

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600



Non-Interior Path-Following Methods for
Complementarity Problems

by

Song Xu

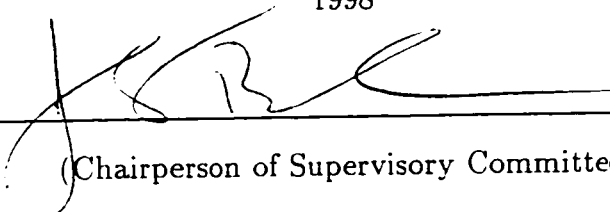
A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

University of Washington

1998

Approved by _____


(Chairperson of Supervisory Committee)

Program Authorized

to Offer Degree _____

Mathematics

Date _____

Aug-7 1998

UMI Number: 9907965

**Copyright 1998 by
Xu, Song**

All rights reserved.

**UMI Microform 9907965
Copyright 1998, by UMI Company. All rights reserved.**

**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

© Copyright 1998

Song Xu

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to University Microfilms, 1490 Eisenhower Place, P. O. Box 975, Ann Arbor, MI 48106, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature Song

Date Aug. 7, 1978

University of Washington

Abstract

Non-Interior Path-Following Methods for
Complementarity Problems

by Song Xu

Chairperson of Supervisory Committee: *Professor James V. Burke*
Department of Mathematics

Because of its excellent numerical performance, non-interior path following methods (also called smoothing methods) have become an important class of methods for solving complementarity problems. However, no rate of convergence results are available for these methods. In this thesis, we bridge this gap between the theory and the practical performance of the methods. Specifically, we focus on the rates of convergence, the complexity, and the implementation of non-interior path following methods.

The thesis introduces new notions of neighborhoods of the central path for non-interior path following methods for linear complementarity problems. These neighborhoods are modeled on similar concepts from the interior point literature and are used to adjust the value of a continuation parameter. However, these neighborhoods are fundamentally different from those used in the interior-point methods. In particular the solution set of the underlying LCP is contained in the interior of these neighborhoods relative to the affine constraints. The new neighborhood concepts have proven to be fundamental for both the theoretical analysis of the algorithms and in their practical implementation. With these new neighborhood concepts, we are able

to establish the first global linear convergence result for non-interior path following methods. In order to accelerate the convergence, we introduce a predictor-corrector strategy. This strategy allows us to construct the first predictor-corrector non-interior path following method that is both globally linearly convergent and locally quadratically convergent. In the thesis, we also make progress toward understanding the computational complexity of these methods. Complexity results are obtained from both the algorithmic and condition-based perspectives. The complexity bounds that we establish are the only results for these methods that are available to date. These results represent a first step toward understanding the complexity of non-interior path following methods.

TABLE OF CONTENTS

List of Figures	iii
Chapter 1: Introduction	1
1.1 Complementarity Problems and Examples	3
1.2 Numerical Methods for Solving Linear Complementarity Problem . .	9
1.3 Numerical Methods for Solving Nonlinear Complementarity Problems	13
1.4 Organization of the Thesis	18
Chapter 2: The Structural Properties of Non-Interior Path Following Methods For LCP	20
2.1 The Smoothing Functions and Their Properties	20
2.2 A New Notion of Neighborhoods	30
2.3 A Prototype Algorithm for Non-Interior Path Following Methods . .	35
2.4 An Example	38
Chapter 3: The Global and Local Convergence of A Predictor-Corrector Non-Interior Path-Following Method for LCP	42
3.1 Predictor and Corrector Search Directions	43
3.2 A Predictor-Corrector Non-Interior Path-Following Algorithm	46
3.3 The Global Linear Convergence	50
3.4 The Local Quadratic Convergence	55
3.5 The Global Convergence	57

Chapter 4:	The Condition-Based Complexity of Non-Interior Path Following Methods for LCP	58
4.1	A Global Error Bound	59
4.2	A Non-Interior Path Following Method based on Centering Direction	63
4.3	Global Linear Convergence	67
4.4	A Condition-Based Complexity Bound for LCP with a P -Matrix	68
4.5	A Condition-Based Complexity for LCP with a Symmetric Positive Definite Matrix	73
Chapter 5:	A Polynomial Time Infeasible Interior Point Algorithm for Monotone LCP	77
5.1	The Rescaled Newton Directions	78
5.2	The Algorithm	81
5.3	Global Linear Convergence and Polynomial Complexity	85
5.4	The Connection to Standard Interior Point Methods	89
Chapter 6:	Numerical Experiments	91
6.1	The Test Problems	92
6.2	An Implementation Based on the Combined Direction	93
6.3	An Implementation Based on the Centering Direction	96
6.4	An Infeasible Implementation Based on the Centering Direction	98
6.5	Conclusions	101
Bibliography		105

LIST OF FIGURES

2.1	The slices of $\mathcal{N}^x(\beta, \mu)$ with $\beta = 4$ and $\mu = 15, 10, 5$	39
2.2	The slices of $\mathcal{N}_{\text{int}}^x(\beta, \mu)$ with $\beta = 0.5$ and $\mu = 15, 10, 5$	40

ACKNOWLEDGMENTS

I would like to express my sincere appreciation to my thesis adviser, Professor James V. Burke for his guidance during my graduate study. Over the past three years, he has been a constant source of enthusiasm, advice and encouragement. I am also grateful to Professor R. T. Rockafellar and Paul Tseng for their support and advice. I thank my wife Yu for her tremendous support and love. Finally, I gratefully acknowledge the financial support I have received as a graduate student. The research and preparation of this dissertation were supported in part by the National Sciences Foundation grant number DMS-9303772 under the supervision of Professor Burke, by the NSF grant number CCR-9311621 under the supervision of Professor Tseng, and by the McFarlan Fellowship from the Department of Mathematics at the University of Washington.

To Yu (Carrie) and William

Chapter 1

INTRODUCTION

Over the past three decades, the field of complementarity problems has seen a rapid development in its theory, algorithms and applications. On one hand, complementarity problems provide a unified framework to study many traditional optimization problems [20, 61]. One such example is linear complementarity problem (LCP) which can be used to unify linear and quadratic programs as well as the bimatrix game. On the other hand, complementarity problems have found diverse applications in engineering, economics, and sciences, in particular, in various equilibrium models [25, 22]. Because of the importance of complementarity problems, enormous effort has been made by the mathematical programming community in developing efficient and robust algorithms for solving these problems. Over the years, many promising approaches have been proposed and significant progress has been achieved [61, 35]. Among them, the non-interior path following methods (also called smoothing methods) have attracted a lot of attention recently [5, 3, 4, 8, 7, 9, 10, 12, 13, 14, 15, 16, 31, 37, 38, 42, 41, 43, 64, 65, 78, 83, 84, 85]. Similar to interior point methods, non-interior path-following methods also follow a path, namely the central path. But unlike interior-point methods, non-interior path following methods can be initiated from any point in the space and do not require the iterates to stay in the interior of the nonnegative orthant. As a result, non-interior path following methods are more flexible and very convenient for numerical implementation. This thesis studies the interaction between interior-point methods

and non-interior path following methods with a focus on the underline ideas behind these two types of methods. It represents a significant contribution to the structural properties, the rates of convergence, the complexity and the implementation of non-interior path following methods.

First, the thesis brings a new notion of neighborhoods for the central path into non-interior path following methods. In the early stages of the development of non-interior path-following methods, there was no systematic procedure to update the smoothing parameter and no rate of convergence results available. This deficiency is overcome in this thesis by introducing a new notion of neighborhoods for the central path that is suitable for non-interior path following methods. Just as with interior point methods, these neighborhoods are used to adjust the value of the continuation parameter. The neighborhood concept has proven to be fundamental for both the theoretical analysis of the algorithms and in their practical implementation. The expansion of the neighborhood concept for the central path has wide spread consequences for these and other methods. The thesis also studies two important structural properties of these neighborhoods. It establishes a boundedness property which yields the boundedness of the sequence of iterates generated by the methods. It also proves that the solution set of linear complementarity problem is contained in the interior of the neighborhoods relative to the affine constraints. Thus, these neighborhoods are fundamentally different from those used in the interior-point literature, where the solution set necessarily lies on the boundary of the neighborhood relative to the affine constraints. This property of the solution set makes it possible to adapt a wide range of Newton step ideas to accelerate the local rate of convergence.

Second, the thesis establishes the first global linear convergence result for non-interior path following methods for the linear complementarity problems in the following two cases: (a) the problem has a P_0 and R_0 matrix, and (b) the problem is monotone and has a feasible interior point. In addition, the thesis introduces a predictor-corrector strategy to non-interior path-following methods and proposes

the first predictor-corrector non-interior path following method that is both globally linearly convergent and locally quadratically convergent.

The thesis has made some progress in the computational complexity of non-interior path following methods. Two types of complexity have been studied, the condition-based complexity and the polynomial complexity. The thesis has obtained a condition-based complexity for a non-interior path following method for LCP with a symmetric positive definite matrix or with a P -matrix. It also establishes a polynomial complexity bound for an interior-point variant of the methods. These results are the only results available to date and represent a first step toward understanding the complexity of non-interior path following methods.

Finally, the thesis implements three different versions of non-interior path-following methods. The preliminary numerical results demonstrate that non-interior path following methods are extremely promising.

1.1 Complementarity Problems and Examples

Complementarity problems have different formulations. The classical *nonlinear complementarity problem*, is defined by

NCP(F): Find $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ satisfying

$$F(x) - y = 0, \quad (1.1)$$

$$x \geq 0, y \geq 0, x^T y = 0, \quad (1.2)$$

where $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a nonlinear function. NCP(F) is called *monotone* if F is a monotone mapping, i.e.,

$$(x - y)^T (F(x) - F(y)) \geq 0, \quad (1.3)$$

for all $x, y \in \mathbb{R}^n$. When F is an affine mapping, i.e., $F(x) = Mx + q$ for some $M \in \mathbb{R}^{n \times n}$ and $q \in \mathbb{R}^n$, the problem reduces to the *linear complementarity problem*

LCP(q, M): Find $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ satisfying

$$Mx - y + q = 0, \quad (1.4)$$

$$x \geq 0, y \geq 0, x^T y = 0. \quad (1.5)$$

LCP(q, M) is called *monotone* if M is positive semi-definite. The lack of symmetry in LCP(q, M) can be remedied by introducing the horizontal linear complementarity problem where x and y play an equal role. The *horizontal linear complementarity problem* is defined as follows

HLCP(q, M, N): Find $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ satisfying

$$Mx + Ny + q = 0, \quad (1.6)$$

$$x \geq 0, y \geq 0, x^T y = 0, \quad (1.7)$$

where $M, N \in \mathbb{R}^{n \times n}$ and $q \in \mathbb{R}^n$. HLCP(q, M, N) is called *monotone* if for any $u, v \in \mathbb{R}^n$ the following condition holds:

$$Mu + Nv = 0 \implies u^T v = 0. \quad (1.8)$$

In many applications, some of the underlying conditions are defined by a system of nonlinear equations, while the complementarity conditions are only applied to some of the variables and functions. This leads to the *mixed nonlinear complementarity problem*, which is defined as

MCP(F): Find $x \in \mathbb{R}^{n+m}$ satisfying

$$x_I \geq 0, F_I(x) \geq 0, x_I^T F_I(x) = 0, \quad (1.9)$$

$$F_J(x) = 0, \quad x_J \text{ free}, \quad (1.10)$$

where $I = \{1, 2, \dots, n\}$ and $J = \{1, 2, \dots, m\}$. When $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n+m}$ is an affine mapping, i.e., $F(x) = Mx + q$ for some

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \in \mathbb{R}^{(n+m) \times (n+m)},$$

where $M_{11} \in \mathbb{R}^{n \times n}$, $M_{22} \in \mathbb{R}^{m \times m}$, $M_{12} \in \mathbb{R}^{n \times m}$ and $M_{21} \in \mathbb{R}^{m \times n}$, and

$$q = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \in \mathbb{R}^{n+m}.$$

where $q_1 \in \mathbb{R}^n$ and $q_2 \in \mathbb{R}^m$, the problem reduces to the *mixed linear complementarity problem* which is defined by

MLCP(q, M): Find $(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$ satisfying

$$\begin{bmatrix} y \\ 0 \end{bmatrix} = M \begin{bmatrix} x \\ z \end{bmatrix} + q, \quad (1.11)$$

$$x \geq 0, y \geq 0, x^T y = 0. \quad (1.12)$$

MLCP(q, M) is called *monotone* if M is positive semi-definite. NCP(F) and MCP(F) are special cases of the *box constrained variational inequality problem*, which is defined as follows

VI($F, [l, u]$): Find $x \in [l, u]$ satisfying

$$(y - x)^T F(x) \geq 0, \quad \text{for all } y \in [l, u], \quad (1.13)$$

where l and u are n -dimensional vectors with $l_i \in [-\infty, \infty)$ and $u_i \in (l_i, \infty]$, and

$$[l, u] = \{x \in \mathbb{R}^n : l \leq x \leq u\}.$$

The condition (1.13) can be rewritten using the normal cone notation.

Definition 1.1.1 Given a closed convex set $C \subseteq \mathbb{R}^n$, the normal cone $N_C(x)$ to the set C at a point $x \in C$ is defined by

$$N_C(x) := \{z \mid \langle z, y - x \rangle \leq 0 \text{ for all } y \in C\}, \quad (1.14)$$

and $N_C(x) := \emptyset$ if $x \notin C$.

With this notation, condition (1.13) is equivalent to

$$-F(x) \in N_{[l, u]}(x). \quad (1.15)$$

It is sometimes convenient to write (1.13) in the form

$$\begin{aligned} F(x) - x_1 + x_2 &= 0, \\ x_1 &\geq 0, \quad z - l \geq 0, \quad x_1^T(z - l) = 0, \\ x_2 &\geq 0, \quad u - z \geq 0, \quad x_2^T(u - z) = 0. \end{aligned} \quad (1.16)$$

It is easy to see that if $l_i = 0$ and $u_i = \infty$ for $i = 1, 2, \dots, n$, then $\text{VI}(F, [0, \infty])$ is equivalent to $\text{NCP}(F)$. If $l_i = 0, u_i = \infty$ for $i \in I$ and $l_i = -\infty, u_i = \infty$ for $i \in J$, then $\text{VI}(F, [l, u])$ is exactly $\text{MCP}(F)$.

Extensive documentation of applications of complementarity problems in engineering, economics, and sciences can be found in Ferris and Pang [25] and Dirkse and Ferris [22]. Here we list a few such examples.

Example 1.1.2 *Convex Quadratic Programming*

Consider the convex quadratic program (QP)

$$\begin{aligned} &\text{minimize} && c^T x + \frac{1}{2} x^T Q x \\ &\text{subject to} && Ax \geq b, \\ &&& x \geq 0, \end{aligned} \quad (1.17)$$

where $Q \in \mathbb{R}^{n \times n}$ is symmetric positive semi-definite, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

The dual for (1.17) is

$$\begin{aligned} &\text{maximize} && c^T y - \frac{1}{2} x^T Q x \\ &\text{subject to} && A^T y - Q x \leq c, \\ &&& y \geq 0. \end{aligned} \quad (1.18)$$

If x is an optimal solution of the program (1.17), then there exists a vector $y \in \mathbb{R}^m$ such that the pair (x, y) satisfies the Karush-Kuhn-Tucker conditions

$$\begin{aligned} u &= c + Qx - A^T y \geq 0, & x &\geq 0, & x^T u &= 0, \\ v &= -b + Ax \geq 0, & y &\geq 0, & y^T v &= 0. \end{aligned} \quad (1.19)$$

The conditions in (1.19) define a monotone LCP(q, M) where

$$q = \begin{bmatrix} c \\ -b \end{bmatrix} \quad \text{and} \quad M = \begin{bmatrix} Q & -A^T \\ A & 0 \end{bmatrix}. \quad (1.20)$$

When $Q = 0$, convex quadratic program (1.17) reduces to a linear program. In this case, the matrix

$$M = \begin{bmatrix} 0 & -A^T \\ A & 0 \end{bmatrix} \quad (1.21)$$

is skew-symmetric and positive semi-definite.

