

# On the Finite Termination of An Entropy Function Based Smoothing Newton Method for Vertical Linear Complementarity Problems

Shu-Cherng Fang <sup>\*</sup>      Jiye Han <sup>† §</sup>      Zhenghai Huang <sup>†</sup>      Ş. İlker Birbil <sup>‡</sup>

February 18, 2003



## Abstract

By using a smooth entropy function to approximate the non-smooth max-type function, a vertical linear complementarity problem (VLCP) can be treated as a family of parameterized smooth equations. A Newton-type method with a testing procedure is proposed to solve such a system. We show that the proposed algorithm finds an exact solution of VLCP in a finite number of iterations, under some conditions milder than those assumed in literature. Some computational results are included to illustrate the potential of this approach.

**Keywords** Vertical linear complementarity problems, smoothing approximation, Newton method, entropy function, finite termination.

---

<sup>\*</sup>Industrial Engineering and Operations Research, North Carolina State University, Raleigh, NC 26695-7906, USA. E-mail: fang@eos.ncsu.edu

<sup>†</sup>Institute of Applied Mathematics, Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences, Beijing, 100080, P.R. China. E-mail: hjyyyf@ihw.com.cn, zhhuang@mail.amt.ac.cn

<sup>‡</sup>Erasmus Research Institute of Management (ERIM), Erasmus University, Postbus 1738, 3000 DR Rotterdam, The Netherlands. E-mail: sibirbil@few.eur.nl

<sup>§</sup>This author's work was partially supported by the National Natural Science Foundation of China (Grant No. 19731001)

# 1 Introduction

Let  $R^n$  be the space of  $n$ -dimensional real column vectors and  $R^{m \times n}$  the space of  $m \times n$  real matrices. We define an index set  $\mathcal{I} := \{1, 2, \dots, n\}$ . Given  $N^i \in R^{m_i \times n}$ ,  $q^i \in R^{m_i}$ , and  $s^i \in R^{m_i}$ , with  $m = \sum_{i=1}^n m_i \geq n$ , define

$$N := \begin{pmatrix} N^1 \\ N^2 \\ \vdots \\ N^n \end{pmatrix} \in R^{m \times n}, \quad q := \begin{pmatrix} q^1 \\ q^2 \\ \vdots \\ q^n \end{pmatrix} \in R^m, \quad \text{and} \quad s := \begin{pmatrix} s^1 \\ s^2 \\ \vdots \\ s^n \end{pmatrix} \in R^m.$$

We say  $N$  is a vertical block matrix of type  $(m_1, \dots, m_n)$ . The vertical linear complementarity problem (VLCP) associated with  $N$  and  $q$  is to find a pair of vectors  $x \in R^n$  and  $s \in R^m$  such that

$$x \geq 0, \quad s^i = N^i x + q^i \geq 0, \quad \text{and} \quad x_i \prod_{j=1}^{m_i} s_j^i = 0, \quad \forall i \in \mathcal{I}, \quad (1)$$

where  $x_i$  and  $s_j^i$  denote the  $i$ th component of  $x$  and the  $j$ th component of  $s^i$ , respectively. This problem was first introduced by Cottle and Dantzig in name of the *generalized linear complementarity problem* [5], since when  $m_i = 1$  for all  $i \in \mathcal{I}$ , the problem reduces to an ordinary linear complementarity problem [6]. VLCP has various applications in nonlinear networks [13], game theory [16], control theory [31] and economics [9]. Good references can be found in [8, 15, 22, 23, 25, 26, 27, 32].

Ebiefung [8] showed that VLCP is equivalent to a nonlinear complementarity problem  $\text{NCP}(F)$  with  $F = (F_1, \dots, F_n)^T$  and  $F_j$ ,  $j = 1, \dots, n$ , being piecewise linear and concave. It can also be shown that VLCP is equivalent to a system of piecewise linear equations, or a multi-objective program. By extending Lemke's pivoting algorithm, Cottle and Dantzig proposed the first algorithm for VLCP [5]. An interior point method for solving extended vertical linear complementarity problems can be found in [34]. Peng and Lin [25] proposed a non-interior continuation method for solving VLCP. They showed that their algorithm converges locally with a Q-quadratic rate under the following assumptions:

- Non-singularity assumption, i.e., the Jacobian matrix involved in Newton equation is nonsingular at the solution point, or the iteration matrices are uniformly nonsingular.
- Strict complementarity assumption, i.e., the solution of the problem concerned is strictly complementary.

In this paper, we are interested in developing a smoothing Newton method for solving VLCP with finite termination. Our approach is based on the entropic smoothing for the max-type function. Let  $g_i : R^n \rightarrow R$ ,  $\forall i \in \mathcal{I}$ , be differentiable and define a max-type function  $g : R^n \rightarrow R$  by

$$g(x) := \max_{i \in \mathcal{I}} g_i(x).$$

Although the function  $g$  is piecewise smooth and locally Lipschitz continuous, it is not differentiable. Given any  $\mu > 0$ , consider the following entropy-type function as a smoothing

approximation function of  $g$ ,

$$g(x, \mu) := \mu \ln \sum_{i=1}^n \exp(g_i(x)/\mu). \quad (2)$$

Note that, for  $\mu > 0$ ,

$$g(x, \mu) = g(x) + \mu \ln \sum_{i=1}^n \exp\left(\frac{g_i(x) - g(x)}{\mu}\right). \quad (3)$$

Moreover,

$$g(x) \leq g(x, \mu) \leq g(x) + \mu \ln(n), \quad \forall x \in R^n \text{ and } \mu > 0. \quad (4)$$

Therefore,  $g(x, \mu) \rightarrow g(x)$  as  $\mu \rightarrow 0$ . This fact allows us to develop iterative methods based on  $g(x, \mu)$  to solve the problem without facing the non-differentiability problem of  $g(x)$ . The function (2) was introduced by Kort and Bertsekas [20] as a penalty function for constrained minimization. Goldstein [14] studied this function intensively and attributed the basic approximation formula (4) to his former student Chang [3]. Since the function (2) can be derived from the dual problem of an entropy optimization problem [11], we call function (2) an entropic smoothing approximation function. Independently, Li [21] discovered a few properties of this function and named it as the *aggregate function*. Related work can be found in [1, 3, 14, 25, 26, 27, 37]. In particular, Peng and Lin [25] also use this function in developing their non-interior continuation method for solving VLCP. Also note that since a lower bound of the value of  $g(x)$  is singled out in the representation (3), it can be used in computation to avoid the potential overflow problem arising from any exponential function evaluation in (2).

The finite termination of iterative methods is an interesting and important research topic. This property has been investigated for various cases including the interior point methods [19, 24, 38], non-smooth Newton methods [12, 33], and smoothing methods [4, 7]. It is our objective to develop a Newton-type method based on the entropic smoothing function for solving VLCP in a finite number of iterations.

It is well-known that many smoothing Newton methods need to use the non-singularity assumption and the strict complementarity assumption to obtain the local superlinear convergence of the methods [2, 25, 28]. The non-singularity assumption has recently been relaxed in a few smoothing Newton methods [10, 18, 36]. However, in order to achieve finite termination for iterative methods, this assumption has been commonly adopted, for example, see [4, 7, 12, 33]. In this paper, by using the entropic approximation function, we present a smoothing Newton method for solving VLCP, in which a test procedure of finding a solution point in the optimal face of the problem is embedded into each iteration. We show that for  $N$  being a vertical block  $P_0$  and  $R_0$  matrix, if either (i) the strict complementarity condition holds, or (ii) the solution set of (1) is a singleton, then the proposed algorithm finds an exact solution to VLCP in a finite number of iterations. It should be noted that the commonly used non-singularity assumption implies that the solution set of the underlying problem is a singleton. Therefore, the hypothesis used in this paper is weaker than those used in finite termination methods [4, 7, 12, 33] and in smoothing Newton methods [2, 25, 28, 29, 35].

The paper is organized as follows. Some basic concepts and properties for VLCP are introduced in Section 2. Then we present in Section 3 a smoothing Newton method based on

the entropic approximation function for solving VLCP. In Section 4, we show the finite termination property of the proposed algorithm. Some numerical results are presented in Section 5.

## 2 Basic Concepts and Properties

A square matrix  $M \in R^{n \times n}$  is said to be a  $P_0$ -matrix, if for all non-zero vector  $x \in R^n$ , there exists a component  $x_i \neq 0$  such that  $x_i(Mx)_i \geq 0$ . For the vertical block matrix  $N$  of type  $(m_1, \dots, m_n)$ , a square submatrix of  $N$  of order  $n$  is said to be a *representative submatrix*, if its  $i$ th row is drawn from the  $i$ th block  $N^i$  of  $N$  for each  $i \in \mathcal{I}$ . The following definition is from [23]:

**Definition 1** Let  $N \in R^{m \times n}$  be a vertical block matrix of type  $(m_1, \dots, m_n)$ .  $N$  is called a vertical block  $P_0$ -matrix, if all its representative submatrices are  $P_0$ -matrices. Moreover,  $N$  is called a vertical block  $R_0$ -matrix, if

$$\begin{pmatrix} \min\{x_1, N_1^1 x, \dots, N_{m_1}^1 x\} \\ \vdots \\ \min\{x_n, N_1^n x, \dots, N_{m_n}^n x\} \end{pmatrix} = 0 \iff x = 0,$$

where  $N_j^i$  denotes the  $j$ th row of  $i$ th block.

