamma(w,y,\mu_k) \end{pmatrix}, \tag{5.12}$$

$$R_k = R(z^k,\mu_k), \quad P_k = E_{n+2m} - R_k(R_k^{\mathrm{T}}R_k)^{-1}R_k^{\mathrm{T}}. \tag{5.13}$$

LEMMA 5.6. *Suppose that assumptions* (A1)–(A6) *hold, then for all sufficiently large k, the following statements are true:*

(i) *The matrixes R_* and R_k are full of column rank.*

(ii) *The matrix*

$$Q_k = \begin{pmatrix} P_k H(z^*,\tilde{v}^*,0) & R_k \\ R_k^T & 0 \end{pmatrix} \tag{5.14}$$

is nonsingular, furthermore, there exists a constant $\bar{c} > 0$ such that $\|Q_k^{-1}\| \leqslant \bar{c}$.

Proof. (i) By Proposition 4.2(ii), we know that $\Gamma_y^* + M^T\Gamma_w^*$ is nonsingular, therefore, the following matrices

$$\begin{pmatrix} M^{\mathrm{T}} & \Gamma_y^* + M^{\mathrm{T}}\Gamma_w^* \\ -E_m & 0 \end{pmatrix}, \quad \begin{pmatrix} M^{\mathrm{T}} & \Gamma_y^* \\ -E_m & \Gamma_w^* \end{pmatrix}$$

are all full of column rank. By assumption (A1)(ii), $A_{I^*}^T$ is full of column rank and so is R_*.

Similarly, we can prove the matrix R_k is also full of column rank.

(ii) In view of $J_k \equiv I(x^*)$, so part (ii) follows from Lemma 2.2.2 in [9]. □

THEOREM 5.7. *Suppose that assumptions* (A1)–(A6) *hold, if* $\mu_k = o(\|\mathrm{d}z^k\|)$, *then the sequence* $\{z^k\}$ *produced by algorithm A superlinearly converges to a stationary point* z^* *of (LCP)(1.1), i.e.* $\|z^{k+1} - z^*\| = o(\|z^k - z^*\|)$.

Proof. Let

$$u^k = D_k(N \quad M)\mathrm{d}s^k + \Gamma_w^k v^k, \quad \tilde{u}^k = \begin{pmatrix} \lambda_{J_k}^k \\ u^k \\ v^k \end{pmatrix}, \quad \tilde{u}^* = \begin{pmatrix} \lambda_{J_k}^* \\ u^* \\ \tilde{v}^* \end{pmatrix},$$

$$\tilde{h}(z,\mu_k) = \begin{pmatrix} A_{J_k}x - b_{J_k} \\ Nx + My - w + q \\ \Phi(y,w,\mu_k) \end{pmatrix}, \quad g(z,\mu_k) = \begin{pmatrix} \nabla f(s) \\ 0 \end{pmatrix} + R(z,\mu_k)\tilde{u}^*, \tag{5.15}$$

where u^* and $\tilde{v}^*$ are defined by Lemma 5.2. From (SLE_2)(3.4), one gets

$$(\nabla f(s^k)^{\mathrm{T}} 0)^{\mathrm{T}} + B_k \mathrm{d}z^k + R_k \tilde{u}^k = 0, \quad R_k^{\mathrm{T}}\mathrm{d}z^k + \tilde{h}(z^k,\mu_k) = 0. \tag{5.16}$$

Since $(z^*, \lambda^*, u^*, \tilde{v}^*)$ is a KKT pair for $\mu = 0$ of problem (2.14), and $J_k \equiv I(x^*)$, so $g(z^*, 0) = 0$,Therefore, by Taylor expansion and (2.18), we have

$$g(z^*,\mu_k) = g(z^*,0) + O(\mu_k) = O(\mu_k), \quad H(z^*,\tilde{v}^*,\mu_k) = H(z^*,\tilde{v}^*,0) + O(\mu_k), \tag{5.17}$$

$$\begin{aligned} g(z^k,\mu_k) &= g(z^*,\mu_k) + \nabla g(z^*,\mu_k)^{\mathrm{T}}(z^k - z^*) + o(\|z^k - z^*\|) \\ &= \nabla g(z^*,\mu_k)(z^k - z^*) + o(\|z^k - z^*\|) + O(\mu_k) \\ &= H(z^*,\tilde{v}^*,\mu_k)(z^k - z^*) + O(\mu_k) + o(\|z^k - z^*\|) \\ &= H(z^*,\tilde{v}^*,0)(z^k - z^*) + o(\|z^k - z^*\|) + O(\mu_k). \end{aligned}$$