Example 1.1.3 Convex Quadratic Program with Box Constraints

Consider the convex quadratic program with box constraints

$$\begin{aligned} \text{minimize} \quad & q(z) := c^T z + \frac{1}{2} z^T Q z, \\ \text{subject to} \quad & l \leq z \leq u, \end{aligned} \quad (1.22)$$

where $Q \in \mathbb{R}^{n \times n}$ is symmetric positive semi-definite, $l = (l_i), u = (u_i) \in \mathbb{R}^n$ are two fixed vectors with $l_i < u_i$ for $i = 1, \dots, n$. If x is an optimal solution of program (1.22), then the following Karush-Kuhn-Tucker condition holds

$$-\nabla q(z) \in N_{[l, u]}(x). \quad (1.23)$$

Condition (1.23) is equivalent to the conditions:

$$\begin{aligned} Qz + b - x_1 + x_2 &= 0, \\ x_1 &\geq 0, \quad z - l \geq 0, \quad x_1^T(z - l) = 0, \\ x_2 &\geq 0, \quad u - z \geq 0, \quad x_2^T(u - z) = 0. \end{aligned} \quad (1.24)$$

Example 1.1.4 *Extended Linear-Quadratic Programming*

The extended linear-quadratic programming was introduced by Rockafellar as a modeling tool for optimization under uncertainty [71], multistage optimization [70] and optimal control problems [69]. Given nonempty, polyhedral sets $X \in \mathbb{R}^n$ and $Y \in \mathbb{R}^m$, vectors $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$, and positive semi-definite matrices $C \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{m \times m}$, the extended linear-quadratic program may be defined in either a primal form or dual form as follows:

$$(\mathcal{P}) \quad \text{minimize}_{x \in X} \quad c^T x + \frac{1}{2} x^T C x + \rho_{Y,B}(b - Ax), \quad (1.25)$$

$$(\mathcal{D}) \quad \text{maximize}_{y \in Y} \quad b^T x - \frac{1}{2} y^T B y - \rho_{X,C}(A^T y - c), \quad (1.26)$$

where $\rho_{Y,B}$ is the monitoring function defined by

$$\rho_{Y,B}(u) = \sup_{y \in Y} \{u^T y - \frac{1}{2} y^T B y\}. \quad (1.27)$$

The Lagrangian for this problem is:

$$L(x, y) = c^T x + \frac{1}{2} x^T C x + b^T y - \frac{1}{2} y^T B y - y^T A x, \quad (1.28)$$

which is defined for $(x, y) \in X \times Y$. If x is an optimal solution to (\mathcal{P}) and y is an optimal solution to (\mathcal{D}) , then from the optimality conditions, we have

$$-\nabla_x L(x, y) \in N_X(x), \quad \nabla_y L(x, y) \in N_Y(y). \quad (1.29)$$

When $X = \mathbb{R}^n$ and $Y = Y_1 \times \dots \times Y_m$ with $Y_i = [l_i, u_i]$ and $l_i < u_i$ for $i = 1, 2, \dots, m$, the optimality conditions can be written as

$$\begin{aligned} Cx - A^T y + c &= 0, \\ Ax + By - w_1 + w_2 - b &= 0, \\ w_1 \geq 0, y - l \geq 0, w_1^T (y - l) &= 0, \\ w_2 \geq 0, u - y \geq 0, w_2^T (u - y) &= 0. \end{aligned} \quad (1.30)$$

1.2 Numerical Methods for Solving Linear Complementarity Problem

In the next two sections, we give a brief description of existing numerical methods for solving linear complementarity problems and nonlinear complementarity problems. The description is not intended to be comprehensive. It only serves the purpose of providing background and motivation for the methods to be studied in this thesis.

The early numerical methods for solving linear complementarity problem are pivoting methods which include the principal pivoting methods [17, 18, 19, 21] and Lemke's method [48]. A detailed description of these methods and other pivoting methods can be found in Cottle, Pang and Stone [20].

1.2.1 Interior-point methods for LCP

Since the publication of Karmarkar's paper [58] in 1984, the field of interior-point methods saw rapid development and expansion. The excitement of interior point methods are due partly to a theoretical property of the algorithms, the polynomial complexity, and due also to their excellent practical performance in solving large scale problems. Most of the early work on interior-point methods was devoted to solving linear programming problems. However, in the past few years, there was intense research on interior-point methods for monotone LCP. There are many variations of interior-point methods. A complete description of these methods can be found in Kojima, Megiddo, Noma and Yoshise [44] for linear complementarity problems, and in Wright [81] for linear programming problems. Here we give a brief description of interior-point path following methods for solving $LCP(q, M)$.

Interior-point path following methods for solving LCP are typically designed to follow the path in the positive orthant, $\mathbb{R}_{++}^n \times \mathbb{R}_{++}^n$, determined by the equations $F_{\theta_\mu}(x, y) = 0$ for $\mu > 0$ where the function $F_{\theta_\mu} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^n$ is given by

$$F_{\theta_\mu}(x, y) := \begin{bmatrix} Mx - y + q \\ \Theta_\mu(x, y) \end{bmatrix}, \quad (1.31)$$

with $\theta_\mu(a, b) = ab - \mu$ and

$$\Theta_\mu(x, y) = \begin{bmatrix} \theta_\mu(x_1, y_1) \\ \dots \\ \theta_\mu(x_n, y_n) \end{bmatrix}. \quad (1.32)$$

This path is called the *central path* [44] and is denoted by \mathcal{C} , that is

$$\mathcal{C} := \{(x, y) : Mx - y + q = 0, Xy = \mu e \text{ with } 0 < x, 0 < y, \text{ and } 0 < \mu\}, \quad (1.33)$$

where following standard usage in the interior point literature, e is the vector each of whose components is 1 and $X = \text{diag}(x)$ the diagonal matrix with i^{th} diagonal entry x_i for $i = 1, 2, \dots, n$. The interior-point path following methods attempt to follow the central path by applying Newton's method to the equations $F_{\theta_\mu}(x, y) = 0$ for decreasing values of μ and by restricting all the iterates to a pre-specified neighborhood of the central path in the interior of the nonnegative orthant. For each $\beta \in (0, 1)$, the β neighborhood of \mathcal{C} is defined to be the set

$$\mathcal{N}_{int}(\beta) := \left\{ (x, y) : Mx - y + q = 0, \frac{\|Xy - \mu e\|}{\mu} \leq \beta, 0 < x, 0 < y, 0 < \mu \right\}. \quad (1.34)$$

Here, the subscript *int* is used to indicate that this is the neighborhood employed in the interior point literature for LCP(q, M). Very loosely speaking, standard interior-point path following methods start with an initial point (x^0, y^0) lying in $\mathcal{N}_{int}(\beta)$ for some $\beta \in (0, 1)$ and an initial value μ_0 satisfying $\|X^0 y^0 - \mu_0 e\| \leq \mu_0 \beta$. A Newton direction $(\Delta x, \Delta y)$ at (x^0, y^0) based on the equations

$$Mx - y + q = 0, Xy = \mu_1 e \quad (1.35)$$

is computed for some $\mu_1 \in (0, \mu_0)$ and the Newton step is then damped to ensure that the update (x^1, y^1) remains strictly positive and satisfies $\|X^1 y^1 - \mu_1 e\| \leq \beta \mu_1$. This process is then iterated to termination. The trick is to implement the method so to ensure the existence of a sequence μ_k converging linearly to zero and satisfying

$\|X^k y^k - \mu_k e\| \leq \mu_k \beta$. This yields the linear convergence of the vector $X^k y^k$ to zero which in turn provides the basis for complexity results. Numerous variations on this basic plan have been proposed. The most notable of which are the infeasible (i.e. $Mx^k - y^k + q$ may or may not be the zero vector) predictor-corrector strategies for which local super-linear convergence can also be established (e.g. see [76, 79, 80, 86]).

1.2.2 Non-interior path following methods for LCP

Non-interior path following methods also follow the central path, but the iterates do not necessarily reside in the positive orthant. The first non-interior path following method for LCP was developed by Chen and Harker [9] and was based on the Chen-Harker-Kanzow-Smale (CHKS) smoothing function (Chen and Harker [9], Kanzow [42] and Smale [72])

$$\phi_\mu(a, b) = a + b - \sqrt{(a - b)^2 + 4\mu^2}. \quad (1.36)$$

Later Kanzow [42] developed non-interior path following methods based on the smoothed Fischer-Burmeister function

$$\psi_\mu(a, b) = a + b - \sqrt{a^2 + b^2 + 2\mu^2}, \quad (1.37)$$

where $a, b \in \mathbb{R}$ and $\mu > 0$ is the smoothing parameter. Other smoothing function classes include the Chen and Magasarian [13] smoothing functions and the Gabriel and Moré [31] smoothing functions. These smoothing functions will be discussed in the next Chapter.

It is easy to show that $\phi_\mu(a, b) = 0$ (or $\psi_\mu(a, b) = 0$) if and only if $0 \leq a$, $0 \leq b$, and $ab = \mu^2$. Thus, the functions ψ_μ and v_μ have an advantage over the function θ_μ which makes them well suited to non-interior path following methods. That is, the condition $\phi_\mu(a, b) = 0$ or $(\psi_\mu(a, b) = 0)$ guarantees the non-negativity of the arguments a and b .

Using ϕ_μ and ψ_μ as building blocks, one defines the functions

$$F_{\phi_\mu}(x, y) := \begin{bmatrix} Mx - y + q \\ \Phi_\mu(x, y) \end{bmatrix}, \quad (1.38)$$

$$F_{\psi_\mu}(x, y) := \begin{bmatrix} Mx - y + q \\ \Psi_\mu(x, y) \end{bmatrix}, \quad (1.39)$$

$$\Phi_\mu(x, y) := \begin{bmatrix} \phi_\mu(x_1, y_1) \\ \dots \\ \phi_\mu(x_n, y_n) \end{bmatrix}, \quad (1.40)$$

and

$$\Psi_\mu(x, y) := \begin{bmatrix} \psi_\mu(x_1, y_1) \\ \dots \\ \psi_\mu(x_n, y_n) \end{bmatrix}. \quad (1.41)$$

Clearly, a point (x, y) is on the central path if and only if $F_{\phi_\mu}(x, y) = 0$ (or $F_{\psi_\mu}(x, y) = 0$).

The reasons for the growing interest in non-interior methods are: (1) these methods can be initiated from any point in the space and therefore are more flexible and very convenient for numerical implementation, (2) these methods are ideally suited for application to nonlinear complementarity problems where the interiority restriction on the iterates is quite severe, and (3) the numerical evidence on the efficiency of these methods is very impressive. However, no algorithmic convergence results are available for non-interior path following methods. The absence of any rate of convergence results for these algorithms is due to the somewhat ad hoc rules for updating the continuation parameter μ . This gap is bridged in this thesis by introducing a new notion of neighborhood for the central path that is suitable for the non-interior path following methods. Just as with interior point methods, this neighborhood is used to adjust the value of the continuation parameter μ between iterations in a manner that insures the linear convergence of the values $\|\Phi_{\mu_k}(x^k, y^k)\|$ (or $\|\Psi_{\mu_k}(x^k, y^k)\|$) to

zero. The new notion of neighborhood has proved to be fundamental for both the theoretical analysis and the practical implementation of non-interior path following methods.. With this notion, we are able to establish the first global linear convergence of a non-interior path following method (see Chapter 3) and study its complexity (see Chapter 4).

1.3 Numerical Methods for Solving Nonlinear Complementarity Problems

While there are many approaches available for solving nonlinear complementarity problem [61, 35, 1], the most successful approaches are based on Newton-type methods. Here we focus on three such approaches: the nonsmooth and semismooth equation approaches, the interior-point path following approach, and the non-interior path following approach.

1.3.1 Nonsmooth and semismooth equation approaches

In the nonsmooth and semismooth equation approaches for NCP, the nonlinear complementarity problem $\text{NCP}(F)$ is reformulated as a system of nonsmooth equations or semismooth equations. The classical Newton's method for smooth equations are then generalized to solve the nonsmooth or the semismooth equations.

General descriptions of nonsmooth equation approach can be found in Pang [59, 60], Pang and Qi [63], Harker and Pang [35] and Harker and Xiao [36]. Various nonsmooth equation formulations for $\text{NCP}(F)$ have been proposed. Here we only give one such formulation which is based on the *min*-mapping

$$H(x) := \min\{x, F(x)\}. \quad (1.42)$$

It is easy to see that x solves $\text{NCP}(F)$ if and only if $H(x) = 0$. However, the mapping is not (Fréchet) differentiable whenever $x_i = f_i(x)$ for some i . To handle

this difficulty, Newton-type methods using the notion of B-derivatives were proposed. For descriptions of these methods, see Pang [59, 60], Gabriel and Pang [32], Pang and Gabriel [62], Pang and Qi [63].

A well-known semismooth equation reformulation of $\text{NCP}(F)$ is based on the Fischer-Burmeister function $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ defined by

$$\psi(a, b) = (a + b) - \sqrt{a^2 + b^2}. \quad (1.43)$$

It is easy to see that

$$\psi(a, b) = 0 \iff a \geq 0, b \geq 0, ab = 0. \quad (1.44)$$

Therefore x solves $\text{NCP}(F)$ if and only if

$$\Psi(x, F(x)) = 0, \quad (1.45)$$

where for $x, y \in \mathbb{R}^n$,

$$\Psi(x, y) := \begin{bmatrix} \psi(x_1, y_1) \\ \dots \\ \psi(x_n, y_n) \end{bmatrix}. \quad (1.46)$$

Note that the mapping Ψ is not (Fréchet) differentiable whenever $x_i = y_i = 0$ for some i but it is semismooth (see Mifflin [54], Qi and Sun [66]). The function ψ (1.43) was introduced by Fischer [27] and later studied in a number of recent articles (Fischer [27, 28], Facchinei and Soares [24, 23], De Luca, Facchinei and Kanzow [49], Geiger and Kanzow [33], Jiang and Qi [39], Kanzow [40], Tseng [77]). A surprising property of the function ψ is that the natural merit function $\|\Psi(x, y)\|^2$ is smooth, so the global convergence through a linesearch procedure can easily be enforced. A review of the use of the function ψ can be found in Fischer [28].

1.3.2 The interior-point path following methods for NCP

The interior-point methods for $\text{NCP}(F)$ try to follow the central path

$$\mathcal{C} := \{(x, y) : F(x) - y = 0, Xy = \mu e \text{ with } 0 < x, 0 < y, \text{ and } 0 < \mu\}, \quad (1.47)$$

by restricting the iterates to a neighborhood of the central path in the positive orthant defined by

$$\mathcal{N}_{int}(\beta) = \{ (x, y) : \|F(x) - y\| + \|Xy - \mu e\| \leq \beta\mu, 0 < x, 0 < y, 0 < \mu \} , \quad (1.48)$$

for some $\beta \in (0, 1)$. This process is then iterated to termination. This framework for solving $NCP(F)$ was first suggested by McLinden [53] and further studied by Kojima, Megiddo and Noma [45], and Kojima, Mizuno and Noma [46, 47]. Interior-point algorithms within this framework have been developed by Wright and Ralph [82] and Tseng [78].

1.3.3 Non-interior path following methods for NCP

The non-interior path following methods (also called smoothing methods) have attracted a lot of attention recently. During the past two years, many non-interior path following methods have been proposed and they distinguish each other in the use of smoothing functions, the neighborhoods, the search directions and their convergence properties.

The smoothing functions used in non-interior path following methods include the smoothed Fischer-Burmeister function (1.36), the Chen-Harker-Kanzow-Smale smoothing function (1.37), the Chen-Magasarian smoothing functions [13] and the Gabriel and Moré smoothing functions [31]. Non-interior path following methods inherit some nice features possessed by nonsmooth-equation approaches and interior-point path following approaches. On one hand, the function ϕ_μ reduces to the min-function (1.42) and ψ_μ reduces to the Fischer-Burmeister function (1.43) when $\mu = 0$ and therefore the functions ϕ_μ (1.36) and ψ_μ (1.37) can be considered as the smoothed version of the min-function and the function ψ (1.43) respectively. On the other hand, similar to interior-point path following methods, non-interior path following methods also follow a path, the central path. The key distinction between the system (1.35) and the system $F_{\phi_\mu}(x, y) = 0$ (1.38) (or $F_{\psi_\mu}(x, y) = 0$ (1.39)) is that a solution to

the system (1.35) may not be strictly positive and so may not lie on the central path. This is the reason why one must initiate interior point methods at strictly positive points and then damp the Newton steps to maintain this property. However, a solution to (1.38) (or (1.39)) must be strictly positive and so will lie on the central path. Thus, the non-negativity of any limit point is automatically assured without imposing additional non-negativity constraints. This is one of the reasons why the functions F_{ϕ_μ} , F_{ψ_μ} and other smoothing functions are so effective in formulating non-interior path following methods.

The first non-interior path following method was developed by Chen and Harker [9] and the method was further studied by Kanzow [42, 41], Kanzow and Jiang [43], Chen and Harker [9, 10]. All these methods are based on either the smoothed Fischer-Burmeister function or the Chen-Harker-Kanzow-Smale smoothing function. A broader class of smoothing functions was introduced by Chen and Magasarian [13]. The Chen-Magasarian class of smoothing functions is derived from the double integration of parameterized probability density function. The Chen-Harker-Kanzow-Smale smoothing function turns out to be a member of the Chen-Magasarian class of smoothing functions. The class of Chen-Magasarian smoothing functions has been used in designing smoothing methods for solving nonlinear and mixed complementarity problems [13] and for solving convex inequalities and LCP [14]. The Chen-Magasarian class of smoothing functions is further extended to a larger class of smoothing functions, called Gabriel and Moré smoothing functions, which are useful in designing smoothing methods for solving variational inequalities [31]. The existence and limiting behavior of the smoothing paths have been studied in Chen and Harker [11], Hotta and Yoshise [37] and recently by Gowda and Tawhid [34].

The connection between non-interior path following methods and interior point path following methods was established by Xu and Burke [85], where the authors established a polynomial complexity bound for an interior-point path following method based on Chen-Harker-Kanzow-Smale smoothing function and the smoothed Fischer-

Burmeister function. The first global linear convergence result for a non-interior path following method for LCP was established in Burke and Xu [5]. The key in obtaining the global linear convergence is the introduction of a new notion of neighborhood for the central path which is suitable for the non-interior path following method. The neighborhood concept was extended to nonlinear complementarity problem by Xu [83]. He also established the first global linear convergence result for a non-interior path following method for NCP. A different notion of neighborhood was introduced by Hotta and Yoshise [37] for the monotone NCP. They also studied some structural properties and establish the global convergence for their algorithm. The first local superlinear convergence result for non-interior methods was established in Chen, Qi, and Sun [15]. The announcement of the papers [15, 5, 83, 85, 37] has initiated a flurry of activity on rate convergence analysis for non-interior path following methods based on different smoothing techniques.