For a VLCP with given  $N$  and  $q$ , we use the definition [25] for the piecewise smooth function  $H : R^n \rightarrow R^n$  with

$$\begin{aligned} H(x) &:= \begin{pmatrix} \min\{x_1, N_1^1 x + q_1^1, \dots, N_{m_1}^1 x + q_{m_1}^1\} \\ \vdots \\ \min\{x_n, N_1^n x + q_1^n, \dots, N_{m_n}^n x + q_{m_n}^n\} \end{pmatrix} \\ &= - \begin{pmatrix} \max\{-x_1, -(N_1^1 x + q_1^1), \dots, -(N_{m_1}^1 x + q_{m_1}^1)\} \\ \vdots \\ \max\{-x_n, -(N_1^n x + q_1^n), \dots, -(N_{m_n}^n x + q_{m_n}^n)\} \end{pmatrix}. \end{aligned} \quad (5)$$

Let  $\mathcal{F}$  be the feasible solution set of VLCP, i.e.,

$$\mathcal{F} = \{(x, s) \in R^n \times R^m : s = Nx + q \geq 0 \text{ and } x \geq 0\}.$$

Also let  $\mathcal{S}$  denote the solution set of VLCP, i.e.,

$$\mathcal{S} = \{(x, s) \in R^n \times R^m : (x, s) \text{ satisfies (1)}\}.$$

Then we know that

$$(x, s) \in \mathcal{S} \text{ if and only if } x \text{ solves } H(x) = 0 \text{ and } s = Nx + q. \quad (6)$$

By applying the entropic approximation to  $H(x)$ , we can define a smooth function, for any  $\mu > 0$ ,

$$H(x, \mu) := - \begin{pmatrix} \mu \ln \left( \exp(-x_1/\mu) + \sum_{j=1}^{m_1} \exp(-(N_j^1 x + q_j^1)/\mu) \right) \\ \vdots \\ \mu \ln \left( \exp(-x_n/\mu) + \sum_{j=1}^{m_n} \exp(-(N_j^n x + q_j^n)/\mu) \right) \end{pmatrix}. \quad (7)$$

Consequently,  $H(x, \mu) \rightarrow H(x)$  as  $\mu \rightarrow 0$ . This fact and (6) indicate that one can solve VLCP by taking the following steps: (i) start with a  $\mu > 0$  and approximate VLCP by the parameterized smooth equations  $H(x, \mu) = 0$  and  $s = Nx + q$ , (ii) solve  $H(x, \mu) = 0$  and maintain  $s = Nx + q$  at each iteration, and (iii) refine the approximation by reducing the parameter  $\mu$  to zero. Since it is usually very difficult to solve  $H(x, \mu) = 0$  in an exact manner, for  $\mu > 0$ , like in other interior point and non-interior continuation methods, we use the following definition of neighborhood:

$$\mathcal{N}(\beta, \mu) := \{x \in R^n : \|H(x, \mu)\| \leq \beta\mu\}, \quad (8)$$

for  $\beta > 0$  and  $\mu > 0$ .

The following lemma whose proof can be found in [25] summarizes some basic properties of the functions  $H(x)$  and  $H(x, \mu)$ .

**Lemma 1** Suppose that  $N \in R^{m \times n}$  is a vertical block matrix of type  $(m_1, \dots, m_n)$ . Let  $H(x)$  and  $H(x, \mu)$  be defined by (5) and (7), respectively. Then

(i) For each  $i \in \mathcal{I}$ ,  $-H_i(x, \mu)$  is convex and monotonically increasing with respect to  $\mu > 0$  and

$$-H_i(x) \leq -H_i(x, \mu) \leq -H_i(x) + \mu \ln(m_i + 1),$$

where  $H_i(x)$  is the  $i$ th component of  $H(x)$  as defined in (5).

(ii) If  $N$  is a vertical block  $P_0$ -matrix, then, for any  $\mu > 0$ ,  $-H_i(x, \mu)$  is an infinite order differentiable convex function with respect to  $x \in R^n$ , and  $\nabla_x H(x, \mu)$  is nonsingular for any  $x \in R^n$  with

$$\nabla_x H(x, \mu) = \begin{pmatrix} \lambda_0^1(x, \mu)e_1^T + \sum_{j=1}^{m_1} \lambda_j^1(x, \mu)N_j^1 \\ \lambda_0^2(x, \mu)e_2^T + \sum_{j=1}^{m_2} \lambda_j^2(x, \mu)N_j^2 \\ \vdots \\ \lambda_0^n(x, \mu)e_n^T + \sum_{j=1}^{m_n} \lambda_j^n(x, \mu)N_j^n \end{pmatrix},$$

where  $e_i$  is the  $i$ th column of the  $n \times n$  identity matrix,

$$\begin{aligned}
\lambda_0^i(x, \mu) &= \frac{\exp\left(-\frac{x_i}{\mu}\right)}{\exp\left(-\frac{x_i}{\mu}\right) + \sum_{l=1}^{m_i} \exp\left(-\frac{N_l^i x + q_l^i}{\mu}\right)} \\
&= \frac{\exp\left(\frac{-x_i + H_i(x)}{\mu}\right)}{\exp\left(\frac{-x_i + H_i(x)}{\mu}\right) + \sum_{l=1}^{m_i} \exp\left(\frac{-N_l^i x - q_l^i + H_i(x)}{\mu}\right)}, \quad i = 1, 2, \dots, n, \\
\lambda_j^i(x, \mu) &= \frac{\exp\left(-\frac{N_j^i x + q_j^i}{\mu}\right)}{\exp\left(-\frac{x_i}{\mu}\right) + \sum_{l=1}^{m_i} \exp\left(-\frac{N_l^i x + q_l^i}{\mu}\right)} \\
&= \frac{\exp\left(\frac{-N_j^i x - q_j^i + H_i(x)}{\mu}\right)}{\exp\left(\frac{-x_i + H_i(x)}{\mu}\right) + \sum_{l=1}^{m_i} \exp\left(\frac{-N_l^i x - q_l^i + H_i(x)}{\mu}\right)}, \quad i = 1, 2, \dots, n.
\end{aligned}$$

(iii)  $N$  is a vertical block  $R_0$ -matrix if and only if  $\lim_{\|x\| \rightarrow \infty} \|H(x)\|/\|x\| \geq c_0$  holds for some constant  $c_0 > 0$ .

(iv) For any  $x, y \in R^n$  and  $\mu > 0$ , there exists a constant  $c_1 > 0$  such that

$$\|H(y, \mu) - H(x, \mu) - \nabla_x H(x, \mu)(y - x)\| \leq \frac{\sqrt{n}c_1}{\mu} \|y - x\|^2.$$

(v) For any  $\mu_1, \mu_2 > 0$ ,

$$\|H(x, \mu_1) - H(x, \mu_2)\| \leq \sqrt{n}(\ln \bar{m})|\mu_1 - \mu_2|,$$

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

(vi) If  $N$  is a vertical block  $R_0$ -matrix and  $\mathcal{S} \neq \emptyset$ , then there exists a constant  $c_2 > 0$  such that

$$\text{dist}((x, s), \mathcal{S}) := \min_{(y, s) \in \mathcal{S}} \|y - x\| \leq c_2 \|H(x)\|$$

for any  $x \in R^n$ .

Note that result (i) implies that  $H(x, \mu) \rightarrow H(x)$  as  $\mu \rightarrow 0$ . Hence, we define  $H(x, 0) := H(x)$  for  $x \in R^n$ .

### 3 Proposed Algorithm

Define an index set  $\mathcal{J} := \{(i, j) : i \in \mathcal{I}, j = 1, 2, \dots, m_i\}$ . The  $j$ th row  $N_j^i$  of  $N^i$  is called the  $(i, j)$  row of matrix  $N$ . Let  $K_1 \subseteq \mathcal{I}$  and  $K_2 \subseteq \mathcal{J}$  be two nonempty sets, then  $x_{K_1}$  and  $s_{K_2}$  denote the vectors obtained from all components  $x_r$  in  $x$  with  $r \in K_1$  and all components  $s_j^i$  in  $s$  with  $(i, j) \in K_2$ , respectively. Moreover,  $N_{K_2 K_1}$  denotes the submatrix of  $N$  induced by those components of  $N$  whose row indices belong to  $K_2$  and column indices belong to  $K_1$ , respectively. In what follows,  $k$  always denotes the iteration number.

**Algorithm 3.1** Given  $\sigma_1, \sigma_2 \in (0, 1)$ ,  $\alpha_1, \alpha_2 \in (0, 1)$ ,  $\gamma \in (0, 1)$ ,  $p \geq 1$ ,  $\mu_0 > 0$ , and  $x^0 \in \mathbb{R}^n$ , choose  $\beta > 0$  such that  $\|H(x^0, \mu_0)\| \leq \beta\mu_0$ . Set  $s^0 := Nx^0 + q$ ,  $k := 0$ .

**Step 1** If  $H(x^k) = 0$ , then stop (output  $x^k$  as a solution).

**Step 2** If  $\mu_k > \gamma$ , then go to Step 3; otherwise, define four sets

$$\begin{aligned} A &:= \{i \in \mathcal{I} : x_i^k > \sqrt{\mu_k}\}, & C &:= \{i \in \mathcal{I} : x_i^k \leq \sqrt{\mu_k}\}, \\ B &:= \{(i, j) \in \mathcal{J} : (s^k)_j^i > \sqrt{\mu_k}\} & D &:= \{(i, j) \in \mathcal{J} : (s^k)_j^i \leq \sqrt{\mu_k}\}. \end{aligned} \quad (9)$$

If one of the following four cases occurs, then stop (output  $x^{k+1}$  as a solution); otherwise, go to Step 3.