So by (5.16) and (5.15), we obtain

$$-B_k \mathrm{d}z^k - R_k \tilde{u}^k = H(z^*, \tilde{v}^*, 0)(z^k - z^*) - R_k \tilde{u}^* + o(\|z^k - z^*\|) + O(\mu_k).$$

Since $P_k R_k = 0$ follows from (5.13), so, multiplying the two sides of the equation above by matrix P_k , we get

$$P_k H(z^*, \tilde{v}^*, 0)(z^k - z^*) = -P_k B_k \mathrm{d}z^k + o(\|z^k - z^*\|) + O(\mu_k). \tag{5.18}$$

From assumption (A6), we have

$$\begin{aligned} &P_k H(z^*, \tilde{v}^*, 0)(z^k + \mathrm{d}\hat{z}^k - z^*) = P_k H(z^*, \tilde{v}^*, 0)(z^k - z^* + \mathrm{d}z^k + \widehat{\mathrm{d}z}^k - \mathrm{d}z^k) \\ &= P_k H(z^*, \tilde{v}^*, 0)(z^k - z^*) + P_k H(z^*, \tilde{v}^*, 0)\mathrm{d}z^k + P_k H(z^*, \tilde{v}^*, 0)(\widehat{\mathrm{d}z}^k - \mathrm{d}z^k) \\ &= P_k (H(z^*, \tilde{v}^*, 0) - B_k)\mathrm{d}z^k + o(\|z^k - z^*\|) + O(\mu_k) \\ &= o(\|\mathrm{d}z^k\|) + o(\|z^k - z^*\|) + O(\mu_k). \end{aligned} \tag{5.19}$$

In view of the definition of $\tilde{h}(z, \mu_k)$, we know $\tilde{h}(z^*, 0) = 0$, $\nabla_z \tilde{h}(z^*, \mu_k) = R_k$, so one has by Taylor expansion,

$$\begin{aligned} \tilde{h}(z^*, \mu_k) &= \tilde{h}(z^k, \mu_k) + \nabla_z \tilde{h}(z^k, \mu_k)^{\mathrm{T}}(z^* - z^k) + o(\|z^k - z^*\|) \\ &= \tilde{h}(z^k, \mu_k) + R_k^{\mathrm{T}}(z^* - z^k) + o(\|z^k - z^*\|). \end{aligned}$$

On the other hand, $\tilde{h}(z^*, \mu_k) = \tilde{h}(z^*, 0) + O(\mu_k) = O(\mu_k)$, so

$$R_k^{\mathrm{T}}(z^k - z^*) = \tilde{h}(z^k, \mu_k) + o(\|z^k - z^*\|) + O(\mu_k).$$

This along with Lemma 5.4 and (5.16) implies

$$\begin{aligned} R_k^{\mathrm{T}}(z^k + \widehat{\mathrm{d}z}^k - z^*) &= R_k^{\mathrm{T}}\mathrm{d}z^k + R_k^{\mathrm{T}}(z^k - z^*) + R_k^{\mathrm{T}}(\widehat{\mathrm{d}z}^k - \mathrm{d}z^k) \\ &= R_k^{\mathrm{T}}\mathrm{d}z^k + \tilde{h}(z^k, \mu_k) + o(\|\mathrm{d}z^k\|) + o(\|z^k - z^*\|) + O(\mu_k) \\ &= o(\|\mathrm{d}z^k\|) + o(\|z^k - z^*\|) + O(\mu_k). \end{aligned} \tag{5.20}$$