The papers [8, 7, 12] modify the notion of neighborhoods in [5, 83, 37] and establish the global linear convergence of their non-interior path following algorithms. In addition, they introduce the idea of an *Approximate Newton Step* to obtain local quadratic or superlinear convergence of their methods. In [12], Chen and Xiu compute both a centering step and an approximate Newton step based on a single matrix factorization. If the approximate Newton step performs better than the centering step, then the new iterate is based on the approximate Newton step. In [8], only the approximate Newton step is used but two backtracking line searches are required to obtain both global linear and local superlinear convergence. In [7], Chen and Chen use a new technique to dynamically update the neighborhood of the central path in order to establish global convergence and local superlinear convergence. Xu [84] studies the boundedness properties of the neighborhoods, the stopping criterion, and provides some preliminary complexity results for a non-interior path-following method for monotone LCP. In [64], Qi and Sun develop a non-interior path following algorithm that uses the neighborhood ideas developed by Hotta and Yoshise [37].

Conditions are given under which the algorithm is globally linearly convergent, or globally convergent and locally superlinearly convergent. Chen and Ye [16] build on the work in [15] and develop a hybrid smoothing Newton method that is globally convergent, locally superlinearly convergent, and possesses a finite termination property for linear variational inequality problems. Jiang [38] develops a generalized Newton method and Gauss-Newton method for complementarity problem and establishes the global and local superlinear convergence. In [65], Qi, Sun, and Zhou apply techniques from non-smooth equations to obtain the global convergence and local superlinear convergence of a smoothing method based on Robinson's normal equations [68]. In [78], Tseng obtains global linear convergence using the neighborhood concept for the central path. The search direction is a combination of a centering step with a step designed to accelerate the method locally. In addition, an active set strategy is introduced that allows one to establish the local superlinear convergence of the method under very mild conditions.

1.4 Organization of the Thesis

The basic components of non-interior path following methods will be laid out in Chapter 2. In particular, we will review the existing smoothing functions, introduce a new notion of neighborhoods, and provide a prototype algorithm for non-interior path following methods and compare it with the interior-point methods. We also prove two important structural properties for the neighborhoods. The first is a boundedness property for the neighborhoods which yields the boundedness of the sequence of the iterates generated by the methods. The second property says that the solution set of the linear complementarity problem is contained in the interior of the neighborhoods relative to the affine constraints. Thus, these neighborhoods are fundamentally different from those used in the interior-point literature, where the solution set necessarily lies on the boundary of the neighborhood relative to the affine constraints.

In Chapter 3, we introduce a predictor-corrector strategy to non-interior path following methods. Based on this strategy, we propose a predictor-corrector path following algorithm for LCP and show that the algorithm is both globally linearly convergent and locally quadratically convergent if either the problem is monotone and there is a feasible interior point, or the matrix M in $LCP(q, M)$ is P_0 and R_0 .

Chapters 4 and 5 are devoted to studying the complexity of non-interior path following methods with Chapter 4 focusing on the condition-based complexity and Chapter 5 the polynomial complexity. In Chapter 4, a condition-based complexity was established for a non-interior path following method for linear complementarity problems with the matrix M being either a symmetric positive definite matrix or a P -matrix. In Chapter 5, we introduce a rescaling technique to the Newton directions used in non-interior path following methods. Based on the rescaled Newton directions, we propose an interior-point path following method for monotone LCP and establish a polynomial complexity of the method.

Finally in Chapter 6, we present some preliminary numerical results on three implementations of non-interior path-following methods for LCP. The implementations are based on the algorithms studied in Chapters 3 and 4. The numerical results show that non-interior path-following methods are very promising.

Chapter 2

THE STRUCTURAL PROPERTIES OF NON-INTERIOR PATH FOLLOWING METHODS FOR LCP

The plan of this chapter is as follows. In section 2.1, we review various smoothing functions used in non-interior path following methods, including the smoothed Fischer-Burmeister function, the Chen-Harker-Kanzow-Smale smoothing function, the Chen-Magasarian smoothing functions, the Gabriel and Moré smoothing functions.

In section 2.2, we introduce a new notion of neighborhoods into non-interior path following methods and study two important structural properties of these neighborhoods. The first is a boundedness property which yields the boundedness of the sequence of iterates generated by the methods. The second property states that the solution set of the linear complementarity problem is contained in the interior of the neighborhood relative to the affine constraints.

In section 2.3, we give a prototype algorithm of non-interior path following methods and compare it with its counterpart in interior-point path following methods.

Finally in section 2.4, we use an example to illustrate the concepts and ideas introduced in this chapter with a focus on the differences between non-interior and interior-point methods.

2.1 The Smoothing Functions and Their Properties

The smoothing functions studied in this section include the smoothed Fischer-Burmeister function, the Chen-Harker-Kanzow-Smale smoothing function, the Chen-Magasarian

smoothing functions, and the Gabriel and Moré smoothing functions.

2.1.1 The Smoothed Fischer-Burmeister function and the Chen-Harker-Kanzow-Smale smoothing function

Recall the smoothed Fischer-Burmeister function defined in (1.37)

$$\psi_\mu(a, b) = a + b - \sqrt{a^2 + b^2 + 2\mu^2}, \quad (2.1)$$

and the Chen-Harker-Kanzow-Smale smoothing function defined in (1.36)

$$\phi_\mu(a, b) = a + b - \sqrt{(a - b)^2 + 4\mu^2}, \quad (2.2)$$

where $a, b \in \mathbb{R}$ and $\mu > 0$ is the smoothing parameter. From time to time, we may want to consider μ as a variable as well as a smoothing parameter. For this purpose, we use the notation

$$\psi(a, b, \mu) := a + b - \sqrt{a^2 + b^2 + 2\mu^2}, \quad (2.3)$$

and

$$\phi(a, b, \mu) := a + b - \sqrt{(a - b)^2 + 4\mu^2}. \quad (2.4)$$

It is easy to show that $\psi(a, b, \mu) = \psi_\mu(a, b) = 0$ (or $\phi(a, b, \mu) = \phi_\mu(a, b) = 0$) if and only if $0 \leq a$, $0 \leq b$, and $ab = \mu^2$.

In the following lemma, we list some of the properties of ψ_μ for later use.

Lemma 2.1.1 *The smoothed Fischer-Burmeister function $\psi_\mu : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ has the following properties:*

1. For every $\mu > 0$, the function ψ_μ is continuously differentiable on \mathbb{R}^2 with

$$\frac{\partial \psi_\mu(a, b)}{\partial a} = 1 - \frac{a}{\sqrt{a^2 + b^2 + 2\mu^2}} \quad \text{and} \quad \frac{\partial \psi_\mu(a, b)}{\partial b} = 1 - \frac{b}{\sqrt{a^2 + b^2 + 2\mu^2}}.$$

2. For any $a, b \in \mathbb{R}$ and $\mu > 0$, one has

$$\|\nabla^2 \psi_\mu(a, b)\|_\infty \leq \frac{\sqrt{2}}{\mu}. \quad (2.5)$$

3. For $\mu_1 \geq 0$, $\mu_2 \geq 0$ and $a, b \in \mathbb{R}$, we have

$$|\psi_{\mu_1}(a, b) - \psi_{\mu_2}(a, b)| \leq \sqrt{2}|\mu_1 - \mu_2|. \quad (2.6)$$

4. For any $a, b \in \mathbb{R}$ and $\mu > 0$, one has

$$0 < \frac{\partial \psi_\mu(a, b)}{\partial a} < 2, \quad 0 < \frac{\partial \psi_\mu(a, b)}{\partial b} < 2, \quad (2.7)$$

and

$$\left(\frac{\partial \psi_\mu(a, b)}{\partial a}\right)^2 + \left(\frac{\partial \psi_\mu(a, b)}{\partial b}\right)^2 \geq 3 - 2\sqrt{2} > 0. \quad (2.8)$$

Proof Part 1 is trivial. Part 3 is given by Kanzow [42] and Part 4 is taken from Proposition 2.3 in Fukushima, Luo and Pang [30]. It remains to prove that Part 2 holds.

It is easy to check that

$$\nabla^2 \psi_\mu(a, b) = \frac{1}{\sqrt{a^2 + b^2 + 2\mu^2}} \begin{bmatrix} -\frac{b^2 + 2\mu^2}{a^2 + b^2 + 2\mu^2} & \frac{ab}{a^2 + b^2 + 2\mu^2} \\ \frac{ab}{a^2 + b^2 + 2\mu^2} & -\frac{a^2 + 2\mu^2}{a^2 + b^2 + 2\mu^2} \end{bmatrix}.$$

Since

$$\begin{aligned} \left| \frac{1}{\sqrt{a^2 + b^2 + 2\mu^2}} \right| &\leq \frac{1}{\sqrt{2}\mu}, & \left| \frac{b^2 + 2\mu^2}{a^2 + b^2 + 2\mu^2} \right| &\leq 1, \\ \left| \frac{a^2 + 2\mu^2}{a^2 + b^2 + 2\mu^2} \right| &\leq 1, & \left| \frac{ab}{a^2 + b^2 + 2\mu^2} \right| &\leq 1. \end{aligned}$$

we have $\|\nabla^2 \psi_\mu(a, b)\|_\infty \leq \frac{\sqrt{2}}{\mu}$. □

We now list some properties of the the function $\psi(a, b, \mu)$ for later use.

Lemma 2.1.2 *The smoothed Fischer-Burmeister function $\psi : \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{++} \rightarrow \mathbb{R}$ defined in (2.3) has the following properties:*

(i) The function $\psi(a, b, \mu)$ is continuously differentiable on $\mathbb{R}^2 \times \mathbb{R}_{++}$.

(ii) The function $\psi(a, b, \mu)$ is concave on $\mathbb{R}^2 \times \mathbb{R}_{++}$.

(iii) For any $(a, b, \mu) \in \mathbb{R}^2 \times \mathbb{R}_{++}$, we have

$$\|\nabla^2 \psi(a, b, \mu)\| \leq \frac{\sqrt{5}}{\sqrt{a^2 + b^2 + 2\mu^2}} \leq \frac{\sqrt{2}}{\mu}.$$

Proof The proof of (i) is trivial.

(ii): It is easy to check that

$$\nabla \psi(a, b, \mu) = \begin{pmatrix} 1 - \frac{a}{\sqrt{a^2 + b^2 + 2\mu^2}} \\ 1 - \frac{b}{\sqrt{a^2 + b^2 + 2\mu^2}} \\ -\frac{2\mu}{\sqrt{a^2 + b^2 + 2\mu^2}} \end{pmatrix}, \quad (2.9)$$

and

$$\nabla^2 \psi(a, b, \mu) = \frac{1}{(a^2 + b^2 + 2\mu^2)^{\frac{3}{2}}} \begin{pmatrix} -(b^2 + 2\mu^2) & ab & 2a\mu \\ ab & -(a^2 + 2\mu^2) & 2b\mu \\ 2a\mu & 2b\mu & -2(a^2 + b^2) \end{pmatrix} \quad (2.10)$$

Since $\nabla^2 \psi(a, b, \mu)$ is negative semi-definite for $(a, b, \mu) \in \mathbb{R}^2 \times \mathbb{R}_{++}$, the function $\psi(a, b, \mu)$ is concave on $\mathbb{R}^2 \times \mathbb{R}_{++}$.

(iii):

$$\begin{aligned} \|\nabla^2 \psi(a, b, \mu)\| &= \sqrt{\frac{(a^2 + b^2 + 2\mu^2)^2 + 4(a^2 + b^2 + \mu^2)^2}{(a^2 + b^2 + 2\mu^2)^3}} \\ &\leq \frac{\sqrt{5}}{\sqrt{a^2 + b^2 + 2\mu^2}} \leq \frac{\sqrt{2}}{\mu}. \end{aligned} \quad (2.11)$$

□

In the following lemma, we list some of the properties of the Chen-Harker-Kanzow-Smale smoothing function ϕ_μ for later use. The proof of the Lemma is similar to Lemma 2.1.1 and is omitted.

Lemma 2.1.3 *The Chen-Harker-Kanzow-Smale smoothing function $\phi_\mu : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ has the following properties:*

1. For every $\mu > 0$, the function ϕ_μ is continuously differentiable on \mathbb{R}^2 with

$$\frac{\partial \phi_\mu(a, b)}{\partial a} = 1 - \frac{a - b}{\sqrt{(a - b)^2 + 4\mu^2}} \quad \text{and} \quad \frac{\partial \phi_\mu(a, b)}{\partial b} = 1 + \frac{a - b}{\sqrt{(a - b)^2 + 4\mu^2}}.$$

2. For any $a, b \in \mathbb{R}$ and $\mu > 0$, one has

$$\|\nabla^2 \phi_\mu(a, b)\|_\infty \leq \frac{1}{\mu}. \quad (2.12)$$

3. For $\mu_1 \geq 0$, $\mu_2 \geq 0$ and $a, b \in \mathbb{R}$, we have

$$|\phi_{\mu_1}(a, b) - \phi_{\mu_2}(a, b)| \leq 2|\mu_1 - \mu_2|. \quad (2.13)$$

4. For any $a, b \in \mathbb{R}$ and $\mu > 0$, one has

$$0 < \frac{\partial \phi_\mu(a, b)}{\partial a} < 2, \quad 0 < \frac{\partial \phi_\mu(a, b)}{\partial b} < 2. \quad (2.14)$$

and

$$\left(\frac{\partial \phi_\mu(a, b)}{\partial a}\right)^2 + \left(\frac{\partial \phi_\mu(a, b)}{\partial b}\right)^2 \geq 2 > 0. \quad (2.15)$$

Finally, We list some properties of the the function $\phi(a, b, \mu)$ for later use.

Lemma 2.1.4 *The CHKS smoothing function $\phi : \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{++} \rightarrow \mathbb{R}$ defined in (2.4) has the following properties:*

(i) [41, 42] *The function $\phi(a, b, \mu)$ is continuously differentiable on $\mathbb{R}^2 \times \mathbb{R}_{++}$.*

(ii) *The function $\phi(a, b, \mu)$ is concave on $\mathbb{R}^2 \times \mathbb{R}_{++}$,*

(iii) [64, Lemma 2] *For any $(a, b, \mu) \in \mathbb{R}^2 \times \mathbb{R}_{++}$, we have*

$$\|\nabla^2 \phi(a, b, \mu)\| \leq \frac{4}{\sqrt{(a - b)^2 + 4\mu^2}} \leq \frac{2}{\mu}.$$

Proof (ii): It is easy to check that

$$\nabla\phi(a, b, \mu) = \begin{pmatrix} 1 - \frac{a-b}{\sqrt{(a-b)^2+4\mu^2}} \\ 1 + \frac{a-b}{\sqrt{(a-b)^2+4\mu^2}} \\ -\frac{4\mu}{\sqrt{(a-b)^2+4\mu^2}} \end{pmatrix}, \quad (2.16)$$

and

$$\nabla^2\phi(a, b, \mu) = \frac{4}{((a-b)^2+4\mu^2)^{\frac{3}{2}}} \begin{pmatrix} -\mu^2 & \mu^2 & (a-b)\mu \\ \mu^2 & -\mu^2 & -(a-b)\mu \\ (a-b)\mu & -(a-b)\mu & -(a-b)^2 \end{pmatrix}. \quad (2.17)$$

Since $\nabla^2\psi(a, b, \mu)$ is negative semi-definite for $(a, b, \mu) \in \mathbb{R}^2 \times \mathbb{R}_{++}$, the function $\phi(a, b, \mu)$ is concave on $\mathbb{R}^2 \times \mathbb{R}_{++}$. \square

Using ϕ_μ and ψ_μ as building blocks, we define $F_{\phi_\mu}(x, y) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{2n}$ and $F_{\psi_\mu}(x, y) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{2n}$ by

$$F_{\phi_\mu}(x, y) := \begin{bmatrix} Mx - y + q \\ \Phi_\mu(x, y) \end{bmatrix}, \quad (2.18)$$

$$F_{\psi_\mu}(x, y) := \begin{bmatrix} Mx - y + q \\ \Psi_\mu(x, y) \end{bmatrix}, \quad (2.19)$$

$$\Phi_\mu(x, y) := \begin{bmatrix} \phi_\mu(x_1, y_1) \\ \dots \\ \phi_\mu(x_n, y_n) \end{bmatrix}, \quad (2.20)$$

and

$$\Psi_\mu(x, y) := \begin{bmatrix} \psi_\mu(x_1, y_1) \\ \dots \\ \psi_\mu(x_n, y_n) \end{bmatrix}. \quad (2.21)$$

Clearly, a point (x, y) is on the central path if and only if $F_{\phi_\mu}(x, y) = 0$ (or $F_{\psi_\mu}(x, y) = 0$).

Depending on the conditions imposed on $LCP(q, M)$, there are two types of existence results for the central path. In order to state these results we need to review some terminology.

Definition 2.1.5 *A matrix $M \in \mathbb{R}^{n \times n}$ is said to be*

- (a) a P_0 matrix if each of its principal minors is non-negative,*
- (b) a P matrix if each of its principal minors is positive,*
- (c) an R_0 matrix if $LCP(0, M)$ has unique solution $(x, y) = (0, 0)$, and*
- (d) a non-degenerate matrix if each of its principal submatrices is nonsingular.*

The set of P_0 matrices clearly contains the set of all positive semi-definite matrices. The positive semi-definite matrices give rise to the *monotone* linear complementarity problems of which both linear and convex quadratic programming are special cases. Every positive definite matrix is a P matrix, and a P matrix is a non-degenerate matrix that is both a P_0 and an R_0 matrix. Under the assumption that the matrix M is an R_0 matrix, it is straightforward to show that the solution set

$$S = \{(x, y) : 0 \leq x, 0 \leq y, Mx - y + q = 0, \text{ and } x^T y = 0\} \quad (2.22)$$

to $LCP(q, M)$ is bounded. The boundedness of S is key to the analysis of the limiting behavior of the central path as the continuation parameter μ tends to zero. This limiting behavior and the existence of the central path is addressed in the following result due to Kanzow [42, Corollary 3.9] (also see Chen and Harker [9, Corollary 3.9]).