**Case (i)** If  $A \neq \emptyset, B \neq \emptyset$  and for any  $i \in A$  there exists at least one index  $(i, j) \in \mathcal{J}$  such that  $(s^k)_j^i \leq \sqrt{\mu_k}$ , then solve the following system of equations:

$$\begin{pmatrix} s_B^k + \Delta s_B^k \\ 0 \end{pmatrix} = \begin{pmatrix} N_{BA} & N_{BC} \\ N_{DA} & N_{DC} \end{pmatrix} \begin{pmatrix} x_A^k + \Delta x_A^k \\ 0 \end{pmatrix} + \begin{pmatrix} q_B \\ q_D \end{pmatrix}. \quad (10)$$

If there exists a solution  $(\Delta x_A^k, \Delta s_B^k)$  such that  $x_A^k + \Delta x_A^k \geq 0$  and  $s_B^k + \Delta s_B^k \geq 0$ , then set

$$x_i^{k+1} := \begin{cases} x_i^k + \Delta x_i^k & \text{if } i \in A \\ 0 & \text{otherwise} \end{cases}, \quad (s^{k+1})_j^i := \begin{cases} (s^k)_j^i + (\Delta s^k)_j^i & \text{if } (i, j) \in B \\ 0 & \text{otherwise} \end{cases};$$

**Case (ii)** If  $A \neq \emptyset$  and  $B = \emptyset$ , then solve the following system of equations:

$$0 = (N_{DA}, N_{DC}) \begin{pmatrix} x_A^k + \Delta x_A^k \\ 0 \end{pmatrix} + q_D.$$

If there exists a solution  $\Delta x_A^k$  such that  $x_A^k + \Delta x_A^k \geq 0$ , then set

$$x_i^{k+1} := \begin{cases} x_i^k + \Delta x_i^k & \text{if } i \in A \\ 0 & \text{otherwise} \end{cases}, \quad s^{k+1} := 0.$$

**Case (iii)** If  $A = \emptyset, B \neq \emptyset$  and  $q_B > 0, q_D = 0$ , then set

$$x^{k+1} := 0, \quad (s^{k+1})_j^i := \begin{cases} q_j^i & \text{if } (i, j) \in B \\ 0 & \text{otherwise} \end{cases}.$$

**Case (iv)** If  $A = \emptyset, B = \emptyset$  and  $q_D = 0$ , then set  $x^{k+1} := 0, s^{k+1} := 0$ .

**Step 3** Find a Newton direction  $\Delta x^k$  by solving  $\nabla_x H(x^k, \mu_k) \Delta x^k = -H(x^k, \mu_k)$ . Let  $\theta_k$  be the maximum value of the set  $\{1, \alpha_1, \alpha_1^2, \dots\}$  such that

$$\|H(x^k + \theta_k \Delta x^k, \mu_k)\| \leq (1 - \sigma_1 \theta_k) \beta \mu_k.$$

Set  $x^{k+1} := x^k + \theta_k \Delta x^k$ . Moreover, let  $\lambda_k$  be the maximum value of the set  $\{\max\{1, \frac{1}{\sigma_2}(1 - \mu_k^p)\}, \alpha_2, \alpha_2^2, \dots\}$  such that

$$x^k + \theta_k \Delta x^k \in \mathcal{N}(\beta, (1 - \sigma_2 \lambda_k) \mu_k).$$

Set  $\mu_{k+1} := (1 - \sigma_2 \lambda_k) \mu_k$ .

**Step 4** Set  $s^{k+1} := Nx^{k+1} + q$  and  $k := k + 1$ . Go to Step 1.

Note that since the Jacobian matrix  $\nabla_x H(x, \mu)$  in Step 3 is guaranteed to be nonsingular for any  $\mu > 0$  and  $x \in R^n$  by the result (ii) of Lemma 1, it is not difficult to see that Algorithm 3.1 is well-defined. The initial value of parameter  $p$  ( $\geq 1$ ) can be selected to be a suitable positive integer. The parameter  $\gamma$  in Step 2 is used to control the quality of final solution. In the next section (see Lemma 4), we show that the four index sets (9) in Step 2 actually coincide with the index sets of a solution to VLCP as  $k$  becomes sufficiently large.

**Theorem 1** If Algorithm 3.1 terminates in either Step 1 or Step 2 for some  $k \geq 0$ , then  $(x^k, s^k)$  or  $(x^{k+1}, s^{k+1})$  is a solution to VLCP, respectively.

**Proof.** If Algorithm 3.1 terminates in Step 1, that is,

$$H(x^k, s^k) = 0 \quad (11)$$

for some  $k \geq 0$ . From the algorithm, it is easy to see that

$$s^k = Nx^k + q \quad (12)$$

for all  $k \geq 0$ . From (6), (11) and (12), it follows that  $(x^k, s^k)$  is a solution to VLCP.

If one stopping criterion in Step 2 is met. Since  $A$  and  $C$  form a partition of  $\mathcal{I}$ , and  $B$  and  $D$  form a partition of  $\mathcal{J}$ , it is not difficult to check from cases (i)-(iv) that  $x^{k+1}$  and  $s^{k+1}$  satisfy the non-negativity condition, feasibility condition, and complementarity condition of the system (1). Consequently,  $(x^{k+1}, s^{k+1})$  is a solution to VLCP. ■

A smoothing Newton method in general generates a sequence of infinitely many iterations. In this case, only an approximate solution is generated. But if the proposed algorithm terminates in either Step 1 or Step 2 for some  $k \geq 0$ , then Theorem 1 guarantees an exact solution to VLCP.

## 4 Finite Termination

In this section we show that, under some milder conditions, the stopping criteria in Step 2 of the proposed algorithm must be met as  $k$  becomes sufficiently large. This implies that the proposed algorithm terminates in a finite number of iterations.

**Theorem 2** Let  $N$  be a vertical block  $P_0$  and  $R_0$  matrix of type  $(m_1, \dots, m_n)$  and  $\{(x^k, s^k, \mu_k)\}$  be the sequence generated by Algorithm 3.1. If  $H(x^k) \neq 0$  for  $k = 0, 1, 2, \dots$ , then

- (i)  $\{(x^k, s^k, \mu_k)\}$  is a bounded infinite sequence,
- (ii) each accumulation point of the sequence  $\{(x^k, s^k)\}$  is a solution to VLCP.

**Proof.** (i) If  $H(x^k) \neq 0$  for all  $k \geq 0$ , then the proposed algorithm will not terminate at Step 2. Otherwise, if the algorithm terminates at Step 2 in  $k_0 \geq 0$  iterations, then from Theorem 1, we know that  $(x^{k_0+1}, s^{k_0+1}) \in \mathcal{S}$ , and hence  $H(x^{k_0+1}) = 0$ . Therefore an infinite sequence  $\{(x^k, s^k, \mu_k)\}$  is generated. From the algorithm itself, it is not difficult to see that  $s^k = Nx^k + q$

and  $x^k \in \mathcal{N}(\beta, \mu_k)$  for  $k = 0, 1, 2, \dots$ . Now, for any  $x^k \in \mathcal{N}(\beta, \mu_k)$ , the result (i) of Lemma 1 implies that

$$\|H(x^k)\| \leq \|H(x^k) - H(x^k, \mu_k)\| + \|H(x^k, \mu_k)\| \leq (\sqrt{n} \ln \bar{m} + \beta) \mu_0,$$

where  $\bar{m} = \max\{m_1, \dots, m_n\} + 1$ . This inequality and the result (iii) of Lemma 1 further imply that  $\{x^k\}$  is bounded. Consequently,  $\{s^k\}$  is bounded because  $s^k = Nx^k + q$ . Note that  $\{\mu_k\}$  obtained in Step 3 is a monotonically decreasing non-negative sequence. Hence the sequence  $\{(x^k, s^k, \mu_k)\}$  is bounded.