In view of the fact that $z^{k+1} = z^k + \widehat{\mathrm{d}z}^k$, so we have from (5.19) and (5.20)

$$\begin{pmatrix} P_k H(z^*, \tilde{v}^*, 0) & R_k \\ R_k^{\mathrm{T}} & 0 \end{pmatrix} \begin{pmatrix} z^{k+1} - z^* \\ 0 \end{pmatrix} = o(\|\mathrm{d}z^k\|) + o(\|z^k - z^*\|) + O(\mu_k),$$

which together with Lemma 5.6(ii) and $\mu_k = o(\|\mathrm{d}z^k\|)$ gives that

$$\begin{aligned} &\|z^{k+1} - z^*\| \leqslant o(\|z^k - z^*\|) + o(\|\mathrm{d}z^k\|) = o(\|z^k - z^*\|) + o(\|\widehat{\mathrm{d}z}^k\|) \\ &= o(\|z^k - z^*\|) + o(\|z^{k+1} - z^k\|) \leqslant o(\|z^{k+1} - z^*\|) + o(\|z^k - z^*\|). \end{aligned}$$

Hence

$$\frac{\|z^{k+1}-z^*\|}{\|z^k-z^*\|}\left(1-\frac{o(\|z^{k+1}-z^*\|)}{\|z^{k+1}-z^*\|}\right)\leqslant\frac{o(\|z^k-z^*\|)}{\|z^k-z^*\|}.$$

Passing to the limit $k\to\infty$, we have

$$\lim_{k\to\infty}\frac{\|z^{k+1}-z^*\|}{\|z^k-z^*\|}=0,\quad \text{i.e. } \|z^{k+1}-z^*\| \;=\; o(\|z^k-z^*\|). \qquad \square$$

6. Numerical Results

In order to test the efficience of the proposed algorithm, we solve some examples. The computing results show that the algorithm is efficient and fast. The following three test problems are taken from [4].

PROBLEM 1

$$\begin{aligned}&\min\ f(x,y)=\frac{1}{2}x^2+\frac{1}{2}xy-95x\\ &\text{s.t. } 0\leqslant x\leqslant 200,\\ &\quad w=\frac{1}{2}x+2y-100,\\ &\quad 0\leqslant w\perp y\geqslant 0.\end{aligned}$$

PROBLEM 2. Let

$$f(x,y)=\frac{1}{2}\left[(x_1+x_2+y_1-15)^2+(x_1+x_2+y_2-15)^2\right],$$

$$A=\begin{pmatrix}1&0\\-1&0\\0&1\\0&-1\end{pmatrix},\quad b=\begin{pmatrix}10\\0\\10\\0\end{pmatrix},\quad M=\begin{pmatrix}2&\frac{8}{3}\\ \frac{5}{4}&2\end{pmatrix},\quad N=\begin{pmatrix}\frac{8}{3}&2\\2&\frac{5}{4}\end{pmatrix}.$$

PROBLEM 3. This is a set of several problems with data generated as follows. The objective function is given by

$$f(x,y)=\frac{1}{2}x^Tx+e^Ty,\ e=(1,\ldots,1)^T,$$

and the matrix M is strictly diagonally dominant (thus P) with off-diagnoal entires being random numbers between 0 and 1, the entires of the matrices N, A and vectors q, b are randomly generated; moreover, q and b are nonnegative vectors and the pair (A,b) is such that the n-vector of all ones satisfies $Ax\leqslant b$. This kind of problems have optimal solutions $x^*=0$, $y^*=0$.