Theorem 2.1.6 *([42, Corollary 3.9]) If M is a P_0 and an R_0 matrix, then the equation $F_{\psi_\mu}(x, y) = 0$ (or $F_{\phi_\mu}(x, y) = 0$) has a unique solution $(x(\mu), y(\mu))$ for all $\mu > 0$. Moreover, the entire sequence $(x(\mu), y(\mu))$ converges to a solution of $LCP(q, M)$ as μ tends to 0.*

Some interesting linear complementarity problems don't satisfy the conditions in Theorem 2.1.6. For example, the LCP derived from linear programming and convex quadratic programming is monotone but is generally not R_0 . Therefore the existence results don't apply to these problems. A different existence result is established by Hotta and Yoshise [37] (see also Kojima, Megiddo, Noma and Yoshise [44]) and is based on the following condition.

Assumption (A): LCP(q, M) has a feasible interior point $(x, y) \in \mathbb{R}^{n \times n}$, i.e.,

$$x > 0, y > 0 \text{ and } Mx - y + q = 0.$$

Theorem 2.1.7 ([37, Theorem 3.3]) *If M is a positive semi-definite and condition (A) holds, then the equation $F_{\psi_\mu}(x, y) = 0$ (or $F_{\phi_\mu}(x, y) = 0$) has a unique solution $(x(\mu), y(\mu))$ for all $\mu > 0$. Moreover, the entire sequence $(x(\mu), y(\mu))$ converges to a solution of LCP(q, M) as μ tends to 0.*

2.1.2 Chen-Magasarian smoothing functions

Chen and Magasarian [13] introduced a class of smoothing functions that approximates the plus function z_+ by twice integrating a parameterized probability density function. The smoothing function is defined by

$$p_\mu(z) = \int_{-\infty}^z \int_{-\infty}^t \frac{1}{\mu} d\left(\frac{x}{\mu}\right) dx dt, \quad (2.23)$$

where $0 < \mu < \infty$ is a smoothing parameter and $d(x)$ is a probability density function, that is it satisfies

$$d(x) \geq 0, \quad \int_{-\infty}^{\infty} d(x) dx = 1.$$

When μ goes to 0, the limit of $\frac{1}{\mu} d\left(\frac{x}{\mu}\right)$ is the Dirac delta function, $\delta(x)$, which satisfies the following properties

$$\delta(x) \geq 0, \quad \int_{-\infty}^{\infty} \delta(x) dx = 1,$$

and the limit of the function $p_\mu(z)$ is the plus function $z_+ := \max\{z, 0\}$. It is well known that the complementarity condition

$$x \geq 0, y \geq 0, x^T y = 0 \quad (2.24)$$

can be rewritten as:

$$\min\{x, y\} = 0 \quad \text{or} \quad x - (x - y)_+ = 0, \quad (2.25)$$

where the plus function is taken component-wise. By replacing the plus function by the smooth function p_μ , we have the following smoothed complementarity condition:

$$x - P_\mu(x - y) = 0, \quad (2.26)$$

where $\mu > 0$ is the smoothing parameter and

$$P_\mu(x - y) := \begin{bmatrix} p_\mu(x_1 - y_1) \\ \dots \\ p_\mu(x_n - y_n) \end{bmatrix}. \quad (2.27)$$

The smoothed LCP then becomes

$$F_{p_\mu}(x, y) := \begin{bmatrix} Mx - y + q \\ x - P_\mu(x - y) \end{bmatrix} = 0. \quad (2.28)$$

In the following, we give several examples of Chen-Magasarian smoothing functions.

Example 2.1.8 *Neural networks smooth function (Chen and Magasarian [13]):*

$$d(x) = \frac{e^{-x}}{(1 + e^{-x})^2}, \quad p_\mu(z) = z + \mu \ln(1 + e^{-\frac{z}{\mu}}). \quad (2.29)$$

Example 2.1.9 *Chen-Harker-Kanzow-Smale smooth function (Chen and Harker [9], Kanzow [42] and Smale [72]):*

$$d(x) = \frac{2}{(x^2 + 4)^{\frac{3}{2}}}, \quad p_\mu(z) = \frac{z + \sqrt{z^2 + 4\mu^2}}{2}. \quad (2.30)$$

The properties of the Chen-Magasarian smoothing functions and the existence of the smooth path defined by (2.28) can be found in Chen and Magasarian [13], Chen and Harker [11], Chen and Xiu [12].

2.1.3 Gabriel and Moré smoothing functions

Gabriel and Moré [31] proposed a class of smoothing functions that approximate the median function. It is shown in [31] that the class of Gabriel and Moré smoothing functions is general enough to include the class of Chen-Magasarian smoothing functions.

It is well-known that the boxed constrained variational inequality problem $\text{VI}(F, [l, u])$ (1.13) can be reformulated as the following nonsmooth equation

$$x - \text{mid}(l, u, x - F(x)) = 0, \quad (2.31)$$

where $\text{mid}(\cdot)$ denotes the component-wise median operator. The smoothed variational inequality problem with parameter $\mu > 0$ is obtained by approximating the median function with the Gabriel-Moré smoothing function $G_\mu(x)$:

$$x - G_\mu(x) = 0, \quad (2.32)$$

where $G_\mu(x)$ is defined by

$$(G_\mu(x))_i = \int_{-\infty}^{\infty} \text{mid}(l_i, u_i, x_i - F_i(x) - \mu t) \rho(t) dt = 0, \quad i = 1, 2, \dots, n, \quad (2.33)$$

and ρ is a probability density function satisfying certain conditions. Typical choices for ρ include the neural network smooth function (2.29) and the Chen-Harker-Kanzow-Smale smooth function (2.30).

The analyses of algorithms based on the smoothed Fischer-Burmeister function ψ_μ and the Chen-Harker-Kanzow-Smale smoothing function ϕ_μ are very similar differing only by a constant here and there. In this thesis, for simplicity, we choose to focus on the CHKS smoothing function ϕ_μ . However, whenever appropriate, we

indicate how the analysis differs when the function ψ_μ is used instead of ϕ_μ . The Chen-Harker-Kanzow-Smale smoothing function is a member of the class of Chen-Magasarian smoothing functions. Therefore it is not surprising that the analysis of the algorithms based on the Chen-Harker-Kanzow-Smale smoothing function can be generalized to the algorithms based on the Chen-Magasarian smoothing functions and more generally to algorithms based on Gabriel and Moré smoothing functions.

2.2 A New Notion of Neighborhoods

For $\beta > 0$ and $\mu > 0$, we define a *slice* of the neighborhood by

$$\mathcal{N}_s(\beta, \mu) := \{ (x, y) : Mx - y + q = 0, \|\Phi_\mu(x, y)\|_\infty \leq \beta\mu \}, \quad (2.34)$$

and take as our *neighborhood of the central path* the union of all slices over $\mu > 0$, i.e.,

$$\mathcal{N}_s(\beta) := \cup_{\mu > 0} \mathcal{N}_s(\beta, \mu), \quad (2.35)$$

here the subscript s is used to distinguish this neighborhood with the neighborhood to be discussed later in this section.

Just as the assumption that the matrix M in $\text{LCP}(q, M)$ is an R_0 matrix can be used to establish the boundedness of the solution set S , this assumption can also be used to assure the boundedness of certain *slices* of the neighborhood $\mathcal{N}_s(\beta)$.

Proposition 2.2.1 *Let $\beta > 0$, $\mu_0 > 0$. If M is an R_0 matrix, then the set*

$$\cup_{0 < \mu \leq \mu_0} \mathcal{N}_s(\beta, \mu)$$

is bounded.

Proof The pattern of proof is identical to that which is used to show the boundedness of S . Suppose to the contrary that there exists an unbounded sequence $\{(x^k, y^k)\} \in \cup_{0 < \mu \leq \mu_0} \mathcal{N}_s(\beta, \mu)$. Then there is also a sequence of scalars $\{\mu_k\}$ such that $\{(x^k, y^k)\} \in$

$\mathcal{N}_s(\beta, \mu_k)$ and $0 < \mu_k \leq \mu_0$. Since the sequence $\{((x^k, y^k) / \|(x^k, y^k)\|_\infty, \mu_k)\}$ is bounded, we may assume without loss of generality that this sequence converges to a point $((x^*, y^*), \mu_*) \in \mathbb{R}^{2n+1}$. By dividing the equation $Mx^k - y^k + q = 0$ through by $\|(x^k, y^k)\|_\infty$ and taking the limit as $k \rightarrow \infty$, we find that

$$Mx^* - y^* = 0. \quad (2.36)$$

In addition, for each $i = 1, \dots, n$, we have

$$\frac{|\phi_{\mu_k}(x_i^k, y_i^k)|}{\|(x^k, y^k)\|_\infty} \leq \frac{\beta \mu_k}{\|(x^k, y^k)\|_\infty} \leq \frac{\beta \mu_0}{\|(x^k, y^k)\|_\infty}.$$

Again, taking the limit in k yields

$$\phi_0(x_i^*, y_i^*) = 0, \text{ for each } i = 1, \dots, n. \quad (2.37)$$

But (2.37) and (2.36) taken together imply that $(x^*, y^*) \neq 0$ is a solution to $\text{LCP}(0, M)$.

This contradiction yields the result. \square

Proposition 2.2.1 shows that the set $\cup_{0 < \mu \leq \mu_0} \mathcal{N}_s(\beta, \mu)$ is bounded when M is an R_0 matrix. However, the set may not be bounded for arbitrary $\beta > 0$ and $\mu_0 > 0$ under condition (A), as shown by the following example suggested by Sun [74].

Example 2.2.2 *Let*

$$n = 1, M = 0, q = 1, \mu_0 = 1, \beta = 2.$$

Then condition (A) holds, but the set $\mathcal{N}_s(\beta, \mu_0)$ is unbounded.

The above example motivates us to modify the neighborhood $\mathcal{N}_s(\beta)$. It turns out that the modified neighborhood has the desired boundedness property under condition (A). The modified neighborhood is defined as

$$\mathcal{N}(\beta) := \left\{ (x, y) \mid \begin{array}{l} Mx - y + q = 0, \Phi(x, y, \mu) \leq 0, \\ \|\Phi(x, y, \mu)\|_\infty \leq \beta \mu \text{ for some } \mu > 0 \end{array} \right\}, \quad (2.38)$$

where $\beta > 0$, $\Phi(x, y, \mu) : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{++} \rightarrow \mathbb{R}^n$ is defined by

$$\Phi(x, y, \mu) := \begin{bmatrix} \phi(x_1, y_1, \mu) \\ \dots \\ \phi(x_n, y_n, \mu) \end{bmatrix},$$

and ϕ by (2.4). This neighborhood can be viewed as the union of the *slices*

$$\mathcal{N}(\beta, \mu) := \{(x, y) : Mx - y + q = 0, \Phi(x, y, \mu) \leq 0, \|\Phi(x, y, \mu)\|_\infty \leq \beta\mu\} \quad (2.39)$$

for $\mu > 0$.

We will show that if the algorithm is initiated in this neighborhood, then the inequality $\Phi(x, y, \mu) \leq 0$ is automatically satisfied at subsequent iterates. Hence this inequality does not complicate the structure of the algorithm. In the monotone case, this inequality is key to establishing the boundedness of the iterates.

Lemma 2.2.3 *Assume that condition (A) holds. Then for any $\beta > 0$ and $\mu_0 > 0$, the set*

$$\cup_{0 < \mu \leq \mu_0} \mathcal{N}(\beta, \mu)$$

is bounded. Indeed, for any $(x, y) \in \cup_{0 < \mu \leq \mu_0} \mathcal{N}(\beta, \mu)$, we have for $i = 1, 2, \dots, n$

$$\begin{aligned} -(\beta\mu_0)/2 \leq x_i &\leq \frac{\bar{x}^T \bar{y} + \frac{\beta\mu_0}{2} (\|\bar{x}\|_1 + \|\bar{y}\|_1) + n \max\{\mu_0^2, \frac{\beta^2 \mu_0^2}{4}\}}{\bar{y}_i} \\ -(\beta\mu_0)/2 \leq y_i &\leq \frac{\bar{x}^T \bar{y} + \frac{\beta\mu_0}{2} (\|\bar{x}\|_1 + \|\bar{y}\|_1) + n \max\{\mu_0^2, \frac{\beta^2 \mu_0^2}{4}\}}{\bar{x}_i}, \end{aligned}$$

where (\bar{x}, \bar{y}) is any feasible interior point, that is, a point satisfying

$$M\bar{x} - \bar{y} + q = 0, \quad \bar{x} > 0, \quad \bar{y} > 0.$$

Proof Let $\beta > 0$, $0 < \mu \leq \mu_0$, and $(x, y) \in \mathcal{N}(\beta, \mu)$ be given, and let (\bar{x}, \bar{y}) be a feasible interior point for LCP(q, M). First observe that if $-\delta \leq \phi(a, b, \mu)$, then $-\delta/2 < \min\{a, b\}$. To see this, note that the condition $-\delta \leq \phi(a, b, \mu)$ implies that

$$0 < \sqrt{[(a + \delta/2) - (b + \delta/2)]^2 + 4\mu^2} \leq (a + \delta/2) + (b + \delta/2). \quad (2.40)$$

Squaring both sides and cleaning up yields $0 < \mu^2 \leq (a + \delta/2)(b + \delta/2)$. Thus, since at least one of $(a + \delta/2)$ and $(b + \delta/2)$ must be positive by (2.40), both must be positive yielding $-\delta/2 < \min\{a, b\}$. This observation implies that

$$\begin{aligned} x_i &> -(\beta\mu)/2 \geq -(\beta\mu_0)/2, \text{ and} \\ y_i &> -(\beta\mu)/2 \geq -(\beta\mu_0)/2. \end{aligned} \quad (2.41)$$

for $i = 1, 2, \dots, n$.

Next, note that if $0 \leq a$ and $0 \leq b$, then the inequality $\phi(a, b, \mu) \leq 0$ implies that $0 \leq a + b \leq \sqrt{(a - b)^2 + 4\mu^2}$. Again, by squaring and cleaning up, we see that this gives $ab \leq \mu^2$. This observation implies that

$$x_i y_i \leq \mu_0^2 \text{ for each } i \in \{1, \dots, n\} \text{ with } 0 < x_i, 0 < y_i. \quad (2.42)$$

We conclude the proof by noting that monotonicity yields $0 \leq (\bar{x} - x)^T(\bar{y} - y)$, or equivalently $\bar{x}^T y + \bar{y}^T x \leq \bar{x}^T \bar{y} + x^T y$. This inequality plus those in (2.41) and (2.42) yield the bound

$$\begin{aligned} \sum_{y_i > 0} \bar{x}_i y_i + \sum_{x_i > 0} \bar{y}_i x_i &\leq \bar{x}^T \bar{y} + x^T y - \left[\sum_{y_i < 0} \bar{x}_i y_i + \sum_{x_i < 0} \bar{y}_i x_i \right] \\ &\leq \bar{x}^T \bar{y} + x^T y + \frac{\beta\mu_0}{2} (\|\bar{x}\|_1 + \|\bar{y}\|_1) \\ &\leq \bar{x}^T \bar{y} + \sum_{\substack{x_i > 0 \\ y_i > 0}} x_i y_i + \sum_{\substack{x_i < 0 \\ y_i < 0}} x_i y_i + \frac{\beta\mu_0}{2} (\|\bar{x}\|_1 + \|\bar{y}\|_1) \\ &\leq \bar{x}^T \bar{y} + n \max\{\mu_0^2, \frac{\beta^2 \mu_0^2}{4}\} + \frac{\beta\mu_0}{2} (\|\bar{x}\|_1 + \|\bar{y}\|_1). \end{aligned}$$

It follows that if $y_i > 0$, then

$$y_i \leq \frac{\bar{x}^T \bar{y} + \frac{\beta\mu_0}{2} (\|\bar{x}\|_1 + \|\bar{y}\|_1) + n \max\{\mu_0^2, \frac{\beta^2 \mu_0^2}{4}\}}{\bar{x}_i},$$

and, if $x_i > 0$, then

$$x_i \leq \frac{\bar{x}^T \bar{y} + \frac{\beta\mu_0}{2} (\|\bar{x}\|_1 + \|\bar{y}\|_1) + n \max\{\mu_0^2, \frac{\beta^2 \mu_0^2}{4}\}}{\bar{y}_i}.$$

□

We now prove an interesting structural property of the neighborhood. We show that the solution set

$$\mathcal{S} = \{(x, y) : Mx - y + q = 0, x \geq 0, y \geq 0, x^T y = 0\}, \quad (2.43)$$

is contained in the interior of the slice $\mathcal{N}(\beta, \mu)$ relative to the affine set

$$\Lambda = \{(x, y) : Mx - y + q = 0\}, \quad (2.44)$$

for all $\mu > 0$ and $\beta > 2$. This property demonstrates that this neighborhood is fundamentally different from those used in the interior-point literature, where the solution set necessarily lies on the boundary of the neighborhood relative to the affine constraints.