(ii) Since the infinite sequence  $\{(x^k, s^k, \mu_k)\}$  is bounded, there exists a convergent subsequence. We may assume without loss of generality that  $\lim_{k \rightarrow \infty} (x^k, s^k, \mu_k) = (x^*, s^*, \mu_*)$ . Because  $s^k = Nx^k + q$  holds for all  $k \geq 0$ , we have

$$s^* = Nx^* + q. \quad (13)$$

Noting that  $\{\mu_k\}$  is a monotonically decreasing non-negative sequence, we know  $\mu_* \geq 0$ . If  $\mu_* = 0$ , the result (i) of Lemma 1 implies that  $H(x^*) = H(x^*, \mu_*)$ . Moreover,  $x^k \in \mathcal{N}(\beta, \mu_k)$  implies that  $x^* \in \mathcal{N}(\beta, \mu_*)$ . Hence  $\|H(x^*, \mu_*)\| = 0$  and  $H(x^*) = 0$ . Together with (13), we know that  $(x^*, s^*) \in \mathcal{S}$ , and the desired result follows. We now show that  $\mu_* > 0$  will not occur. Assume that  $\mu_* > 0$ , then the result (ii) of Lemma 1 implies that  $\nabla_x H(x^k, \mu_k)$  is nonsingular and its norm is uniformly bounded below by a positive constant for all  $k \geq 0$ . In other words, there exists a constant  $c_3 > 0$  such that  $\|[\nabla_x H(x^k, \mu_k)]^{-1}\| \leq c_3$ . By Step 3 of the proposed algorithm, we have

$$\|\Delta x^k\| = \|[\nabla_x H(x^k, \mu_k)]^{-1} H(x^k, \mu_k)\| \leq c_3 \|H(x^k, \mu_k)\| \leq c_3 \beta \mu_k \text{ for } k \geq 0. \quad (14)$$

For  $\alpha \in (0, 1)$ , define

$$r^k(\alpha) := H(x^k + \alpha \Delta x^k, \mu_k) - H(x^k, \mu_k) - \alpha \nabla_x H(x^k, \mu_k) \Delta x^k. \quad (15)$$

It follows from the result (iv) of Lemma 1 and (14) that

$$\|r^k(\alpha)\| \leq \frac{\sqrt{n} \alpha^2 c_1}{\mu_k} \|\Delta x^k\|^2 \leq \sqrt{n} \alpha^2 c_1 c_3^2 \beta \|H(x^k, \mu_k)\|.$$

If we let  $\bar{\alpha} = \min\{\frac{1-\sigma_1}{\sqrt{n} \beta c_1 c_3^2}, 1\}$ , then

$$\|r^k(\alpha)\| \leq (1 - \sigma_1) \alpha \|H(x^k, \mu_k)\| \text{ for any } \alpha \in (0, \bar{\alpha}). \quad (16)$$

Combining (15) and (16), we see

$$\begin{aligned} & \|H(x^k + \alpha \Delta x^k, \mu_k)\| - (1 - \sigma_1 \alpha) \|H(x^k, \mu_k)\| \\ & \leq (1 - \alpha) \|H(x^k, \mu_k)\| + \|r^k(\alpha)\| - (1 - \sigma_1 \alpha) \|H(x^k, \mu_k)\| \\ & = (\sigma_1 - 1) \alpha \|H(x^k, \mu_k)\| + \|r^k(\alpha)\| \\ & \leq 0. \end{aligned} \quad (17)$$

Let  $l_1$  be the minimum integer such that  $\alpha_1^{l_1} \leq \bar{\alpha}$ , then  $\theta_k \geq \alpha_1^{l_1}$  follows from the algorithm. Hence there exists a constant  $\theta_* > 0$  such that  $\theta_k \geq \theta_*$  for all  $k \geq 0$ . It follows from (17) that

$$\|H(x^{k+1}, \mu_k)\| \leq (1 - \sigma_1 \theta_k) \|H(x^k, \mu_k)\| \leq (1 - \sigma_1 \theta_*) \|H(x^k, \mu_k)\|.$$

Using the result (v) of Lemma 1 and the above inequality, we know that, for any  $\lambda \in (0, 1)$ ,

$$\begin{aligned} \frac{\|H(x^{k+1}, (1 - \sigma_2\lambda)\mu_k)\|}{(1 - \sigma_2\lambda)\mu_k} &\leq \frac{\|H(x^{k+1}, \mu_k)\| + \sqrt{n}(\ln \bar{m})\lambda\sigma_2\mu_k}{(1 - \sigma_2\lambda)\mu_k} \\ &\leq \frac{(1 - \sigma_1\theta_*)\|H(x^k, \mu_k)\| + \sqrt{n}(\ln \bar{m})\lambda\sigma_2\mu_k}{(1 - \sigma_2\lambda)\mu_k} \\ &\leq \frac{(1 - \sigma_1\theta_*)\beta + \sqrt{n}(\ln \bar{m})\lambda\sigma_2}{1 - \sigma_2\lambda}. \end{aligned}$$

For  $\frac{(1 - \sigma_1\theta_*)\beta + \sqrt{n}(\ln \bar{m})\lambda\sigma_2}{1 - \sigma_2\lambda} \leq \beta$ ,  $\lambda \leq \bar{\lambda} := \frac{\sigma_1\theta_*\beta}{\sqrt{n}(\ln \bar{m})\sigma_2 + \sigma_2\beta}$ . If we let  $l_2$  be the minimum integer such that  $\alpha_2^{l_2} \leq \min\{\bar{\lambda}, 1\}$ , then a similar argument assures that there exists a constant  $\lambda_* > 0$  such that  $\lambda_k \geq \lambda_*$  for any  $k \geq 0$ . In this case,  $\mu_{k+1} = (1 - \sigma_2\lambda_k)\mu_k \leq (1 - \sigma_2\lambda_*)\mu_k$ , which further implies that  $\mu_k \rightarrow 0$  as  $k \rightarrow \infty$ . This contradicts the hypothesis of  $\mu_* > 0$ . ■

**Lemma 2** Let  $N$  be a vertical block  $P_0$  and  $R_0$  matrix of type  $(m_1, \dots, m_n)$ . Then the solution set  $\mathcal{S}$  of VLCP is nonempty and compact.

*Proof.* Since  $N$  be a vertical block  $P_0$  and  $R_0$  matrix of type  $(m_1, \dots, m_n)$ , Theorem 5.8 in [25] implies that  $\mathcal{S}$  is nonempty. In addition, it is not difficult to see that  $\mathcal{S}$  is closed. Thus, it suffices to show that  $\mathcal{S}$  is bounded. If not, then there exists an unbounded solution sequence  $\{(x^r, s^r)\} \in \mathcal{S}$  for all  $r \geq 0$ . It follows from (6) that

$$H(x^r) = 0 \quad \text{and} \quad s^r = Nx^r + q, \quad \text{for all } r \geq 0.$$

Consequently,  $\lim_{r \rightarrow \infty} \|H(x^r)\|/\|x^r\| = 0$ . However, the result (iii) of Lemma 1 shows that there exists a constant  $c_0 > 0$  such that  $\lim_{r \rightarrow \infty} \|H(x^r)\|/\|x^r\| \geq c_0$ . This contradicts the hypothesis. ■

When  $N$  is a vertical block  $P_0$  and  $R_0$  matrix of type  $(m_1, \dots, m_n)$ , if  $H(x^k) \neq 0$  for all  $k \geq 0$ , the result (i) of Theorem 2 says that Algorithm 3.1 generates a bounded infinite sequence  $\{(x^k, s^k, \mu_k)\}$ . Let  $\{(x^{\bar{k}}, s^{\bar{k}}, \mu_{\bar{k}})\}$  be a convergent subsequence with a limit point  $(x^*, s^*, \mu_*)$ . Then the result (ii) of Theorem 2 says that  $\mu_* = 0$  and  $(x^*, s^*) \in \mathcal{S}$ . By making use of the sequence  $\{(x^{\bar{k}}, s^{\bar{k}}, \mu_{\bar{k}})\}$ , for each  $k \geq 0$ , we define four index sets:

$$\begin{aligned} \bar{A}_k &:= \{i \in \mathcal{I} : x_i^{\bar{k}} > \sqrt{\mu_{\bar{k}}}\}, & \bar{C}_k &:= \{i \in \mathcal{I} : x_i^{\bar{k}} \leq \sqrt{\mu_{\bar{k}}}\}, \\ \bar{B}_k &:= \{(i, j) \in \mathcal{J} : (s^{\bar{k}})^i_j > \sqrt{\mu_{\bar{k}}}\}, & \bar{D}_k &:= \{(i, j) \in \mathcal{J} : (s^{\bar{k}})^i_j \leq \sqrt{\mu_{\bar{k}}}\}. \end{aligned} \tag{18}$$

Clearly,  $\bar{A}_k$  and  $\bar{C}_k$  partition the index set  $\mathcal{I}$  and  $\bar{B}_k$  and  $\bar{D}_k$  partition the index set  $\mathcal{J}$ . In what follows, we discuss the finite termination of Algorithm 3.1 in two cases.

## Case 1: Finite Termination under Strict Complementarity

Assume that  $(x^*, s^*)$  satisfies the strict complementarity condition [25]. Define

$$\begin{aligned}\mathcal{B} &:= \{i \in \mathcal{I} : x_i^* = 0\}, \\ \mathcal{N} &:= \{i \in \mathcal{I} : x_i^* > 0, (s^*)_{j_{i_0}}^i = 0 \text{ for some } j_{i_0} \text{ and } (s^*)_j^i > 0 \text{ for all } j \neq j_{i_0}, 1 \leq j \leq m_i\}.\end{aligned}$$

Since  $(x^*, s^*) \in \mathcal{S}$  satisfies the strict complementarity condition, it follows that  $\mathcal{B} \cup \mathcal{N} = \mathcal{I}$  and  $\mathcal{B} \cap \mathcal{N} = \emptyset$ .