Table 1. Performance of algorithms A and FLP on the three test problems

Problem	Size (p,m,n)	Initial point (x^0, y^0)	Updating μ_k	Formula for comput ing B_k	N_{A}	$f(x^*,y^*)$	N_{FLP}
				(6.1)	20	−3.2666666665e + 003	23
			$\mu_{k+1}=0.5\mu_k$	(6.2)	18	−3.2666666660e + 003	22
1	(2,1,1)	(0,0)		(6.3)	19	−3.2666666663e + 003	23
				(6.1)	23	−3.2666666667e + 003	Fails
			$\mu_{k+1}=0.5\mu_k^{1.5}$	(6.2)	20	−3.2666666667e + 003	Fails
				(6.3)	5	−3.2666666666e + 003	8
				(6.1)	Fails		48
			$\mu_{k+1}=0.5\mu_k$	(6.2)	42	1.1501591998e-027	Fails
2	(4,2,2)	(0,0)		(6.3)	Fails		28
				(6.1)	Fails		Fails
			$\mu_{k+1}=0.5\mu_k^{1.5}$	(6.2)	18	0.0000000000e + 000	Fails
				(6.3)	Fails		Fails
				(6.1)	36	404566452855e-010	50
			$\mu_{k+1}=0.5\mu_k$	(6.2)	35	5.9148158177e-010	30
3	(30,30,50)	(ones,zeros)		(6.3)	37	1.7414574636e-010	28
				(6.1)	7	2.5837711355e016	11
			$\mu_{k+1}=0.5\mu_k^{1.5}$	(6.2)	7	2.5716508351e015	11
				(6.3)	7	5.1736973325e-016	12
				(6.1)	41	3.8260317239e-011	50
				(6.2)	39	1.0485304714e-010	31
3	(50,60,50)	(ones,zeros)	$\mu_{k+1}=0.5\mu_k$	(6.3)	38	1.9690694013e-010	30
				(6.1)	7	3.4554261581e-016	12
			$\mu_{k+1}=0.5\mu_k^{1.5}$	(6.2)	7	1.1889320698e-015	11
				(6.3)	7	5.8833421524e-016	11
				(6.1)	38	2.1989377005e-010	51
			$\mu_{k+1}=0.5\mu_k$	(6.2)	38	2.2583316969e-010	32
3	(100,70,70)	(ones,zeros)		(6.3)	40	8.9947090307e-011	31
				(6.1)	7	2.8473364919e-015	11
			$\mu_{k+1}=0.5\mu_k^{1.5}$	(6.2)	7	5.8214664360e-016	12
				(6.3)	7	6.4785283839e-015	11
				(6.1)	42	3.7324915964e-011	50
			$\mu_{k+1}=0.5\mu_k$	(6.2)	44	1.0236235664e-011	38
3	(150,100,100)	(ones,zeros)		(6.3)	40	9.1641172875e-011	38
				(6.1)	7	9.3123502743e-016	11
			$\mu_{k+1}=0.5\mu_k^{1.5}$	(6.2)	7	1.8769541428e-015	11
				(6.3)	7	5.5426324526e-016	11

In the test processes, the termination rule and parameters are chosen as follows:

$$\|\mathrm{d}z_0^k\|_\infty \leqslant 10^{-8} \quad \text{and} \quad \lambda_{0,J_k}^k \geqslant 0,$$

$$\hat{\varepsilon} = 10^{-8},\ \lambda = 0,\ \alpha_{-1} = 10,\ \varepsilon = 1,\ \beta = 0.5,\ \delta = 10,\ \sigma = 0.1, \bar{\eta} = 0.5,$$

and we choose the matrices B_k, i.e., C_k, and D_k in (2.21) by one of following forms:

$$C_k \equiv \begin{pmatrix} I_n & 0 \\ 0 & 0 \end{pmatrix}, \ D_k \equiv 0, \tag{6.1}$$

$$C_k \equiv I_{n+m}, \ D_k \equiv 0, \tag{6.2}$$

$$C_k = \begin{pmatrix} \nabla^2_{xx} f(s^k) & \nabla^2_{xy} f(s^k) \\ \nabla^2_{yx} f(s^k) & \nabla^2_{yy} f(s^k) + \mathrm{diag}\left(v_i^{k-1} \frac{\partial^2 \phi(t_i^k, \mu_k)}{\partial a^2} \right) \end{pmatrix},$$

$$D_k = \mathrm{diag}\left(v_i^{k-1} \frac{\partial^2 \phi(t_i^k, \mu_k)}{\partial b^2} \right)_{m \times m,} \tag{6.3}$$

with $v^{-1} = (-1, \ldots, -1)^T \in \Re^m$.

The numerical results are given in Table 1 below where we make a simple comparison with the algorithm which is proposed by Fukushima, Luo and Pang in [4] (FLP). The word "fails" means that the associated algorithms can not achieve the given precision 10^{-8} or the number of iterations is too large. N_{A} and N_{FLP} represent the total number of iteration of Algorithm A, Algorithm FLP, respectively. From the computational results of Table 1, the proposed Algorithm A with superlinear convergence is more efficient and faster than FLP algorithm. Specifically, FLP algorithm fails in many cases of choice of matrix B_k and perturbed parmeter μ_k, but our Algorithm A is efficient still in these cases.