Theorem 2.2.4 *For any $\mu > 0$ and $\beta > 2$, the solution set \mathcal{S} is contained in the interior of the slice $\mathcal{N}(\beta, \mu)$ relative to the affine set Λ .*

Proof First note that if $x_i y_i \leq \mu^2$, then

$$\begin{aligned} \phi(x_i, y_i, \mu) &= x_i + y_i - \sqrt{(x_i - y_i)^2 + 4\mu^2} \\ &= x_i + y_i - \sqrt{(x_i + y_i)^2 + 4(\mu^2 - x_i y_i)} \\ &\leq 0. \end{aligned}$$

Also, if $x_i \geq 0, y_i \geq 0$ and $x_i y_i = 0$, then either $x_i = 0$ or $y_i = 0$. If $x_i = 0$ and $y_i \geq 0$, then

$$\begin{aligned} \phi(x_i, y_i, \mu) &= x_i + y_i - \sqrt{(x_i - y_i)^2 + 4\mu^2} \\ &= y_i - \sqrt{y_i^2 + 4\mu^2} \\ &\geq y_i - (y_i + 2\mu) \\ &= -2\mu > -\beta\mu. \end{aligned}$$

Similarly, if $y_i = 0$ and $x_i \geq 0$, then again $\phi(x_i, y_i, \mu) > -\beta\mu$. Now if $(x^*, y^*) \in S$, then by the continuity of function ϕ , there is an $\delta > 0$ such that for all (x, y) in

$$\mathcal{O}(\delta) = \{(x, y) : \|x - x^*\|_\infty \leq \delta, \|y - y^*\|_\infty \leq \delta\}$$

we have $\phi(x_i, y_i, \mu) \geq -\beta\mu$ and $x_i y_i \leq \mu^2$ for all $i = 1, \dots, n$. Therefore for any $(x, y) \in \mathcal{O}(\delta)$, we have

$$\Phi(x, y, \mu) \leq 0, \|\Phi(x, y, \mu)\|_\infty \leq \beta\mu,$$

and so (x^*, y^*) is in the interior of the set $\mathcal{N}(\beta, \mu)$ relative to the affine set Λ . \square

Corollary 2.2.5 *For any $\mu > 0$ and $\beta > 2$, the solution set S is contained in the interior of the slice $\mathcal{N}_s(\beta, \mu)$ relative to the affine set Λ .*

Proof The result follows from Theorem 2.2.4 and the fact that

$$\mathcal{N}(\beta, \mu) \subseteq \mathcal{N}_s(\beta, \mu).$$

\square

2.3 A Prototype Algorithm for Non-Interior Path Following Methods

In this section, we give a prototype algorithm for non-interior path following methods and compare it with a prototype algorithm for interior-point path following methods.

A Prototype Algorithm for Non-Interior Path-Following Methods

Step 0 (Initialization)

Let $\mu_0 > 0$, $\beta > 0$, and $(x^0, y^0) \in \mathbb{R}^{2n}$ be given so that $(x^0, y^0) \in \mathcal{N}(\beta, \mu_0)$.

Step 1 (Computation of the Newton Direction)

For $k = 1, 2, 3, \dots$ compute a Newton direction $(\Delta x^k, \Delta y^k)$ at (x^k, y^k) from the equations

$$Mx - y + q = 0, \quad \Phi_{\sigma_k \mu_k}(x, y) = 0, \quad (2.45)$$

where $\sigma_k \in [0, 1]$.

Step 2 (Backtracking Line Search)

Set

$$(x^{k+1}, y^{k+1}) = (x^k, y^k) + \alpha_k (\Delta x^k, \Delta y^k),$$

where α_k is chosen such that

$$(x^{k+1}, y^{k+1}) \in \mathcal{N}(\beta, \mu_{k+1}), \quad (2.46)$$

with $\mu_{k+1} \in (0, \mu_k)$.

Remarks 1. For LCP(q, M) with R_0 matrix, we can replace the set $\mathcal{N}(\beta, \mu)$ by the set $\mathcal{N}_s(\beta, \mu)$ in the algorithm.

2. The parameter β can be chosen to be any positive number.
3. In the initialization step, setting

$$\mu_0 > \sqrt{\max_{\substack{i \in \{1, \dots, n\} \\ 0 < x_i^0, 0 < y_i^0}} x_i^0 y_i^0}$$

guarantees that the inequality $\Phi(x^0, y^0, \mu_0) < 0$ is satisfied. For example, one can choose $(x^0, y^0) = (0, q)$ in which case μ_0 can be taken to be any positive number.

4. The choices of search directions and the backtracking line search steps will be discussed in details in the next chapter.

We now give a prototype algorithm for interior-point path following methods.

A Prototype Algorithm for Interior Point Path-Following Methods

Step 0 (Initialization)

Let $\mu_0 > 0$, $\beta > 0$, and $(x^0, y^0) \in \mathbb{R}_{++}^{2n}$ be given so that $(x^0, y^0) \in \mathcal{N}_{int}(\beta, \mu_0)$.

Step 1 (Computation of the Newton Direction)

For $k = 1, 2, 3, \dots$, compute a Newton direction $(\Delta x^k, \Delta y^k)$ at (x^k, y^k) from the equations

$$Mx - y + q = 0, \quad Xy - \sigma_k \mu_k I = 0, \quad (2.47)$$

where $\sigma_k \in [0, 1]$.

Step 2 (Backtracking Line Search)

Set

$$(x^{k+1}, y^{k+1}) = (x^k, y^k) + \alpha_k (\Delta x^k, \Delta y^k),$$

where α_k is chosen such that

$$(x^{k+1}, y^{k+1}) > 0 \quad \text{and} \quad (x^{k+1}, y^{k+1}) \in \mathcal{N}_{int}(\beta, \mu_{k+1}), \quad (2.48)$$

with $\mu_{k+1} \in (0, \mu_k)$.

Remarks 1. The set $\mathcal{N}_{int}(\beta, \mu)$ is a slice of the neighborhood $\mathcal{N}_{int}(\beta)$ defined in (1.34), that is

$$\mathcal{N}_{int}(\beta, \mu) = \{(x, y) : Mx - y + q = 0, \frac{\|Xy - \mu e\|}{\mu} \leq \beta, 0 < x, 0 < y, \}.$$

2. The parameter β in the algorithm has to be in $(0, 1)$.

3. In practice, it may be difficult to find an initial point satisfying $(x^0, y^0) > 0$ and $(x^0, y^0) \in \mathcal{N}_{int}(\beta, \mu_0)$. To handle this difficulty, various infeasible interior-point methods are proposed where $Mx^k - y^k + q$ may or may not be 0.

4. Unlike non-interior path following methods, interior point methods require all intermediate steps to stay in the interior of the nonnegative orthant.

2.4 An Example

In this section, we use an example to illustrate the prototype algorithms for non-interior path following methods and interior-point path following methods.

Consider the LCP(q, M) given in [13, Example 5.1], where

$$M = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, \quad q = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$$

The unique solution of this problem is $(x_1, x_2) = (1, 0)$ and $(y_1, y_1) = (0, 1)$.

For $\mu > 0$, let $\mathcal{N}^x(\beta, \mu)$ be the projection of the slice $\mathcal{N}(\beta, \mu)$ onto the x -coordinate, i.e., the set

$$\{(x_1, x_2) : \exists y \text{ such that } Mx - y + q = 0, \Phi(x, y, \mu) \leq 0, \|\Phi(x, y, \mu)\|_\infty \leq \beta\mu\}.$$

In figure 1, several slices of $\mathcal{N}^x(\beta, \mu)$ are drawn with $\beta = 4$ and $\mu = 15, 10, 5$ respectively.

For $\mu > 0$, let $\mathcal{N}_{int}^x(\beta, \mu)$ be the projection of the slice $\mathcal{N}_{int}(\beta, \mu)$ onto the x -coordinate, i.e., the set

$$\{(x_1, x_2) : \exists y \text{ such that } x > 0, y > 0, Mx - y + q = 0, \|\Phi(x, y, \mu)\|_\infty \leq \beta\mu\}.$$

Here the infinity norm is used in $\mathcal{N}_{int}(\beta, \mu)$ instead of Euclidean norm.

In figure 2, several slices of $\mathcal{N}_{int}^x(\beta, \mu)$ are drawn with $\beta = 0.5$ and $\mu = 15, 10, 5$ respectively.

Remarks 1. The iterates in both interior-point path following methods and non-interior path following methods move from one slice to another and eventually lead to a solution.

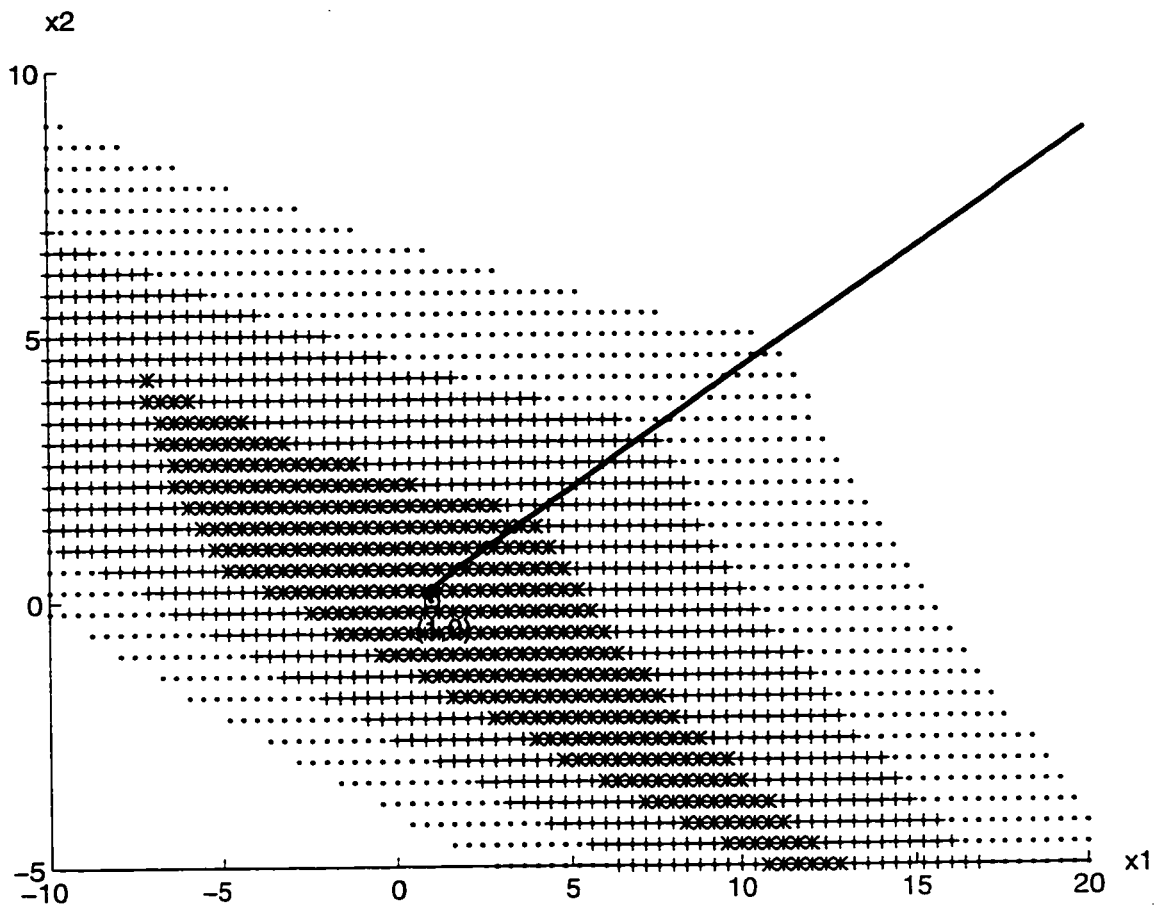


Figure 2.1: The slices of $\mathcal{N}^x(\beta, \mu)$ with $\beta = 4$ and $\mu = 15, 10, 5$.

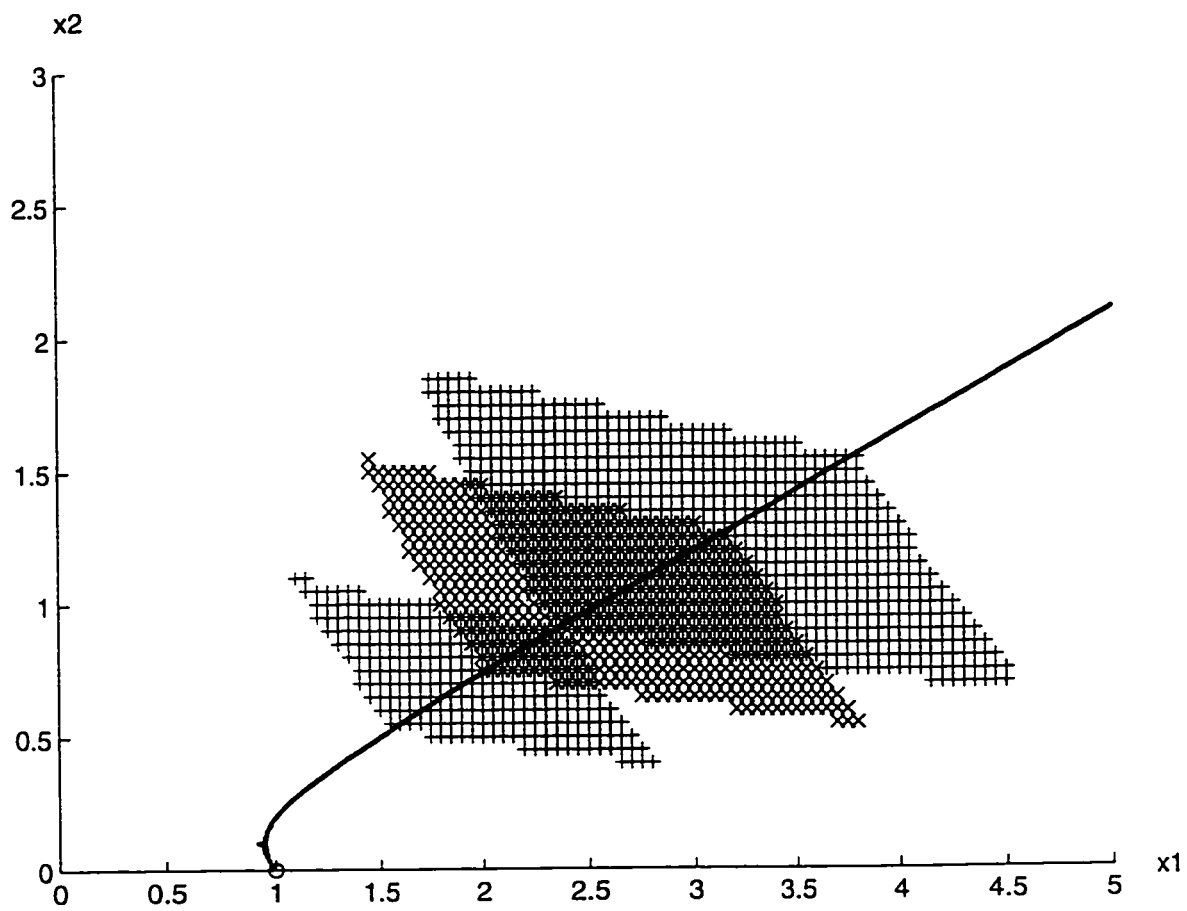


Figure 2.2: The slices of $\mathcal{N}_{int}^x(\beta, \mu)$ with $\beta = 0.5$ and $\mu = 15, 10, 5$.

2. From Figure 2.1 and Figure 2.2 we see that the neighborhood used in interior-point methods reside in the interior of the nonnegative orthant, but the neighborhood used in non-interior path following methods does not necessarily reside in the interior of the nonnegative orthant.
3. As shown in Figure 2.2, the solutions to $\text{LCP}(q, M)$ are on the boundary of the interior-point neighborhood $\mathcal{N}_{int}(\beta)$ (1.34) relative to the affine set Λ (2.44). On the other hand, from Figure 2.1, we see that the solutions to $\text{LCP}(q, M)$ are contained in the interior of every slice of the neighborhood $\mathcal{N}(\beta)$ (2.38) relative to the affine set Λ . This confirms the result in Theorem 2.2.4.

Chapter 3

THE GLOBAL AND LOCAL CONVERGENCE OF A PREDICTOR-CORRECTOR NON-INTERIOR PATH-FOLLOWING METHOD FOR LCP

In this chapter, we introduce a predictor-corrector strategy for non-interior path following methods and propose a predictor-corrector algorithm for linear complementarity problem. Under standard hypotheses, we show that the proposed algorithm is both globally linearly convergent and locally quadratically convergent. In addition, we show that the algorithm is globally convergent under a relatively mild condition. The algorithm and the convergence results presented in this chapter are based on the Chen-Harker-Kanzow-Smale smoothing function. However, it should be pointed out that the algorithms and the convergence analysis remain valid if the Chen-Harker-Kanzow-Smale smoothing function is replaced by the smoothed Fischer-Burmeister function.

In this chapter, the parameter μ in the definition of ϕ plays a dual role. It is considered both as a smoothing parameter and as a variable. Therefore, we choose to use the notation

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

instead of ϕ_μ .

The plan of this chapter is as follows. In section 3.1, we introduce the predictor and corrector search directions for non-interior path following methods. In section 3.2, we propose a predictor-corrector non-interior path following algorithm and prove

that the algorithm is well defined. The global linear convergence of the algorithm is presented in section 3.3 and the local quadratic convergence is given in section 3.4. In section 3.5, we show that the algorithm is globally convergent under a relatively mild condition.

3.1 Predictor and Corrector Search Directions

We obtain the predictor and corrector directions by applying Newton's method to equations of the form

$$F_\phi(x, y, \mu) = v \quad (3.1)$$

for various choices of the right hand side v where the function $F : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{++} \rightarrow \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{++}$ is given by

$$F_\phi(x, y, \mu) := \begin{bmatrix} Mx - y + q \\ \Phi(x, y, \mu) \\ \mu \end{bmatrix}, \quad (3.2)$$

with

$$\Phi(x, y, \mu) = \begin{bmatrix} \phi(x_1, y_1, \mu) \\ \dots \\ \phi(x_n, y_n, \mu) \end{bmatrix}. \quad (3.3)$$

Note that

$$F_\phi(x, y, \mu) = 0 \quad (3.4)$$

if and only if (x, y) solves $\text{LCP}(q, M)$, and

$$F_\phi(x, y, \mu) = \begin{bmatrix} 0 \\ 0 \\ \bar{\mu} \end{bmatrix} \quad \text{with } \bar{\mu} \neq 0 \quad (3.5)$$

if and only if (x, y) is on the central path \mathcal{C} corresponding to the smoothing parameter $\bar{\mu}$.