For any  $w = (x, s) \in R^n \times R^m$ , let  $s_{\mathcal{N}_0}^{\mathcal{N}}$  denote a vector with  $i$ th component being  $s_{j_{i_0}}^i$  for  $i \in \mathcal{N}$ . Define

$$G(w) := \begin{pmatrix} s - Nx - q \\ x_{\mathcal{B}} \\ s_{\mathcal{N}_0}^{\mathcal{N}} \end{pmatrix} \quad (19)$$

and

$$\mathcal{S}_0 := \{w \in R^{n+m} : G(w) = 0\}. \quad (20)$$

Similar to Lemma 5.1 in [18], we have the following result:

**Lemma 3** Let

$$\varepsilon = \min \left\{ \min_{i \in \mathcal{N}} \left\{ x_i^*, \min_{1 \leq j \leq m_i, j \neq j_{i_0}} \{(s^*)_j^i\} \right\}, \min_{i \in \mathcal{B}, 1 \leq j \leq m_i} \{(s^*)_j^i\} \right\}$$

and

$$\Delta = \{w = (x, s) \in R^n \times R^m : |x_i - x_i^*| \leq \varepsilon/3, |s_j^i - (s^*)_j^i| \leq \varepsilon/3, i \in \mathcal{I}, (i, j) \in \mathcal{J}\}.$$

Then for any  $w \in \Delta \cap \mathcal{F}$  there exists a constant  $\lambda > 0$  such that

$$\|H(w)\| = \|G(w)\| \geq \lambda \cdot \text{dist}(w, \mathcal{S}_0), \quad (21)$$

where  $H(\cdot)$  and  $G(\cdot)$  are defined by (5) and (19), respectively.

**Proof.** Denote  $(x^*, s^*)$  by  $w^*$ . Since  $G(w^*) = 0$ ,  $G(w) = 0$  is solvable and  $\mathcal{S}_0 \neq \emptyset$ . By Hoffman's result on error bound of linear systems [17], there exists a positive number  $\lambda > 0$  such that for any  $w \in R^{n+m}$

$$\|G(w)\| \geq \lambda \cdot \text{dist}(w, \mathcal{S}_0). \quad (22)$$

For any  $w \in \Delta$ , if  $i \in \mathcal{N}$ , then

$$\begin{aligned}x_i &= x_i^* + x_i - x_i^* \geq \varepsilon - \frac{1}{3}\varepsilon = \frac{2}{3}\varepsilon, \\ |s_{j_{i_0}}^i| &\leq \frac{1}{3}\varepsilon, \\ s_j^i &= (s^*)_j^i + s_j^i - (s^*)_j^i \geq \varepsilon - \frac{1}{3}\varepsilon = \frac{2}{3}\varepsilon \quad \text{for } 1 \leq j \leq m_i, j \neq j_{i_0}.\end{aligned}$$

If  $i \in \mathcal{B}$ , then

$$\begin{aligned} |x_i| &\leq \frac{1}{3}\varepsilon, \\ s_j^i &= (s^*)^i_j + s_j^i - (s^*)^i_j \geq \varepsilon - \frac{1}{3}\varepsilon = \frac{2}{3}\varepsilon \quad \text{for } 1 \leq j \leq m_i. \end{aligned}$$

These imply that for each  $i \in \mathcal{I}$ ,

$$\min\{x_i, s_1^i, s_2^i, \dots, s_{m_i}^i\} = \begin{cases} x_i, & \forall i \in \mathcal{B} \\ s_{j_{i0}}^i, & \forall i \in \mathcal{N}. \end{cases}$$

Hence,

$$\|H(x)\| = \|G(w)\| \quad \text{for } w \in \Delta \cap \mathcal{F}. \quad (23)$$

(21) follows from (23) and (22). ■

Lemma 3 indicates that if  $w \in \mathcal{S}_0$  for  $w$  being sufficiently close to  $w^*$ , then  $w$  solves (1), i.e.,  $w \in \mathcal{S}$ . If we define

$$\bar{\mathcal{S}}_0 := \Delta \cap \mathcal{F} \cap \mathcal{S}_0,$$

then

$$\bar{\mathcal{S}}_0 \subset \mathcal{S},$$

which implies that  $\bar{\mathcal{S}}_0$  is bounded by Lemma 2. Denote  $w^{\bar{k}} := (x^{\bar{k}}, s^{\bar{k}})$ . It is not difficult to see that, for each  $\bar{k} \geq 0$ , there exists a point  $w^{\bar{k}*} = (x^{\bar{k}*}, s^{\bar{k}*}) \in \bar{\mathcal{S}}_0$  such that

$$\|w^{\bar{k}} - w^{\bar{k}*}\| = \min_{w \in \bar{\mathcal{S}}_0} \|w^{\bar{k}} - w\| = \text{dist}(w^{\bar{k}}, \bar{\mathcal{S}}_0). \quad (24)$$

Note that when  $\bar{k}$  is sufficiently large, the point  $w^{\bar{k}*}$  is also the projection of point  $w^{\bar{k}}$  on  $\mathcal{S}_0$ . Therefore, from (21), there exists a constant  $\bar{\rho} > 0$  such that

$$\text{dist}(w^{\bar{k}}, \bar{\mathcal{S}}_0) = \text{dist}(w^{\bar{k}}, \mathcal{S}_0) \leq \bar{\rho} \|H(x^{\bar{k}})\| \quad (25)$$

for any  $\bar{k}$  being sufficiently large.

Corresponding to the solution point  $(x^{\bar{k}*}, s^{\bar{k}*})$ , we define four index sets:

$$\begin{aligned} \bar{A}_k^* &:= \{i \in \mathcal{I} : x_i^{\bar{k}*} > 0\}, & \bar{C}_k^* &:= \{i \in \mathcal{I} : x_i^{\bar{k}*} = 0\}, \\ \bar{B}_k^* &:= \{(i, j) \in \mathcal{J} : (s^{\bar{k}*})^i_j > 0\}, & \bar{D}_k^* &:= \{(i, j) \in \mathcal{J} : (s^{\bar{k}*})^i_j = 0\}. \end{aligned} \quad (26)$$

Obviously,  $\bar{A}_k^*$  and  $\bar{C}_k^*$  form a partition of the index set  $\mathcal{I}$  and  $\bar{B}_k^*$  and  $\bar{D}_k^*$  form a partition of the index set  $\mathcal{J}$ .

The following lemma presents a special property of the four index sets defined above.

**Lemma 4** Let  $N$  be a vertical block  $P_0$  and  $R_0$  matrix of type  $(m_1, \dots, m_n)$  and the index sets  $\bar{A}_k, \bar{B}_k, \bar{C}_k, \bar{D}_k$  and  $\bar{A}_k^*, \bar{B}_k^*, \bar{C}_k^*, \bar{D}_k^*$  be defined as in (18) and (26), respectively. Assume

that the strict complementarity condition holds. If  $H(x^k) \neq 0$  for all  $k \geq 0$ , then  $\bar{A}_k = \bar{A}_k^*$ ,  $\bar{B}_k = \bar{B}_k^*$ ,  $\bar{C}_k = \bar{C}_k^*$ , and  $\bar{D}_k = \bar{D}_k^*$ , when  $\bar{k}$  becomes sufficiently large.

**Proof.** (i) We first show that  $\bar{A}_k \subseteq \bar{A}_k^*$ . For any  $i_0 \in \bar{A}_k$ , i.e.,  $x_{i_0}^{\bar{k}} > \sqrt{\mu_{\bar{k}}}$ , by using  $\|H(x^{\bar{k}}, \mu_{\bar{k}})\| \leq \beta \mu_{\bar{k}}$  and the result (i) of Lemma 1, we have

$$|H_i(x^{\bar{k}})| \leq |H_i(x^{\bar{k}}, \mu_{\bar{k}})| + \mu_{\bar{k}} \ln(m_i + 1) \leq (\beta + \ln(m_i + 1))\mu_{\bar{k}} \text{ for } i \in \mathcal{I}. \quad (27)$$

Thus, for any  $\bar{k}$  being sufficiently large,

$$\begin{aligned} \|x^{\bar{k}} - x^{\bar{k}*}\| &\leq \|w^{\bar{k}} - w^{\bar{k}*}\| = \text{dist}(w^{\bar{k}}, \bar{\mathcal{S}}_0) \quad (\text{by (24)}) \\ &\leq \bar{\rho} \|H(x^{\bar{k}})\| \quad (\text{by (25)}) \\ &= \bar{\rho} \sqrt{\sum_{i=1}^n |H_i(x^{\bar{k}})|^2} \\ &\leq c_5 \mu_{\bar{k}}, \quad (\text{by (27)}) \end{aligned} \quad (28)$$

where  $c_5 := \bar{\rho}(\sum_{i=1}^n [\beta + \ln(m_i + 1)]^2)^{1/2}$ . Remembering that  $x_{i_0}^{\bar{k}} > \sqrt{\mu_{\bar{k}}}$ , we have

$$x_{i_0}^{\bar{k}*} \geq x_{i_0}^{\bar{k}} - c_5 \mu_{\bar{k}} > \sqrt{\mu_{\bar{k}}} - c_5 \mu_{\bar{k}} > 0, \quad (29)$$

where the first inequality follows from (28) and the third inequality from  $\lim_{\bar{k} \rightarrow \infty} \mu_{\bar{k}} = 0$ . Obviously, (29) implies  $i_0 \in \bar{A}_k^*$  and, consequently,  $\bar{A}_k \subseteq \bar{A}_k^*$ .

Next we show that  $\bar{A}_k^* \subseteq \bar{A}_k$ . For any  $i_0 \in \bar{A}_k^*$ , i.e.,  $x_{i_0}^{\bar{k}*} > 0$ , from the proof of Theorem 2, we know

$$\lim_{\bar{k} \rightarrow \infty} \mu_{\bar{k}} = 0,$$

and, by (28),

$$\|x^{\bar{k}} - x^{\bar{k}*}\| \rightarrow 0.$$

Consequently,

$$\|x^{\bar{k}*} - x^*\| \leq \|x^{\bar{k}*} - x^{\bar{k}}\| + \|x^{\bar{k}} - x^*\| \rightarrow 0$$

as  $\bar{k} \rightarrow 0$ . Since  $x_{i_0}^{\bar{k}*} > 0$ , by the proof of Lemma 3, there is a constant  $\xi > 0$  such that  $x_{i_0}^{\bar{k}*} \geq \xi > 0$ , for  $\bar{k}$  being sufficiently large. Note that  $\mu_{\bar{k}} \rightarrow 0$  as  $k \rightarrow \infty$ . We know  $x_{i_0}^{\bar{k}} > \sqrt{\mu_{\bar{k}}}$  for  $k$  being sufficiently large. This implies  $i_0 \in \bar{A}_k$  and hence  $\bar{A}_k^* \subseteq \bar{A}_k$ .