7. Concluding Remarks

In this paper, we have first transformed the discussed problem (LCP)(1.1) into a family of general nonlinear optimization problems (2.14) containing parameters, then have established an SSLE algorithm for it. We have tested the proposed algorithm on some examples, and the results have shown that the algorithm is numerically doable. Moreover, the three systems of equations solved at each iteration have the same coefficients and the computational amount of the proposed algorithm is less than that of existing SQP type algorithm. So we feel that the proposed algorithm is an effective method for (LCP)(1.1) and will be further studied to turn it into a practical tool for solving large-scale engineering and economic applications of LCP.

Meanwhile, we point out that the proposed algorithm is rather sensitive to the way the perturbed parameter μ_k is updated. As expected, the matrix

B_k has a dramatic effect on the performance of the algorithm; In fact, in some cases a careless choice of B_k has caused the algorithm to fail.

Acknowledgement

This work was supported by the National Natural Science Foundation (10261001) and the Guangxi Province Natural Science Foundation (0236001, 0249003) of China.

Reference

1. Chen, B. and Harker, P.T. (1993), A non-interior-point continuation method for linear complementarity problems. *SIAM Journal on Matrix Analysis and Application* 14, 1168–1190.
2. Faccshinei, F., Jiang, H. and Qi, L. (1999), A smoothing method for mathematical programs with equilibrium constraints. *Mathematical Programming* 85, 107–134.
3. Facchinei, F. (1997), Robust recursive quadratic programming algorithm model with global and superlinear convergence properties. *Journal of Optimization Theory and Application* 92(3), 543–579.
4. Fukushima, M., Luo, Z.Q. and Pang, J. (1998), A globally convergent sequential quadratic programming algorithm for mathematical programming with linear complementarity constraints. *Computational Optimization and Applications* 10, 5–34.
5. Gao, Z.Y., He, G.P. and Wu, F. (1997), Sequential Systems of Linear Equations Algorithm for Nonlinear Optimization Problems with General Constraints. *Journal of Optimization Theory and Application* 95(2), 371–397.
6. Gao, Z.Y., He, G.P. and Wu, F. (1997), A sequential systems of linear equations method with arbitrary initial point. *Science in China* 27(1), 24–33.
7. Han, S.P. (1976), Superlinearly convergent variable metric algorithms for general nonlinear programming problems. *Mathematical Programming* 11, 263–282.
8. Jian, J.B. (2002), A Feasible Method for Superlinearly and Quadratically Convergent Sequential Systems of Equations. *Acta Mathematica Sinica* 45(6), 1137–1146.
9. Jian, J.B. (2000), Researches on superlinearly and quadratically convergent algorithms for nonlinearly constrained optimization, Ph.D.Thesis, Xi'an Jiaotong University, Xi'an, China.
10. Jiang, H. and Ralph, D. (2000), Smooth SQP methods for mathematical programs with nonlinear complementarity constraints. *SIAM Journal on Optimization* 10(3), 779–808.
11. Kanzow, C. (1996), Some noninterior continuation methods for linear complementarity problems. *SIAM Journal on Matrix Analysis and Application* 17, 851–868.
12. Kocvara, M. and Outrata, J.V. (1994), On optimization systems governed by implicit complementarity problems. *Numerical Functional Analysis and Optimization* 15, 869–887.
13. Kocvara, M. and Outrata, J.V. (1995), A nonsmooth approach to optimization problems with equilibrium constraints, In: Ierns, M.c. and Pang, J.D. (eds.), *Proceedings of the international conference on Complementarity Problems*, pp. 148–164. SIAM Publications, Baltimore, Maryland.
14. Kojima, M., Megiddo, N., et al. (1991), *A Unified Approach to Interior Point Algorithms for Linear Complementarity problems*, Springer-Verlag, Berlin, Heidelberg.
15. Luo, Z.Q., Pang, J.S. and Ralph, D.M. (1996), Mathematical Programs with Equilibrium Constraints[M], Cambrige University Press, London.

16. Outrata, J.V., Kocvara, M. and Zowe, J. (1998), Nonsmooth Approach to Optimization Problems with Equilibrium Constraints[M], Kluwer Academic Publishers, Netherlands.
17. Panier, E.R., Tits, A.L. and Herskovits, J.N. (1988), A QP-free global convergent, locally superlinearly convergent algorithm for inequality constrained optimization. *SIAM Journal on Control and Optimization* 26(4), 788–811.