If we choose $v = 0$, then the Newton direction at $(\bar{x}, \bar{y}, \bar{\mu})$ to the equation (3.1) gives us the *predictor direction*, that is, the direction $(\Delta x, \Delta y, \Delta \mu)$ satisfying the equations

$$F_\phi(\bar{x}, \bar{y}, \bar{\mu}) + \nabla F_\phi(\bar{x}, \bar{y}, \bar{\mu})^T \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta \mu \end{bmatrix} = 0. \quad (3.6)$$

It is helpful to take a closer look at the Newton equations (3.6). In (3.6), we have $\Delta \mu = -\bar{\mu}$ and so system (3.6) reduces to

$$\begin{aligned} M\Delta x - \Delta y &= 0, \\ \nabla_x \Phi(\bar{x}, \bar{y}, \bar{\mu})^T \Delta x + \nabla_y \Phi(\bar{x}, \bar{y}, \bar{\mu})^T \Delta y &= -\Phi(\bar{x}, \bar{y}, \bar{\mu}) + \bar{\mu} \nabla_\mu \Phi(\bar{x}, \bar{y}, \bar{\mu}). \end{aligned} \quad (3.7)$$

If we choose

$$v = \begin{bmatrix} 0 \\ 0 \\ \bar{\mu} \end{bmatrix},$$

then the Newton direction at $(\bar{x}, \bar{y}, \bar{\mu})$ for the equation (3.1) gives us the *centering direction* (or corrector direction), that is, the direction $(\Delta x, \Delta y, \Delta \mu)$ satisfying the equations

$$F_\phi(\bar{x}, \bar{y}, \bar{\mu}) + \nabla F_\phi(\bar{x}, \bar{y}, \bar{\mu})^T \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta \mu \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \bar{\mu} \end{bmatrix}. \quad (3.8)$$

It is easy to see that $\Delta \mu = 0$, so the system (3.8) reduces to

$$\begin{aligned} M\Delta x - \Delta y &= 0, \\ \nabla_x \Phi(\bar{x}, \bar{y}, \bar{\mu})^T \Delta x + \nabla_y \Phi(\bar{x}, \bar{y}, \bar{\mu})^T \Delta y &= -\Phi(\bar{x}, \bar{y}, \bar{\mu}). \end{aligned} \quad (3.9)$$

Finally, if we choose

$$v = \begin{bmatrix} 0 \\ 0 \\ (1 - \sigma)\bar{\mu} \end{bmatrix},$$

for some $\sigma \in (0, 1)$, then the Newton direction at $(\bar{x}, \bar{y}, \bar{\mu})$ to the equation (3.1) gives us a *combined direction*, that is, a direction $(\Delta x, \Delta y, \Delta \mu)$ satisfying the equations

$$F_\phi(\bar{x}, \bar{y}, \bar{\mu}) + \nabla F_\phi(\bar{x}, \bar{y}, \bar{\mu})^T \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta \mu \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ (1 - \sigma)\bar{\mu} \end{bmatrix}. \quad (3.10)$$

In this case, $\Delta \mu = -\sigma \bar{\mu}$. Therefore the system (3.10) reduces to

$$\begin{aligned} M \Delta x - \Delta y &= 0 \\ \nabla_x \Phi(\bar{x}, \bar{y}, \bar{\mu})^T \Delta x + \nabla_y \Phi(\bar{x}, \bar{y}, \bar{\mu})^T \Delta y &= -\Phi(\bar{x}, \bar{y}, \bar{\mu}) + \sigma \bar{\mu} \nabla_\mu \Phi(\bar{x}, \bar{y}, \bar{\mu}). \end{aligned} \quad (3.11)$$

When $\sigma = 1$, this yields the predictor direction and when $\sigma = 0$, it yields the centering direction. Therefore, when $0 < \sigma < 1$, the combined direction can be considered as a combination of the predictor direction and the centering direction. Most existing non-interior path following methods [5, 8, 7, 9, 12, 42, 41, 43, 83, 84] are based on the centering direction. The predictor direction and the combined direction are new and are used in the predictor-corrector non-interior path following algorithm to be studied in this chapter.

The general outline of our predictor-corrector algorithm can now be described. First, a predictor step is computed. This is the Newton step based on the equation (3.4) at the current iterate. The predictor step in the x and y variables is either accepted or rejected depending on whether it is in a pre-specified neighborhood of the central path for the current value of the smoothing parameter. If it is in this neighborhood, then the predictor step is accepted and a backtracking routine is applied to the smoothing parameter μ to reduce its value as much as is possible subject to remaining in the neighborhood of the central path. Next, a corrector step is computed. This is the Newton step based on the equation (3.5) at the iterate obtained from the predictor step with $\bar{\mu}$ taken to be a fixed fraction of the value for μ obtained from the predictor step. Again, a step length is determined to return the update to the neighborhood of the central path. The technique for choosing the corrector

step guarantees the global linear convergence of the method, while the technique for choosing the predictor step provides for the local quadratic convergence of the iterates. The line search routines for both the predictor and corrector steps are based on finitely terminating backtracking procedures and as such are easily implemented.

3.2 A Predictor-Corrector Non-Interior Path-Following Algorithm

In this section, we state our predictor-corrector algorithm for LCP and show that it is well-defined.

The Predictor-Corrector Non-Interior Path-Following Algorithm

Step 0: (Initialization)

Choose $x^0 \in \mathbb{R}^n$, set $y^0 = Mx^0 + q$, and let $\mu_0 > 0$ be such that $\Phi(x^0, y^0, \mu_0) < 0$.
 Choose $\beta > 2$ so that $\|\Phi(x^0, y^0, \mu_0)\|_\infty \leq \beta\mu_0$. We now have $(x^0, y^0) \in \mathcal{N}(\beta, \mu_0)$.
 Choose $\bar{\sigma}$, α_1 , and α_2 from $(0, 1)$.

Step 1: (The Predictor Step)

Let $(\Delta x^k, \Delta y^k, \Delta \mu_k)$ solve the equation

$$F_\phi(x^k, y^k, \mu_k) + \nabla F_\phi(x^k, y^k, \mu_k)^T \begin{bmatrix} \Delta x^k \\ \Delta y^k \\ \Delta \mu_k \end{bmatrix} = 0. \quad (3.12)$$

If $\|\Phi(x^k + \Delta x^k, y^k + \Delta y^k, 0)\|_\infty = 0$, **STOP**, $(x^k + \Delta x^k, y^k + \Delta y^k)$ solves LCP(q, M);
 else if

$$\|\Phi(x^k + \Delta x^k, y^k + \Delta y^k, \mu_k)\|_\infty > \beta\mu_k,$$

set

$$\hat{x}^k := x^k, \quad \hat{y}^k := y^k, \quad \hat{\mu}_k := \mu_k, \quad \text{and} \quad \eta_k = 1; \quad (3.13)$$

else let $\eta_k = \alpha_1^s$ where s is the positive integer such that

$$\|\Phi(x^k + \Delta x^k, y^k + \Delta y^k, \alpha_1^t \mu_k)\|_\infty \leq \alpha_1^t \beta \mu_k, \quad t = 0, 1, 2, \dots, s, \quad \text{and} \quad (3.14)$$

$$\|\Phi(x^k + \Delta x^k, y^k + \Delta y^k, \alpha_1^{s+1} \mu_k)\|_\infty > \alpha_1^{s+1} \beta \mu_k. \quad (3.15)$$

Set

$$\hat{x}^k := x^k + \Delta x^k, \quad \hat{y}^k := y^k + \Delta y^k, \quad \hat{\mu}_k := \eta_k \mu_k. \quad (3.16)$$

Step 2: (The Corrector Step)

Let $(\Delta \hat{x}^k, \Delta \hat{y}^k, \Delta \hat{\mu}_k)$ solve the equation

$$F_\phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) + \nabla F_\phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)^T \begin{bmatrix} \Delta \hat{x}^k \\ \Delta \hat{y}^k \\ \Delta \hat{\mu}_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ (1 - \bar{\sigma}) \hat{\mu}_k \end{bmatrix} \quad (3.17)$$

and let $\hat{\lambda}_k$ be the maximum of the value $1, \alpha_2, \alpha_2^2, \dots$, such that

$$\|\Phi(\hat{x}^k + \hat{\lambda}_k \Delta \hat{x}^k, \hat{y}^k + \hat{\lambda}_k \Delta \hat{y}^k, (1 - \bar{\sigma} \hat{\lambda}_k) \hat{\mu}_k)\|_\infty \leq (1 - \bar{\sigma} \hat{\lambda}_k) \beta \hat{\mu}_k. \quad (3.18)$$

Set

$$x^{k+1} = \hat{x}^k + \hat{\lambda}_k \Delta \hat{x}^k, \quad y^{k+1} = \hat{y}^k + \hat{\lambda}_k \Delta \hat{y}^k, \quad \mu_{k+1} = (1 - \bar{\sigma} \hat{\lambda}_k) \hat{\mu}_k, \quad (3.19)$$

and return to Step 1.

Remarks 1. Note that if the null step (3.13) is taken in Step 1, then the Newton equations (3.12) and (3.17) have the same coefficient matrix. Therefore only one matrix factorization is needed to implement both Steps 1 and 2 in this case.

2. In the initialization step, setting

$$\mu_0 > \sqrt{\max_{\substack{i \in \{1, \dots, n\} \\ 0 < x_i^0, 0 < y_i^0}} x_i^0 y_i^0}$$

guarantees that the inequality $\Phi(x^0, y^0, \mu_0) < 0$ is satisfied. For example, one can choose $(x^0, y^0) = (0, q)$ in which case μ_0 can be taken to be any positive number.

3. The condition that $\beta > 2$ is only employed in proof of local quadratic convergence. It is not required to verify the global linear convergence of the method.
4. Observe that the function F has nonsingular Jacobian at a given point if and only if $\nabla_{(x,y)}F$ is nonsingular at that point. In [42, Theorem 3.5] it is shown that if $\mu > 0$ and M is a P_0 -matrix, then $\nabla_{(x,y)}F(\bar{x}, \bar{y}, \mu)$ is nonsingular for all $(\bar{x}, \bar{y}) \in \mathbb{R}^{2n}$. Therefore if M is a P_0 -matrix, the Jacobians $\nabla F(x^k, y^k, \mu_k)$ and $\nabla F(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)$ are always nonsingular and so the Newton equations (3.12) and (3.17) yield unique solutions whenever (x^k, y^k, μ_k) and $(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)$ are well-defined. In addition, since $y^0 = Mx^0 + q$, we have $y^k = Mx^k + q$ and $\hat{y}^k = M\hat{x}^k + q$ for all well-defined iterates.

Theorem 3.2.1 *Consider the algorithm described above with M being a P_0 -matrix. If $(x^k, y^k) \in \mathcal{N}(\beta, \mu_k)$ with $\mu_k > 0$, then either $(x^k + \Delta x^k, y^k + \Delta y^k)$ solves $LCP(q, M)$ or both $(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)$ and $(x^{k+1}, y^{k+1}, \mu_{k+1})$ are well-defined with the backtracking routines in Steps 1 and 2 finitely terminating. In the latter case, we have $(\hat{x}^k, \hat{y}^k) \in \mathcal{N}(\beta, \hat{\mu}_k)$ and $(x^{k+1}, y^{k+1}) \in \mathcal{N}(\beta, \mu_{k+1})$ with $0 < \mu_{k+1} < \hat{\mu}_k$. Since $(x^0, y^0) \in \mathcal{N}(\beta, \mu_0)$ with $\mu_0 > 0$, this shows that the algorithm is well-defined.*

Proof Let $(x^k, y^k) \in \mathcal{N}(\beta, \mu_k)$ with $\mu_k > 0$. By Remark 4 above, $(\Delta x^k, \Delta y^k, \Delta \mu_k)$ exists and is unique. Since $y^k + \Delta y^k = M(x^k + \Delta x^k) + q$, we have

$$\|\Phi(x^k + \Delta x^k, y^k + \Delta y^k, 0)\|_\infty = 0$$

if and only if $(x^k + \Delta x^k, y^k + \Delta y^k)$ solves $LCP(q, M)$. Therefore, if $(x^k + \Delta x^k, y^k + \Delta y^k)$ does not solve $LCP(q, M)$, then by continuity, there exist $\epsilon > 0$ and $\bar{\mu} > 0$ such that $\|\Phi(x^k + \Delta x^k, y^k + \Delta y^k, \mu)\|_\infty > \epsilon$ for all $\mu \in [0, \bar{\mu}]$. In this case, the

backtracking routine described in (3.14) and (3.15) of Step 2 is finitely terminating. Hence $(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)$ is well-defined, with $0 < \hat{\mu}_k \leq \mu_k$, and $(\Delta\hat{x}^k, \Delta\hat{y}^k, \Delta\hat{\mu}_k)$ is uniquely determined by (3.17). To see that the backtracking routine in Step 3 is finitely terminating, define $\theta(x, y, \mu) = \|\Phi(x, y, \mu)\|_\infty$. This is a convex composite function [2]. By (3.17)

$$\begin{aligned}
& \theta'((\hat{x}^k, \hat{y}^k, \hat{\mu}_k); (\Delta\hat{x}^k, \Delta\hat{y}^k, \Delta\hat{\mu}_k)) \\
&= \inf_{\lambda > 0} \lambda^{-1} \left(\left\| \Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) + \lambda \nabla \Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)^T \begin{bmatrix} \Delta\hat{x}^k \\ \Delta\hat{y}^k \\ \Delta\hat{\mu}_k \end{bmatrix} \right\|_\infty - \|\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)\|_\infty \right) \\
&\leq \left\| \Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) + \nabla \Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)^T \begin{bmatrix} \Delta\hat{x}^k \\ \Delta\hat{y}^k \\ \Delta\hat{\mu}_k \end{bmatrix} \right\|_\infty - \|\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)\|_\infty \\
&= -\|\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)\|_\infty \\
&< 0.
\end{aligned}$$

Therefore, (3.18) can be viewed as an instance of a standard backtracking line search routine and as such is finitely terminating with $0 < \mu_{k+1} < \hat{\mu}_k \leq \mu_k$ (indeed, one can replace the value of $\bar{\sigma}$ on the right hand side of (3.18) by any number in the open interval $(0, 1)$).

Since $(x^k, y^k) \in \mathcal{N}(\beta, \mu_k)$, the argument given above implies that either $(x^k + \Delta x^k, y^k + \Delta y^k)$ solves $\text{LCP}(q, M)$ or $\hat{y}^k = M\hat{x}^k + q$ with $\|\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)\|_\infty \leq \beta\hat{\mu}_k$ and $y^{k+1} = Mx^{k+1} + q$ with $\|\Phi(x^{k+1}, y^{k+1}, \mu_{k+1})\|_\infty \leq \beta\mu_{k+1}$. Thus, if $(x^k + \Delta x^k, y^k + \Delta y^k)$ does not solve $\text{LCP}(q, M)$, we need only show that $\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) \leq 0$ and $\Phi(x^{k+1}, y^{k+1}, \mu_{k+1}) \leq 0$ in order to have $(\hat{x}^k, \hat{y}^k) \in \mathcal{N}(\beta, \hat{\mu}_k)$ and $(x^{k+1}, y^{k+1}) \in \mathcal{N}(\beta, \mu_{k+1})$. First note that the component-wise concavity of Φ (see Lemma 2.1.4) implies that for any $(x, y, \mu) \in \mathbb{R}^{2n+1}$ with $\mu > 0$, and $(\Delta x, \Delta y, \Delta \mu) \in \mathbb{R}^{2n+1}$ one

has

$$\Phi(x + \Delta x, y + \Delta y, \mu + \Delta\mu) \leq \Phi(x, y, \mu) + \nabla\Phi(x, y, \mu)^T \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta\mu \end{bmatrix}.$$

Hence, in the case of the predictor step, either (3.13) holds or

$$\begin{aligned} & \Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) \\ = & \Phi(x^k + \Delta x^k, y^k + \Delta y^k, \eta_k \mu_k) \\ \leq & \Phi(x^k, y^k, \mu_k) + \nabla\Phi(x^k, y^k, \mu_k)^T \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ (\eta_k - 1)\mu_k \end{pmatrix} \\ = & \Phi(x^k, y^k, \mu_k) + \nabla\Phi(x^k, y^k, \mu_k)^T \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ -\mu_k \end{pmatrix} + \eta_k \mu_k \nabla_\mu \Phi(x^k, y^k, \mu_k) \\ = & \eta_k \mu_k \nabla_\mu \Phi(x^k, y^k, \mu_k) \\ \leq & 0. \end{aligned}$$

In either case, $\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) \leq 0$. For the corrector step, we have

$$\begin{aligned} & \Phi(x^{k+1}, y^{k+1}, \mu_{k+1}) \\ = & \Phi(\hat{x}^k + \hat{\lambda}_k \Delta \hat{x}^k, \hat{y}^k + \hat{\lambda}_k \Delta \hat{y}^k, \hat{\mu}_k - \bar{\sigma} \hat{\lambda}_k \Delta \hat{\mu}_k) \\ \leq & \Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) + \hat{\lambda}_k \nabla\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)^T \begin{bmatrix} \Delta \hat{x}^k \\ \Delta \hat{y}^k \\ -\bar{\sigma} \hat{\mu}_k \end{bmatrix} \\ = & (1 - \hat{\lambda}_k) \Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) \leq 0, \end{aligned}$$

since we have already shown that $\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) \leq 0$. This completes the proof. \square

3.3 The Global Linear Convergence

To establish the global linear convergence, we require the following key assumption:

Assumption (B): Given $\beta > 0$ and $\mu_0 > 0$, there exists a $C > 0$ such that

$$\|\nabla_{(x,y)} F_\phi(\bar{x}, \bar{y}, \mu)^{-1}\|_\infty \leq C, \quad (3.20)$$

for all $0 < \mu \leq \mu_0$ and $(\bar{x}, \bar{y}) \in \mathcal{N}(\beta, \mu)$.