(ii) Since  $(s^{\bar{k}})_j^i = N_j^i x^{\bar{k}} + q_j^i$  and  $(s^{\bar{k}*})_j^i = N_j^i x^{\bar{k}*} + q_j^i$ , it follows that

$$|(s^{\bar{k}})_j^i - (s^{\bar{k}*})_j^i| = |N_j^i x^{\bar{k}} - N_j^i x^{\bar{k}*}| \leq \|N\| \|x^{\bar{k}} - x^{\bar{k}*}\| \leq c_5 \|N\| \mu_{\bar{k}}.$$

Thus, similar to the proof of (i), we can show that  $\bar{B}_k = \bar{B}_k^*$  for  $\bar{k}$  being sufficiently large.

(iii) From (i), (ii) and

$$\begin{aligned} \bar{A}_k \cup \bar{C}_k &= \mathcal{I} = \bar{A}_k^* \cup \bar{C}_k^*, & \bar{A}_k \cap \bar{C}_k &= \emptyset = \bar{A}_k^* \cap \bar{C}_k^*, \\ \bar{B}_k \cup \bar{D}_k &= \mathcal{J} = \bar{B}_k^* \cup \bar{D}_k^*, & \bar{B}_k \cap \bar{D}_k &= \emptyset = \bar{B}_k^* \cap \bar{D}_k^*, \end{aligned}$$

it is easy to see that  $\bar{C}_k = \bar{C}_k^*$  and  $\bar{D}_k = \bar{D}_k^*$ . ■

The following result is a direct consequence of Lemma 4:

**Corollary 1** In the same setting of Lemma 4, there exists a constant  $c_6 > 0$  such that, for any  $\bar{k}$  being sufficiently large,

$$x_i^{\bar{k}} \geq c_6 \text{ for all } i \in \bar{A}_k \quad \text{and} \quad (s^{\bar{k}})_j^i \geq c_6 \text{ for all } (i, j) \in \bar{B}_k.$$

The following lemma characterizes a solution to VLCP in terms of the index sets. Since the proof is simple, we omit it.

**Lemma 5**  $(x^{\bar{k}^*}, s^{\bar{k}^*}) \in \mathcal{S}$  if and only if  $x^{\bar{k}^*} \geq 0$ ,  $s^{\bar{k}^*} = Nx^{\bar{k}^*} + q \geq 0$  and one of the following conditions is satisfied:

- (i) If  $\bar{A}_k^* \neq \emptyset$ , then for any  $i \in \bar{A}_k^*$  there exists at least one index  $(i, j) \in \mathcal{J}$  such that  $(s^{\bar{k}^*})_j^i = 0$ . In this case,  $\bar{B}_k^*$  can be either empty or nonempty.
- (ii) If  $\bar{A}_k^* = \emptyset$ , then  $q_{\bar{B}_k^*} > 0$  and  $q_{\bar{D}_k^*} = 0$  when  $\bar{B}_k^* \neq \emptyset$ , and  $q_{\bar{D}_k^*} = 0$  when  $\bar{B}_k^* = \emptyset$ .

Now we are ready to show the following main result:

**Theorem 3** Let  $N$  be a vertical block  $P_0$  and  $R_0$  matrix of type  $(m_1, \dots, m_n)$ . Assume that the strict complementarity condition holds. Then Algorithm 3.1 terminates with an exact solution to VLCP in a finite number of iterations.

**Proof.** Suppose that Algorithm 3.1 does not terminate in a finite number of generation, but instead generates an infinite sequence  $\{(x^k, s^k, \mu_k)\}$ . Then we know  $H(x^k) \neq 0$  for all  $k \geq 0$  and the stopping criteria in Step 2 are inactive all the time. Otherwise, if  $H(x^{k_0}) = 0$  for some  $k_0 \geq 0$ , then, by noting that  $s^k = Nx^k + q$  for all  $k \geq 0$ , we have  $(x^{k_0}, s^{k_0}) \in \mathcal{S}$  and the algorithm terminates here. From Theorem 2, the sequence  $\{(x^k, s^k, \mu_k)\}$  is bounded and hence has a convergent subsequence by  $\{(x^{\bar{k}}, s^{\bar{k}}, \mu_{\bar{k}})\}$ . Let the index sets  $\bar{A}_k, \bar{B}_k, \bar{C}_k$ , and  $\bar{D}_k$  be defined as in (18). Then one of the following three cases will happen:

- (i)  $\bar{A}_k \neq \emptyset, \bar{B}_k \neq \emptyset$ .
- (ii)  $\bar{A}_k \neq \emptyset, \bar{B}_k = \emptyset$ .
- (iii) Either  $\bar{A}_k = \emptyset, \bar{B}_k \neq \emptyset$  or  $\bar{A}_k = \emptyset = \bar{B}_k$ .

From Lemmas 4 and 5, if case (i) happens, then, for any  $i \in \bar{A}_k$ , we have  $i \in \bar{A}_k^*$ , i.e.,  $x_i^{\bar{k}^*} > 0$ . Since  $(x^{\bar{k}^*}, s^{\bar{k}^*}) \in \mathcal{S}$  when  $\bar{k}$  is sufficiently large, there exists an index  $(i, j) \in \mathcal{J}$  such that  $(s^{\bar{k}^*})_j^i = 0$ . Consequently,  $(i, j) \in \bar{D}_k^*$ . Therefore,  $(i, j) \in \bar{D}_k$ , i.e.,  $(s^{\bar{k}})_j^i \leq \sqrt{\mu_{\bar{k}}}$  holds for all  $\bar{k}$  being sufficiently large.

Suppose that case (i) indeed happens at infinitely many  $\bar{k}$ . Since  $\bar{A}_k, \bar{C}_k$  form a partition of the index set  $\mathcal{I}$ , and  $\bar{B}_k, \bar{D}_k$  form a partition of the index set  $\mathcal{J}$ , the equation  $s^{\bar{k}} = Nx^{\bar{k}} + q$  can be written as

$$\begin{pmatrix} s_{\bar{B}_k}^{\bar{k}} \\ s_{\bar{D}_k}^{\bar{k}} \end{pmatrix} = \begin{pmatrix} N_{\bar{B}_k \bar{A}_k} & N_{\bar{B}_k \bar{C}_k} \\ N_{\bar{D}_k \bar{A}_k} & N_{\bar{D}_k \bar{C}_k} \end{pmatrix} \begin{pmatrix} x_{\bar{A}_k}^{\bar{k}} \\ x_{\bar{C}_k}^{\bar{k}} \end{pmatrix} + \begin{pmatrix} q_{\bar{B}_k} \\ q_{\bar{D}_k} \end{pmatrix}. \quad (30)$$

Consider the subsequence  $\{(x^{\bar{k}}, s^{\bar{k}}, \mu_{\bar{k}})\}$ , (10) becomes

$$\begin{pmatrix} s_{\bar{B}_k}^{\bar{k}} + \Delta s_{\bar{B}_k}^{\bar{k}} \\ 0 \end{pmatrix} = \begin{pmatrix} N_{\bar{B}_k \bar{A}_k} & N_{\bar{B}_k \bar{C}_k} \\ N_{\bar{D}_k \bar{A}_k} & N_{\bar{D}_k \bar{C}_k} \end{pmatrix} \begin{pmatrix} x_{\bar{A}_k}^{\bar{k}} + \Delta x_{\bar{A}_k}^{\bar{k}} \\ 0 \end{pmatrix} + \begin{pmatrix} q_{\bar{B}} \\ q_{\bar{D}} \end{pmatrix}. \quad (31)$$

Subtracting (30) from (31) yields

$$\begin{pmatrix} \Delta s_{\bar{B}_k}^{\bar{k}} \\ -s_{\bar{D}_k}^{\bar{k}} \end{pmatrix} = \begin{pmatrix} N_{\bar{B}_k \bar{A}_k} & N_{\bar{B}_k \bar{C}_k} \\ N_{\bar{D}_k \bar{A}_k} & N_{\bar{D}_k \bar{C}_k} \end{pmatrix} \begin{pmatrix} \Delta x_{\bar{A}_k}^{\bar{k}} \\ -x_{\bar{C}_k}^{\bar{k}} \end{pmatrix}. \quad (32)$$

Let  $I$  and  $0$  denote the identity matrix and zero matrix with appropriate dimensionality, respectively. Define

$$y^{\bar{k}} := ((\Delta x_{\bar{A}_k}^{\bar{k}})^T, (\Delta s_{\bar{B}_k}^{\bar{k}})^T)^T, \quad z^{\bar{k}} := ((x_{\bar{C}_k}^{\bar{k}})^T, (s_{\bar{D}_k}^{\bar{k}})^T)^T, \\ P := \begin{pmatrix} -N_{\bar{B}_k \bar{A}_k} & I \\ -N_{\bar{D}_k \bar{A}_k} & 0 \end{pmatrix}, \quad Q := \begin{pmatrix} -N_{\bar{B}_k \bar{C}_k} & 0 \\ -N_{\bar{D}_k \bar{C}_k} & I \end{pmatrix}$$