We now give a condition under which Assumption (B) holds.

Definition 3.3.1 Let S be the set of solutions to $LCP(q, M)$ and set

$$J := \{j \mid \text{there exists } (x, y) \in S \text{ such that } y_j = 0\}.$$

We say that the problem $LCP(q, M)$ satisfies the FLP condition if the principal submatrix M_{JJ} is non-degenerate.

Remarks 1. The FLP condition extends a similar notion due to Fukushima, Luo, and Pang [30, Assumption (A2)].

2. If M is a P matrix, then clearly $LCP(q, M)$ satisfies the FLP condition.

The key step in establishing the uniform boundedness of $\nabla_{(x,y)} F_\phi(x, y, \mu)^{-1}$ is provided by the following technical lemma due to Fukushima, Luo, and Pang.

Lemma 3.3.2 [30, Proposition 3.2] Let $\beta > 0$ and $\mu_0 > 0$ be given and assume that M is a P_0 matrix. Let (x^k, y^k, μ_k) be a sequence in \mathbb{R}^{2n+1} such that $(x^k, y^k) \in \mathcal{N}_s(\beta, \mu_k)$ and $0 < \mu_k \leq \mu_0$ for all k . If the limit

$$\lim_{k \rightarrow \infty} (\nabla_x \Phi(x^k, y^k, \mu_k), \nabla_y \Phi(x^k, y^k, \mu_k)) = (D_x, D_y)$$

exists and the principal submatrix M_{JJ} is non-degenerate, where

$$J = \{i : (D_x)_{ii} = 0\},$$

then the limiting matrix

$$\begin{bmatrix} M & -I \\ D_x & D_y \end{bmatrix}$$

is nonsingular.

Proposition 3.3.3 *Let $\mu_0 > 0, \beta > 0$ and assume that M is a P_0 and an R_0 matrix for which $LCP(q, M)$ satisfies the FLP condition. Then for all $0 < \mu$ and $x, y \in \mathbb{R}^n$ satisfying*

$$0 < \mu \leq \mu_0, \text{ and } (x, y) \in \mathcal{N}_s(\beta, \mu)$$

there exists a constant $C > 0$ such that

$$\|\nabla_{(x,y)} F_\phi(x, y, \mu)^{-1}\|_\infty \leq C. \quad (3.21)$$

Proof Assume to the contrary that there is a sequence $\{(x^k, y^k, \mu_k)\}$ such that $0 < \mu_k \leq \mu_0$, $(x^k, y^k) \in \mathcal{N}_s(\beta, \mu_k)$, and $\|\nabla_{(x,y)} F_\phi(x^k, y^k, \mu_k)^{-1}\|_\infty \geq k$. By Proposition 2.2.1, the sequence $\{(x^k, y^k, \mu_k)\}$ is bounded, hence we can assume that the sequence converges to some point (x^*, y^*, μ_*) . If $\mu_* > 0$, then $\nabla_{(x,y)} F_\phi(x^*, y^*, \mu_*)$ is non-singular which implies the boundedness of the sequence $\|\nabla_{(x,y)} F_\phi(x^k, y^k, \mu_k)^{-1}\|_\infty$. Hence it must be the case that $\mu_* = 0$. Therefore $(x^*, y^*) \in S$. In addition, by Lemma (2.1.3), the sequence $(\nabla_x \Phi(x^k, y^k, \mu_k), \nabla_y \Phi(x^k, y^k, \mu_k))$ is also bounded, so with no loss in generality

$$\lim_{k \rightarrow \infty} (\nabla_x \Phi(x^k, y^k, \mu_k), \nabla_y \Phi(x^k, y^k, \mu_k)) = (D_x, D_y)$$

for some positive diagonal matrices D_x and D_y . Let $I = \{i | (D_x)_{ii} = 0\}$. It is easy to check that $I \subseteq J$ (see Lemma (2.1.3)). Therefore, M_{II} is non-degenerate since M_{JJ} is. Hence Lemma 3.3.2 implies that $\lim_k \nabla_{(x,y)} F_\phi(x^k, y^k, \mu_k)$ exists and is non-singular. But then again the sequence $\|(\nabla_{(x,y)} F_\phi(x^k, y^k, \mu_k))^{-1}\|_\infty$ must be bounded. This contradiction yields the result. \square

We are now ready to establish the global linear convergence of the algorithm. This result depends only on the corrector step (Step 2 of the algorithm) and is independent of whether or not the predictor step (Step 1 of the algorithm) is implemented on any given iteration.

Theorem 3.3.4 (Global Linear Convergence) *Suppose that M is a P_0 -matrix and that Assumption (B) holds. Let $\{(x^k, y^k, \mu_k)\}$ be the sequence generated by the algo-*

rithm. If the algorithm does not terminate finitely at a solution to $LCP(q, M)$, then for $k = 0, 1, \dots$,

$$(x^k, y^k) \in \mathcal{N}(\beta, \mu_k), \quad (3.22)$$

$$(1 - \bar{\sigma}\hat{\lambda}_{k-1})\eta_{k-1} \dots (1 - \bar{\sigma}\hat{\gamma}_0)\eta_0\mu_0 = \mu_k, \quad (3.23)$$

with

$$\hat{\lambda}_k \geq \bar{\lambda} := \min\left\{1, \frac{\alpha_2(1 - \bar{\sigma})\beta}{2C^2(\beta + 2\bar{\sigma})^2 + \bar{\sigma}^2 + \bar{\sigma}(1 - \bar{\sigma})\beta}\right\}, \quad (3.24)$$

where C is the constant defined in (3.20). Therefore μ_k converges to 0 at a global linear rate. In addition, the sequence $\{(x^k, y^k)\}$ converges to a solution of $LCP(q, M)$.

Proof (i) The inclusion (3.22) has already been established in Theorem 3.2.1 and the relation (3.23) follows by construction.

(ii) For sake of simplicity, set $(x, y, \mu) = (\hat{x}_k, \hat{y}_k, \hat{\mu}_k)$ and $(\Delta x, \Delta y) = (\Delta \hat{x}^k, \Delta \hat{y}^k)$.

Then for $i \in \{1, \dots, n\}$ and $\lambda \in [0, 1]$, Lemma 2.1.4 and (3.17) imply that

$$\begin{aligned} & |\phi(x_i + \lambda\Delta x_i, y_i + \lambda\Delta y_i, (1 - \bar{\sigma}\lambda)\mu)| \\ = & \left| \phi(x_i, y_i, \mu) + \lambda \nabla \phi(x_i, y_i, \mu)^T \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ -\bar{\sigma}\mu \end{pmatrix} + \right. \\ & \left. \frac{\lambda^2}{2} \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ -\bar{\sigma}\mu \end{pmatrix}^T \nabla^2 \phi(x_i + \theta_i \lambda \Delta x_i, y_i + \theta_i \lambda \Delta y_i, (1 - \theta_i \bar{\sigma} \lambda) \mu) \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ -\bar{\sigma}\mu \end{pmatrix} \right| \\ \leq & (1 - \lambda) |\phi(x_i^k, y_i^k, \mu)| + \frac{\lambda^2}{2} \left\| \nabla^2 \phi(x_i + \theta_i \lambda \Delta x_i, y_i + \theta_i \lambda \Delta y_i, (1 - \theta_i \bar{\sigma} \lambda) \mu) \right\| \left\| \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ -\bar{\sigma}\mu \end{pmatrix} \right\|^2 \\ \leq & (1 - \lambda) |\phi(x_i, y_i, \mu)| + \frac{\lambda^2}{(1 - \bar{\sigma}\lambda)\mu} \|(\Delta x_i, \Delta y_i, -\bar{\sigma}\mu)\|^2 \end{aligned}$$

for some $\theta_i \in [0, 1]$. Set $t_i := \|(\Delta x_i, \Delta y_i, -\bar{\sigma}\mu)\|^2$ for $i = 1, \dots, n$, then

$$\begin{aligned}
& \|\Phi(x + \lambda\Delta x, y + \lambda\Delta y, (1 - \bar{\sigma}\lambda)\mu)\|_\infty \\
& \leq (1 - \lambda) \|\Phi(x, y, \mu)\|_\infty + \frac{\lambda^2}{(1 - \bar{\sigma}\lambda)\mu} \left\| \begin{pmatrix} t_1 \\ \dots \\ t_n \end{pmatrix} \right\|_\infty \\
& \leq (1 - \lambda) \|\Phi(x, y, \mu)\|_\infty + \frac{\lambda^2}{(1 - \bar{\sigma}\lambda)\mu} \left(2 \left\| \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\|_\infty^2 + \bar{\sigma}^2 \mu^2 \right) \\
& \leq (1 - \lambda)\beta\mu + \frac{\lambda^2}{1 - \bar{\sigma}\lambda} (2C^2(\beta + 2\bar{\sigma})^2 + \bar{\sigma}^2) \mu, \tag{3.25}
\end{aligned}$$

where the last inequality follows from (3.7) which yields the bound

$$\begin{aligned}
\left\| \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\|_\infty & \leq \|\nabla_{(x,y)} F^{-1}(x, y, \mu)\|_\infty (\|\Phi(x, y, \mu)\|_\infty + \bar{\sigma}\mu \|\nabla_\mu \Phi(x, y, \mu)\|_\infty) \\
& \leq C(\beta + 2\bar{\sigma})\mu. \tag{3.26}
\end{aligned}$$

It is easily verified that

$$(1 - \lambda)\beta\mu + \frac{\lambda^2}{1 - \bar{\sigma}\lambda} [2C^2(\beta + 2\bar{\sigma})^2 + \bar{\sigma}^2] \mu \leq (1 - \bar{\sigma}\lambda)\beta\mu,$$

whenever

$$\lambda \leq \frac{(1 - \bar{\sigma})\beta}{2C^2(\beta + 2\bar{\sigma})^2 + \bar{\sigma}^2 + \bar{\sigma}(1 - \bar{\sigma})\beta}.$$

Therefore

$$\hat{\lambda}_k \geq \min \left\{ 1, \frac{\alpha_2(1 - \bar{\sigma})\beta}{2C^2(\beta + 2\bar{\sigma})^2 + \bar{\sigma}^2 + \bar{\sigma}(1 - \bar{\sigma})\beta} \right\}.$$

Just as in (3.26), the relations (3.7) and (3.9) yield the bounds

$$\left\| \begin{pmatrix} \Delta x^k \\ \Delta y^k \end{pmatrix} \right\|_\infty \leq C(\beta + 2)\mu_k \quad \text{and} \quad \left\| \begin{pmatrix} \Delta \hat{x}^k \\ \Delta \hat{y}^k \end{pmatrix} \right\|_\infty \leq C(\beta + 2)\mu_k,$$

since $0 < \bar{\sigma} < 1$ and $0 < \eta_k \leq 1$ for all k . Therefore, (3.23) and (3.24) imply that

$$\begin{aligned}
\|(x^{k+1}, y^{k+1}) - (x^k, y^k)\|_\infty & \leq \left\| \begin{pmatrix} \Delta x^k \\ \Delta y^k \end{pmatrix} \right\|_\infty + \hat{\lambda}_k \left\| \begin{pmatrix} \Delta \hat{x}^k \\ \Delta \hat{y}^k \end{pmatrix} \right\|_\infty \\
& \leq 2C(\beta + 2)\mu_k \leq 2C(\beta + 2)(1 - \bar{\sigma}\bar{\lambda})^k \mu_0.
\end{aligned}$$

Hence, $\{(x^k, y^k)\}$ is a Cauchy sequence and so must converge to a solution of $\text{LCP}(q, M)$.

□

3.4 The Local Quadratic Convergence

We now establish the local quadratic convergence of the μ_k 's under the assumption that the iterates converge to a solution of $\text{LCP}(q, M)$ at which strict complementary slackness is satisfied.

Theorem 3.4.1 (*Local Quadratic Convergence*) *Suppose Assumption (B) holds and that the sequence $\{(x^k, y^k, \mu_k)\}$ generated by the algorithm converges to $\{(x^*, y^*, 0)\}$ where $\{(x^*, y^*)\}$ is a solution to $\text{LCP}(q, M)$. If it is further assumed that the strict complementary slackness condition $0 < x^* + y^*$ is satisfied, then*

$$\mu_{k+1} = O(\mu_k^2), \quad (3.27)$$

that is, μ_k converges quadratically to zero.

Proof First observe that due to the strict complementarity of $\{(x^*, y^*)\}$, Part (iii) of Lemma 2.1.4 indicates that there exist constants $\epsilon > 0$ and $L > 0$ such that

$$\|\nabla^2 \phi(x, y, \mu)\| \leq L, \quad \text{whenever } \|(x, y, \mu) - (x^*, y^*, 0)\|_\infty \leq \epsilon. \quad (3.28)$$

Hence, for all k sufficient large and $\eta \in (0, 1]$, we have for each $i \in \{1, \dots, n\}$ that

$$\begin{aligned} & |\phi(x_i^k + \Delta x_i^k, y_i^k + \Delta y_i^k, \eta \mu_k)| \\ = & \left| \phi(x_i^k, y_i^k, \mu_k) + \nabla^T \phi(x_i^k, y_i^k, \mu_k) \begin{pmatrix} \Delta x_i^k \\ \Delta y_i^k \\ (\eta - 1)\mu_k \end{pmatrix} \right| + \\ & \frac{1}{2} \left| \begin{pmatrix} \Delta x_i^k \\ \Delta y_i^k \\ (\eta - 1)\mu_k \end{pmatrix}^T \nabla^2 \phi(x_i^k + \theta_i \Delta x_i^k, y_i^k + \theta_i \Delta y_i^k, (1 + \theta_i(\eta - 1))\mu_k) \begin{pmatrix} \Delta x_i^k \\ \Delta y_i^k \\ (\eta - 1)\mu_k \end{pmatrix} \right| \end{aligned}$$

$$\begin{aligned}
&= \left| \phi(x_i^k, y_i^k, \mu_k) + \nabla^T \phi(x_i^k, y_i^k, \mu_k) \begin{pmatrix} \Delta x_i^k \\ \Delta y_i^k \\ -\mu_k \end{pmatrix} + \eta \mu_k \nabla_\mu \phi(x^k, y^k, \mu_k) \right| + \\
&\quad \frac{1}{2} \left| \begin{pmatrix} \Delta x_i^k \\ \Delta y_i^k \\ (\eta - 1)\mu_k \end{pmatrix}^T \nabla^2 \phi(x_i^k + \theta_i \Delta x_i^k, y_i^k + \theta_i \Delta y_i^k, (1 + \theta_i(\eta - 1))\mu_k) \begin{pmatrix} \Delta x_i^k \\ \Delta y_i^k \\ (\eta - 1)\mu_k \end{pmatrix} \right| \\
&\leq \eta \mu_k |\nabla_\mu \phi(x^k, y^k, \mu_k)| + \frac{L}{2} \|(\Delta x_i^k, \Delta y_i^k, (\eta - 1)\mu_k)\|^2 \\
&= 2\eta \mu_k + \frac{L}{2} \|(\Delta x_i^k, \Delta y_i^k, (\eta - 1)\mu_k)\|^2
\end{aligned}$$

Now using an argument similar to that used to obtain (3.25), we have

$$\begin{aligned}
\|\Phi(x^k + \Delta x^k, y^k + \Delta y^k, \eta \mu_k)\|_\infty &\leq 2\eta \mu_k + \frac{L}{2}(2C^2(\beta + 2)^2 + (1 - \eta)^2)\mu_k^2 \\
&\leq 2\eta \mu_k + \frac{L}{2}(2C^2(\beta + 2)^2 + 1)\mu_k^2. \quad (3.29)
\end{aligned}$$

Hence, since $\beta > 2$, the inequality (3.14) in Step 1 of the algorithm holds with $t = 0$ for all k sufficiently large. It is easy to verify that

$$2\eta \mu_k + \frac{L}{2}(2C^2(\beta + 2)^2 + 1)\mu_k^2 \leq \eta \beta \mu_k, \quad (3.30)$$

whenever

$$\eta \geq \frac{L}{2(\beta - 2)}[2C^2(\beta + 2)^2 + 1]\mu_k.$$

Hence, by (3.15), we have

$$\alpha_1 \eta_k \leq \frac{L}{2(\beta - 2)}[2C^2(\beta + 2)^2 + 1]\mu_k,$$

and so

$$\eta_k \leq \frac{L}{2\alpha_1(\beta - 2)}[C^2(\beta + 2)^2 + 1]\mu_k, \quad (3.31)$$

for all k sufficient large. Therefore, by (3.16),

$$\mu_{k+1} = O(\mu_k^2).$$

□

3.5 The Global Convergence

We now show that under relatively mild condition, the algorithm is globally convergent.