Then (32) becomes

$$Py^{\bar{k}} = Qz^{\bar{k}}. \quad (33)$$

Therefore, when  $\bar{k}$  is sufficiently large, the system (33) is solvable for  $y^{\bar{k}}$ . By applying Gaussian elimination on (33), the linearly dependent rows and columns of  $P$  can be eliminated. Let  $\bar{P}$  be a largest possible nonsingular submatrix of  $P$  and  $\bar{y}^{\bar{k}}$  be the corresponding variable. Then we need to solve

$$\bar{P}\bar{y}^{\bar{k}} = \bar{Q}\bar{z}^{\bar{k}},$$

where the rows of  $\bar{Q}$  correspond to the rows of  $\bar{P}$ . Since  $\bar{P}$  is nonsingular, we have

$$\|\bar{y}^{\bar{k}}\| = \|\bar{P}^{-1}\bar{Q}\bar{z}^{\bar{k}}\| \leq \|\bar{P}^{-1}\| \|\bar{Q}\| \|\bar{z}^{\bar{k}}\|. \quad (34)$$

Noting that the definitions of  $\bar{C}_k, \bar{D}_k, z^{\bar{k}}$  and the fact that  $\bar{z}^{\bar{k}}$  is a subvector of  $z^{\bar{k}}$ , it is not difficult to see that there exists a constant  $c_7 > 0$  such that  $\|\bar{z}^{\bar{k}}\| \leq c_7 \sqrt{\mu_{\bar{k}}}$ . Moreover,  $\|\bar{P}^{-1}\|$  is bounded above (see [24, 30]). It follows from (34) that there exists a constant  $c_8 > 0$  such that  $\|\bar{y}^{\bar{k}}\| \leq c_8 \sqrt{\mu_{\bar{k}}}$ . Let the components of  $y^{\bar{k}}$  that were removed during Gaussian elimination be zero. Then there exists a solution to (33), denoted by  $((\Delta x_{\bar{A}_k}^{\bar{k}})^T, (\Delta s_{\bar{B}_k}^{\bar{k}})^T)^T$ , such that

$$\begin{cases} |\Delta x_i^{\bar{k}}| \leq c_9 \sqrt{\mu_{\bar{k}}} & \forall i \in \bar{A}_k \\ |\Delta (s^{\bar{k}})_j^i| \leq c_9 \sqrt{\mu_{\bar{k}}} & \forall (i, j) \in \bar{B}_k, \end{cases} \quad (35)$$

where  $c_9 > 0$  is a constant. By using Corollary 1 and the definitions of  $\bar{A}_k, \bar{B}_k$ , we know that

$$\begin{cases} x_i^{\bar{k}} > c_6 & \forall i \in \bar{A}_k \\ (s^{\bar{k}})_j^i > c_6 & \forall (i, j) \in \bar{B}_k. \end{cases} \quad (36)$$

Combining (35), (36), and the fact that  $\lim_{\bar{k} \rightarrow \infty} \mu_{\bar{k}} = 0$ , we know

$$x_{\bar{A}_k}^{\bar{k}} + \Delta x_{\bar{A}_k}^{\bar{k}} > 0 \quad \text{and} \quad x_{\bar{B}_k}^{\bar{k}} + \Delta x_{\bar{B}_k}^{\bar{k}} > 0$$

for all  $\bar{k}$  being sufficiently large. This indicates that one of the stopping criteria in Step 2 is met for some sufficiently large  $\bar{k}$ . This is a contradiction. Hence  $H(x^{k_0}) = 0$  for some  $k_0 \geq 0$ .

Similar arguments can be developed for cases (ii) and (iii). Hence Algorithm 3.1 terminates in a finite number of iterations.

Following Theorem 3.1, the proposed algorithm finds an exact solution to VLCP when it terminates. ■

## Case 2: Finite Termination under Singleton Assumption

This time let us assume that the solution set  $\mathcal{S}$  of (1) is a singleton, say,

$$\mathcal{S} = \{(x^*, s^*)\}.$$

We will show that Algorithm 3.1 terminates with the unique solution even without the strict complementarity assumption.

Using  $(x^*, s^*)$ , we define four index sets:

$$\begin{aligned} A^* &:= \{i \in \mathcal{I} : x_i^* > 0\}, & C^* &:= \{i \in \mathcal{I} : x_i^* = 0\}, \\ B^* &:= \{(i, j) \in \mathcal{J} : (s^*)_j^i > 0\}, & D^* &:= \{(i, j) \in \mathcal{J} : (s^*)_j^i = 0\}. \end{aligned} \quad (37)$$

Again,  $A^*$  and  $C^*$  form a partition of the index set  $\mathcal{I}$ , and  $B^*$  and  $D^*$  form a partition of the index set  $\mathcal{J}$ .

Since  $\mathcal{S} = \{(x^*, s^*)\}$ , the result (vi) of Lemma 1 assures that there is constant  $c_{10} > 0$  such that

$$\|x^{\bar{k}} - x^*\| \leq c_{10} \|H(x^{\bar{k}})\|.$$

Following a similar proof of Lemma 4, we can show the following result:

**Lemma 6** Let  $N$  be a vertical block  $P_0$  and  $R_0$  matrix of type  $(m_1, \dots, m_n)$  and the index sets  $\bar{A}_k, \bar{B}_k, \bar{C}_k, \bar{D}_k$  and  $A^*, B^*, C^*, D^*$  be defined by (18) and (37), respectively. Assume that the solution set of (1) is a singleton. If  $H(x^k) \neq 0$  for all  $k \geq 0$ , then  $\bar{A}_k = A^*$ ,  $\bar{B}_k = B^*$ ,  $\bar{C}_k = C^*$ , and  $\bar{D}_k = D^*$ , when  $\bar{k}$  becomes sufficiently large.

Furthermore, using a similar proof of Theorem 3, we have the following main theorem:

**Theorem 4** Let  $N$  be a vertical block  $P_0$  and  $R_0$  matrix of type  $(m_1, \dots, m_n)$ . Assume that the solution set of (1) is a singleton. Then Algorithm 3.1 terminates at the unique solution of VLCP in a finite number of iterations.

## 5 Numerical Results

To test the performance and illustrate the potential of the proposed method, we have implemented Algorithm 3.1 in MATLAB on a 1000 MHz Pentium III personal computer running Linux. The eight test problems found in Peng and Lin [25] (some of them are from the literature [8, 26]) were used in our computational experiment. For easy comparison, the order of these eight problems is kept the same as in Peng and Lin's paper. For all test problems, the vertical block matrices are  $P_0$ -matrices. Moreover, Problems 1 to 5 hold the strict complementarity assumption. Problems 6, 7, and 8 were modified so that they may not be strictly complementary, but they do satisfy the singleton assumption.

The following parameters were chosen for all test problems:  $\sigma_1 = 0.005$ ,  $\sigma_2 = 0.001$ ,  $\alpha_1 = 0.9$ ,  $\alpha_2 = 0.85$ ,  $\gamma = 1.0\text{e-}3$ ,  $p = 1.0$ ,  $\mu_0 = 0.0005$ , and  $\beta = \|H(x^0, \mu_0)\|/\mu_0 + 1.0\text{e-}5$ . An initial point  $x^0$  was set to be a contact vector  $(a, \dots, a)^T \in R^n$ . We used the criterion  $\|H(x^k)\|_1 \leq 1.0\text{e-}20$ , where  $\|\cdot\|_1$  denotes  $l_1$ -norm, to stop the algorithm.

Table 1 shows our test results of Algorithm 3.1. The second and the third columns indicate the dimension of the problem and the constant used for the initial point, respectively. The  $l_1$ -norm of  $H(\cdot)$  at the initial point is given in column 4. Then,  $k^*$  in column 5 denotes the number of iterations required to achieve the result given in column 6. Last column shows under which conditions each run is terminated.

Several observations can be made here:

1. For every test problem, the proposed method indeed finds a solution point meeting the desired accuracy in very few iterations. In particular, the exact solutions have been found in many cases, when  $\|H(x^k)\|_1 = 0$ . Compared with other known methods as reported in [8, 25, 26], our method converges in fewer iterations to achieve the known results.
2. Problems 6 and 8 were studied by Peng and Lin [25] only, but they were not able to solve these two problems effectively. Our results are much better. Exact solutions have been found in less than four iterations.
3. The last column of Table 5.1 shows that the algorithm terminates either by satisfying the condition at Case (i) of Step 2 or by meeting the stopping criterion. Furthermore, whenever the stopping criterion is met, the algorithm finds an exact solution. This supports the finite termination results proved in Section 4.
4. In our experiments, we see an overflow problem may occur in (7) when the exponential function  $\exp(-x_i/\mu)$  or  $\exp(-(N_j^i x + q_j^i)/\mu)$  is computed with a very large (negative) argument. But this potential problem can be handled effectively by using the following equality:

$$H_i(x, \mu) = -\mu \ln \left( \exp\left(-\frac{x_i + h_i}{\mu}\right) + \sum_{j=1}^{m_i} \exp\left(-\frac{(N_j^i x + q_j^i) + h_i}{\mu}\right) \right) + h_i, \quad (38)$$

where  $h_i \leq \min\{x_i, N_1^i x + q_1^i, \dots, N_{m_i}^i x + q_{m_i}^i\}$ .

Problem	$n$	$a$	$\ H(x^0)\ _1$	$k^*$	$\ H(x^{k^*})\ _1$	Termination
1	2	1.0	1.0	2	0.0	Step 2 Case (i)
		10.0	10.0	2	0.0	Step 2 Case (i)
		-10.0	22.0	3	0.0	Step 2 Case (i)
2	6	1.0	9.0	2	0.0	Stopping criterion
		10.0	18.0	3	0.0	Stopping criterion
		-10.0	215.0	3	0.0	Stopping criterion
3	50	5.0	86.5	2	2.2e-16	Step 2 Case (i)
	100	5.0	171.5	2	1.1e-15	Step 2 Case (i)
	200	5.0	341.5	2	3.4e-15	Step 2 Case (i)
	100	-5.0	1506	4	1.3e-15	Step 2 Case (i)
	200	-5.0	3005.9	4	2.9e-15	Step 2 Case (i)
4	50	5.0	56.0	2	1.8e-12	Step 2 Case (i)
	100	5.0	106.0	2	7.8e-12	Step 2 Case (i)
	200	5.0	206.0	2	7.3e-11	Step 2 Case (i)
	100	-5.0	2080.0	4	0.0	Stopping criterion
	200	-5.0	4180.0	4	0.0	Stopping criterion
5	50	5.0	52.9	3	0.0	Stopping criterion
	100	5.0	102.95	4	0.0	Step 2 Case (i)
	200	5.0	202.97	3	0.0	Stopping criterion
	100	-5.0	2095.0	5	0.0	Step 2 Case (i)
	200	-5.0	4195.0	5	0.0	Step 2 Case (i)
6	6	1.0	10.0	2	0.0	Step 2 Case (i)
		10.0	16.0	3	0.0	Step 2 Case (i)
		-10.0	215.0	4	0.0	Step 2 Case (i)
7	6	1.0	10.0	2	0.0	Step 2 Case (i)
		10.0	16.0	1	0.0	Stopping criterion
		-10.0	215.0	3	0.0	Step 2 Case (i)
8	6	1.0	9.0	3	0.0	Step 2 Case (i)
		10.0	18.0	3	0.0	Step 2 Case (i)
		-10.0	223.0	2	0.0	Step 2 Case (i)

Table 1: The performance of the proposed method for the test problems in [25].

## 6 Conclusion

In this paper we have proposed an entropy function based smoothing Newton method for solving vertical linear complementarity problems. It has been shown that the proposed method finds an exact solution in a finite number of iterations under either the strict complementarity assumption or singleton assumption. This result is more general than those reported. The computational results we obtained can also confirm the theoretic findings and illustrated its potential.

## References

- [1] D.P. Bertsekas, Approximation procedures based on the method of multipliers, *J. Optim. Theory Appl.*, 23(1977): 487–510.
- [2] J. Burke and S. Xu, The global linear convergence of a non-interior path-following algorithm for linear complementarity problem, *Math. Oper. Res.*, 23(3)(1998): 719–734.
- [3] P.L. Chang, A Minimax Approach to Nonlinear Programming, *Ph.D. dissertation*, Department of Mathematics, University of Washington, Seattle, WA, 1980.
- [4] B. Chen and X. Chen, A global linear and local quadratic continuation smoothing method for variational inequalities with box constraints, *Comput. Optim. Appl.*, 13(2000): 131–158.
- [5] R.W. Cottle and G.B. Dantzig, A generalization of the linear complementarity problem, *Journal of Combinatorial Theory*, 8(1970): 79–90.
- [6] R.W. Cottle, J.S. Pang, and R.E. Stone, *The Linear Complementarity Problem*, Academic Press, Boston, MA, 1992.
- [7] X. Chen and Y. Ye, On homotopy-smoothing methods for variational inequalities, *SIAM J. Control Optim.*, 37(1999): 589–616.
- [8] A.A. Ebiefung, Nonlinear mappings associated with the generalized linear complementarity problem, *Math. Program.*, 69(1995): 255–268.
- [9] A.A. Ebiefung and M.M. Kostreva, The generalized Leontief input-output model and its application to the choice of the new technology, *Ann. Oper. Res.*, 44(1993): 161–172.
- [10] S. Engelke and C. Kanzow, Predictor-corrector smoothing methods for the solution of linear programming, Preprint, Department of Mathematics, University of Hamburg, Germany, March 2000.
- [11] S.-C. Fang, J.R. Rajasekera and H.-S. J. Tsao, *Entropy Optimization and Mathematical Programming*, Kluwer Academic Publishers, Boston/London/Dordrecht, 1997.
- [12] A. Fischer and C. Kanzow, On the finite termination of an iterative method for linear complementarity problems, *Math. Program.*, 74(1996): 279–292.

- [13] T. Fujisawa and E.S. Kuh, Piecewise-linear theory of nonlinear networks, *SIAM J. Appl. Math.*, 22(1972): 307–328.
- [14] A.A. Goldstein, Chebyshev approximation and linear inequalities via exponentials, Report, Department of Mathematics, University of Washington, Seattle, WA, 1997.
- [15] M.S. Gowda and R. Sznajder, The generalized order linear complementarity problem, *SIAM J. Matrix Anal. Appl.*, 15(1994), 779–795.
- [16] M.S. Gowda and R. Sznajder, A generalization of the Nash equilibrium theorem on bimatrix games, *Internat. J. Game Theory*, 25(1996): 1–12.
- [17] A.J. Hoffman, On approximate solutions of systems of linear equalities, *J. Research Nat. Bur. Standards*, 49(1952), 263–265.
- [18] Z.H. Huang, L. Qi and D. Sun, Sub-quadratic convergence of a smoothing Newton algorithm for the  $P_0$ - and monotone LCP, *Manuscript*, August, 2001. <http://www.math.nus.edu.sg/~matsundf>.
- [19] T. Illés, J.M. Peng, C. Roos and T. Terlaky, A strongly polynomial procedure yielding a maximally complementarity solution for  $P_*(\kappa)$  linear complementarity problems, *SIAM J. Optim.*, 11(2)(2001): 320–340.
- [20] B.W. Kort and D.P. Bertsekas, A new penalty function for constrained minimization, *Proceedings of the 1972 IEEE Conference on Decision and Control*, New Orleans, Louisiana, 1972.
- [21] X.S. Li, An aggregate function method for nonlinear programming, *Sci. China (Ser. A)*, 34(12)(1991): 1467–1473.
- [22] O.L. Mangasarian, Generalized linear complementarity problems as linear programming, *Oper. Res. Verfahren*, 31(1979): 393–402.
- [23] S.R. Mohan, S.K. Neogy, and R. Sridhar, The generalized linear complementarity problem revisited, *Math. Program.*, 74(1996): 197–218.
- [24] S. Mehrotra and Y. Ye, On finding the optimal facet of linear programs, *Math. Program.*, 62(1993): 497–515.
- [25] J.M. Peng and Z. Lin, A non-interior continuation method for generalized linear complementarity problems, *Math. Program.*, 86(1999): 533–563.
- [26] H.D. Qi and L.Z. Liao, A smoothing Newton method for extended vertical linear complementarity problems, *SIAM J. Matrix Anal. Appl.*, 21(1999): 45–66.
- [27] H.D. Qi, L.Z. Liao and Z. Lin, Regularized smoothing approximations to vertical nonlinear complementarity problems, *J. Math. Anal. Appl.*, 230(1999): 261–276.
- [28] L. Qi and D. Sun, Improving the convergence of non-interior point algorithm for nonlinear complementarity problems, *Math. Comput.*, 69(2000), 283–304.
- [29] L. Qi, D. Sun and G. Zhou, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequality problems, *Math. Program.*, 87(2000), 1–35.

- [30] A. Schrijver, Theory of Linear and Integer Programming, Wiley, New York, 1986.
- [31] M. Sun, Singular control problems in bounded intervals, *Stochastics*, 21(1987): 303–344.
- [32] M. Sun, Monotonicity of Mangasarian’s iterative algorithm for the generalized linear complementarity problem, *J. Math. Anal. Appl.*, 144(1989): 473–485.
- [33] D. Sun, J. Han and Y.B. Zhao, On the finite termination of the damped-Newton algorithm for the linear complementarity problem, *Acta Mathematica Applicatae Sinica*, 21(1)(1998): 148–154.
- [34] R. Sznajder and M.S. Gowda, Generalizations of  $P_0$  and  $P$ –properties, extended vertical and horizontal LCP’s, *Linear Algebra Appl.*, 223/224(1995): 695–716.
- [35] P. Tseng, Analysis of a non-interior continuation method based on Chen-Mangasarian smoothing functions for complementarity problems, in *Reformulation-Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, edited by M. Fukushima, and L. Qi, Kluwer Academic Publishers, Boston, 1999, pp. 381–404.
- [36] P. Tseng, Error bounds and superlinear convergence analysis of some Newton-type methods in optimization, in *Nonlinear Optimization and Related Topics*, edited by G. Di Pillo and F. Giannessi, Kluwer Academic Publishers, Boston, 2000, pp. 445–462.
- [37] P. Tseng and D.P. Bertsekas, On the convergence of the exponential multiplier method for convex programming, *Math. Program.*, 60(1993): 1–19.
- [38] Y. Ye, On the finite convergence of interior-point algorithms for linear programming, *Math. Program.*, 57(1992): 325–335.