Theorem 3.5.1 *Suppose that M is a P_0 matrix and that Assumption (A) holds. Let $\{x^k, y^k, \mu^k\}$ be a sequence generated by the Algorithm. Then*

- (i) *The sequence $\{\mu_k\}$ is monotonically decreasing and converges to 0 as $k \rightarrow \infty$.*
- (ii) *The sequence $\{(x^k, y^k)\}$ is bounded and every accumulation point of $\{(x^k, y^k)\}$ is a solution of the LCP.*

Proof Since M is a P_0 matrix, the algorithm is well defined. By the construction of the algorithm, we can see that $\mu_{k+1} < \mu_k$ for $(k = 0, 1, \dots)$. Hence the sequence $\{\mu_k\}$ is monotonically decreasing. Since $\mu_k \geq 0$ ($k = 0, 1, \dots$), there is a $\bar{\mu} \geq 0$ such that $\mu_k \rightarrow \bar{\mu}$. Since $\{(x^k, y^k)\} \in \mathcal{N}(\beta, \mu_k)$, the sequence (x^k, y^k) is bounded, by taking a subsequence if necessary, we may assume that $\{(x^k, y^k)\}$ converges to some point (\bar{x}, \bar{y}) . If $\bar{\mu} = 0$, it follows from $\{(x^k, y^k)\} \in \mathcal{N}(\beta, \mu_k)$ that (\bar{x}, \bar{y}) is a solution of the LCP and we obtain the desired results. Suppose that $\bar{\mu} > 0$. Since $\nabla_{(x,y)} F(\bar{x}, \bar{y}, \bar{\mu})$ is nonsingular, there exist $\epsilon > 0$, $L > 0$ and $C > 0$ such that

$$\|\nabla_{(x,y)} F(x, y, \mu)^{-1}\|_{\infty} \leq C, \quad (3.32)$$

$$\|\nabla^2 \phi(x_i, y_i, \mu)\|_2 \leq L, \quad (3.33)$$

for all $(x, y, \mu) \in \mathcal{O}_{(\bar{x}, \bar{y}, \bar{\mu})} = \{(x, y, \mu) : \|(x, y, \mu) - (\bar{x}, \bar{y}, \bar{\mu})\| \leq \epsilon\}$. Similar to the proof of the global linear convergence result, we can show that there exists a $\bar{\lambda}$ such that $\hat{\lambda}_k \geq \bar{\lambda}$ for sufficient large k . Therefore, for sufficient large k , $\mu_{k+1} \leq c\mu_k$ for some constant $c \in (0, 1)$, we get a contradiction. \square

Chapter 4

THE CONDITION-BASED COMPLEXITY OF NON-INTERIOR PATH FOLLOWING METHODS FOR LCP

It is well known that interior-point methods have an excellent property, the polynomial complexity. It is natural to ask what is the complexity of non-interior path following methods. In the next two chapters, we provide some partial results to this question. In this chapter, we establish a condition-based complexity for a non-interior path following method for the following two types of problems: (1) the LCP has a P -matrix, and (2) the LCP has a symmetric positive definite matrix. In the next chapter, we will study the worst-case complexity. We show that an interior-point method based on a rescaled Newton directions has polynomial complexity. These results are the only complexity results available to date and represent a first step toward understanding the complexity of non-interior path following methods.

In the Turing Machine Model, the complexity of an algorithm is given as a function of the numbers of variables, the number of constraints, and the bit-size of the data of an instance. However, complexity in terms of the size of the problem does not really measure the difficulty of an instance of the problem. Two instances with the same size may result in different performance for the same algorithm. Therefore, it is valuable to study the condition-based complexity. For references on the condition-based complexity for interior-point methods, see Renegar [67], Filipowski [26], Freund and Vera [29], Todd and Ye [75]. The goal of this chapter is to study the condition-based complexity for non-interior path following methods. The algorithm in this

chapter uses only a centering direction. However, it should be pointed out that similar results can also be established for the predictor-corrector algorithm studied in the previous chapter.

Unlike the previous chapter, we consider μ as a smoothing parameter in the definition of the functions ϕ and ψ . To reflect this interpretation, we use the notation ϕ_μ and ψ_μ in this chapter.

The plan of this chapter is as follows. In section 4.1, we establish a global error bound for the Chen-Harker-Kanzow-Smale smoothing function and the smoothed Fischer-Burmeister function in terms of certain “natural” residue. Such error bound is useful in designing stopping criterion for non-interior path following methods. We propose in section 4.2 a non-interior path following method which uses only a centering direction and prove its global linear convergence in section 4.3. A condition-based complexity will be established for LCP with a P -matrix in section 4.4 and for LCP with symmetric positive matrix in section 4.5.

4.1 A Global Error Bound

The following two lemmas relating the growth behavior of $|\phi_\mu(a, b)|$ and $|\psi_\mu(a, b)|$ with $|\min\{a, b\}|$ suggests that we can use the norm of certain “natural” residual as the stopping criterion in non-interior path following methods.

Lemma 4.1.1 *For any $a, b \in \mathbb{R}$ and $\mu \geq 0$, we have*

$$|\min(a, b)| \leq \mu + \frac{1 + \sqrt{2}}{2} |\phi_\mu(a, b)|. \quad (4.1)$$

Proof Without loss of generality, we assume that $b \geq a$. Then $\min(a, b) = a$. We can further assume that $a \neq 0$ since in this case (4.1) trivially holds.

In case $a + b > 0$, we have $b > |a| > 0$ and

$$|\phi_\mu(a, b)| = |a + b - \sqrt{(a - b)^2 + 4\mu^2}|$$

$$\begin{aligned}
&= \frac{|(a+b)^2 - (a-b)^2 - 4\mu^2|}{a+b + \sqrt{(a-b)^2 + 4\mu^2}} \\
&= \frac{4|ab - \mu^2|}{a+b + \sqrt{(a-b)^2 + 4\mu^2}} \\
&= \frac{4|a - \frac{\mu^2}{b}|}{\frac{a}{b} + 1 + \sqrt{(\frac{a}{b} - 1)^2 + 4\frac{\mu^2}{b^2}}} \\
&\geq \frac{4|a - \frac{\mu^2}{b}|}{2 + \sqrt{4 + 4\frac{\mu^2}{b^2}}}.
\end{aligned}$$

So

$$4|a| - 4\frac{\mu^2}{b} \leq 4|a - \frac{\mu^2}{b}| \leq (2 + 2\sqrt{1 + \frac{\mu^2}{b^2}})|\phi_\mu(a, b)|,$$

and

$$|a| \leq \frac{\mu^2}{b} + \frac{1 + \sqrt{1 + \frac{\mu^2}{b^2}}}{2} |\phi_\mu(a, b)| \leq \frac{\mu^2}{|a|} + \frac{1 + \sqrt{1 + \frac{\mu^2}{a^2}}}{2} |\phi_\mu(a, b)|. \quad (4.2)$$

We claim that

$$|\min(a, b)| = |a| \leq \mu + \frac{1 + \sqrt{2}}{2} |\phi_\mu(a, b)|. \quad (4.3)$$

Indeed, if $\frac{\mu^2}{a^2} \leq 1$, then (4.3) follows from (4.2); and if $\frac{\mu^2}{a^2} \geq 1$, then $|\min(a, b)| = |a| \leq \mu$.

In case $a + b \leq 0$, we have $a \leq -|b|$ and

$$\begin{aligned}
|\phi_\mu(a, b)| &= |a + b - \sqrt{(a-b)^2 + 4\mu^2}| \\
&= \sqrt{(a-b)^2 + 4\mu^2} - (a+b) \\
&\geq \sqrt{(a-b)^2} - (a+b) \\
&= b - a - a - b = -2a = 2|a|.
\end{aligned}$$

So

$$|\min(a, b)| = |a| \leq \frac{|\phi_\mu(a, b)|}{2}. \quad (4.4)$$

The inequality (4.1) then follows from (4.3) and (4.4).

framebox

Corollary 4.1.2 For all $\mu \geq 0$ and $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, we have

$$\|\min\{x, y\}\|_\infty \leq \mu + \frac{1 + \sqrt{2}}{2} \|\Phi_\mu(x, y)\|_\infty. \quad (4.5)$$

Lemma 4.1.3 For any $a, b \in \mathbb{R}$ and $\mu \geq 0$, we have

$$|\min(a, b)| \leq \mu + 2|\psi_\mu(a, b)|. \quad (4.6)$$

Proof Without loss of generality, we assume that $b \geq a$. Then $\min(a, b) = a$. We can further assume that $a \neq 0$ since in this case (4.6) trivially holds.

In case $a + b > 0$, we have $b > |a| > 0$ and

$$\begin{aligned} |\psi_\mu(a, b)| &= \frac{|a + b - \sqrt{a^2 + b^2 + 2\mu^2}|}{|a + b + \sqrt{a^2 + b^2 + 2\mu^2}|} \\ &= \frac{2|ab - \mu^2|}{a + b + \sqrt{a^2 + b^2 + 2\mu^2}} \\ &= \frac{2|a - \frac{\mu^2}{b}|}{\frac{a}{b} + 1 + \sqrt{\frac{a^2}{b^2} + 1 + 2\frac{\mu^2}{b^2}}} \\ &\geq \frac{2|a - \frac{\mu^2}{b}|}{2 + \sqrt{2 + 2\frac{\mu^2}{b^2}}}. \end{aligned}$$

So

$$2|a| - 2\frac{\mu^2}{b} \leq 2|a - \frac{\mu^2}{b}| \leq (2 + \sqrt{2 + 2\frac{\mu^2}{b^2}})|\psi_\mu(a, b)|,$$

and

$$|a| \leq \frac{\mu^2}{b} + \frac{2 + \sqrt{2 + 2\frac{\mu^2}{b^2}}}{2} |\psi_\mu(a, b)| \leq \frac{\mu^2}{|a|} + \frac{2 + \sqrt{2 + 2\frac{\mu^2}{a^2}}}{2} |\psi_\mu(a, b)|. \quad (4.7)$$

We claim that

$$|\min(a, b)| = |a| \leq \mu + 2|\psi_\mu(a, b)|. \quad (4.8)$$

Indeed, if $\frac{\mu^2}{a^2} \leq 1$, then (4.8) follows from (4.7); and if $\frac{\mu^2}{a^2} \geq 1$, then $|\min(a, b)| = |a| \leq \mu$.

In case $a + b \leq 0$, we have $a \leq -|b|$ and

$$\begin{aligned} |\psi_\mu(a, b)| &= |a + b - \sqrt{a^2 + b^2 + 2\mu^2}| \\ &= \sqrt{a^2 + b^2 + 2\mu^2} - (a + b) \\ &\geq \sqrt{a^2 + b^2 + 2\mu^2} \\ &\geq |a|. \end{aligned}$$

So

$$|\min(a, b)| = |a| \leq |\psi_\mu(a, b)|. \quad (4.9)$$

The inequality (4.6) then follows from (4.8) and (4.9). \square

Corollary 4.1.4 *For all $\mu \geq 0$ and $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, we have*

$$\|\min\{x, y\}\|_\infty \leq \mu + 2 \|\Psi_\mu(x, y)\|_\infty. \quad (4.10)$$

Remarks 1. In the non-interior path-following algorithm in the next section, a sequence of iterates $\{x^k, y^k, \mu_k\}$ is generated such that $y^k = Mx^k + q$, $\mu_k \rightarrow 0$ and $\|\Phi(x^k, y^k, \mu_k)\|_\infty \rightarrow 0$ (or $\|\Psi(x^k, y^k, \mu_k)\|_\infty \rightarrow 0$). By Corollary 4.1.4, the algorithm actually reduces the “natural” residue, $\|\min\{x^k, Mx^k + q\}\|_\infty$, to zero.

2. It is well-known that $\|\min\{x, Mx + q\}\|_\infty$ is a global error bound for $\text{LCP}(q, M)$ when M is an R_0 -matrix [50, 51], in particular when M is a P -matrix [52]. In this case, there exists a constant c which is independent of μ such that

$$\min_{(x^*, y^*) \in \mathcal{S}} \|(x, y) - (x^*, y^*)\|_\infty \leq c [\mu + 2 \|\Phi_\mu(x, y)\|_\infty],$$

$$\min_{(x^*, y^*) \in \mathcal{S}} \|(x, y) - (x^*, y^*)\|_\infty \leq c [\mu + 2 \|\Psi_\mu(x, y)\|_\infty].$$

3. When $\mu = 0$, similar result was established by Tseng [77] for the function ψ .

4.2 A Non-Interior Path Following Method based on Centering Direction

The Algorithm Based on Centering Direction

Step 0 (Initialization)

Let $\mu_0 > 0$, $\beta > 0$, and $(x^0, y^0) \in \mathbb{R}^{2n}$ be given so that $(x^0, y^0) \in \mathcal{N}_s(\beta, \mu_0)$, and choose $\sigma_i \in (0, 1]$ and $\alpha_i \in (0, 1)$ for $i = 1, 2$.

Step 1 (Computation of the Newton Direction)

Let $(\Delta x^k, \Delta y^k)$ solve the equation

$$F_{\phi_{\mu_k}}(x^k, y^k) + \nabla F_{\phi_{\mu_k}}(x^k, y^k)^T \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = 0. \quad (4.11)$$

Step 2 (Backtracking Line Search)

If $\Phi_{\mu_k}(x^k, y^k) = 0$, set $(x^{k+1}, y^{k+1}) = (x^k, y^k)$; otherwise, let λ_k be the maximum of the values $1, \alpha_1, \alpha_1^2, \dots$ such that

$$\|\Phi_{\mu_k}(x^k + \lambda_k \Delta x^k, y^k + \lambda_k \Delta y^k)\|_{\infty} \leq (1 - \sigma_1 \lambda_k) \|\Phi_{\mu_k}(x^k, y^k)\|_{\infty}, \quad (4.12)$$

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

Step 3 (Update the Continuation Parameter)

Let γ_k be the maximum of the values $1, \alpha_2, \alpha_2^2, \dots$ such that

$$\|\Phi_{(1-\sigma_2\gamma_k)\mu_k}(x^{k+1}, y^{k+1})\|_{\infty} \leq \beta(1 - \sigma_2\gamma_k)\mu_k, \quad (4.13)$$

and set $\mu_{k+1} = (1 - \sigma_2\gamma_k)\mu_k$, $k = k + 1$, and return to Step 1.

Remarks 1. The line search in Step 2 is a standard backtracking line search.

2. The parameter σ_2 in Step 3 can be taken to be 1. However, in this case the choice $\gamma_k = 1$ will probably never be accepted. In our experiments, we have

found that by choosing $\sigma_2 \approx 1$ the choice $\gamma_k = 1$ is often taken and μ_k is rapidly reduced.

We now prove the algorithm is well-defined. The proof needs the following lemma.

Lemma 4.2.1 *Let $0 \leq \lambda \leq 1$ and $(\Delta x, \Delta y)$ be the solution to the equation*

$$F_{\phi_\mu}(x, y) + \nabla F_{\phi_\mu}(x, y)^T \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = 0.$$

Then

$$\|\Phi_\mu(x + \lambda\Delta x, y + \lambda\Delta y)\|_\infty \leq (1 - \lambda) \|\Phi_\mu(x, y)\|_\infty + \frac{\lambda^2}{\mu} \|(\Delta x, \Delta y)\|_\infty^2.$$

Proof By Taylor expansion, we have for $i \in \{1, \dots, n\}$

$$\begin{aligned} & \phi_\mu(x_i + \lambda\Delta x_i, y_i + \lambda\Delta y_i) \\ = & \phi_\mu(x_i, y_i) + \lambda \nabla \phi_\mu(x_i, y_i) \begin{pmatrix} \Delta x_i \\ \Delta y_i \end{pmatrix} + \frac{\lambda^2}{2} (\Delta x_i, \Delta y_i) \nabla^2 \phi_\mu(x, y) \begin{pmatrix} \Delta x_i \\ \Delta y_i \end{pmatrix} \\ = & (1 - \lambda) \phi_\mu(x_i, y_i) + \frac{\lambda^2}{2} (\Delta x_i, \Delta y_i) \nabla^2 \phi_\mu(x, y) \begin{pmatrix} \Delta x_i \\ \Delta y_i \end{pmatrix}. \end{aligned}$$

Using Lemma 2.1.3, we have

$$|\phi_\mu(x_i + \lambda\Delta x_i, y_i + \lambda\Delta y_i)| \leq (1 - \lambda) |\phi_\mu(x_i, y_i)| + \frac{\lambda^2}{2\mu} \|(\Delta x, \Delta y)\|_\infty^2,$$

and so

$$\|\Phi_\mu(x + \lambda\Delta x, y + \lambda\Delta y)\|_\infty \leq (1 - \lambda) \|\Phi_\mu(x, y)\|_\infty + \frac{\lambda^2}{\mu} \|(\Delta x, \Delta y)\|_\infty^2.$$

□

We are now in a position to show that the algorithm is well-defined and implementable when it is assumed that M is a P_0 matrix and the condition (B) holds.

Theorem 4.2.2 *Let $\beta > 0$ and $(x^k, y^k) \in \mathcal{N}_s(\beta, \mu_k)$ for some $\mu_k > 0$. Suppose that M is a P_0 matrix and that assumption (B) holds.*

1. *The Jacobian $\nabla F_{\phi_{\mu_k}}(x^k, y^k)$ is non-singular. Hence, the Newton step in Step 1 of the algorithm exists and is unique.*

2. *If $\Phi_{\mu_k}(x^k, y^k) \neq 0$ then $\lambda_k \geq \bar{\lambda}$, where*

$$\bar{\lambda} = \alpha_1 \tilde{\lambda} \quad \text{and} \quad \tilde{\lambda} = \min\left\{1, \frac{1 - \sigma_1}{\beta C^2}\right\}. \quad (4.14)$$

Hence, the backtracking procedure for evaluating λ_k in Step 2 is finitely terminating.

3. *$\gamma_k \geq \bar{\gamma}$, where*

$$\bar{\gamma} = \min\{1, \sigma_2^{-1} \alpha_2 \tilde{\gamma}\}, \quad \text{and} \quad \tilde{\gamma} = \frac{\sigma_1 \bar{\lambda} \beta}{2 + \beta}. \quad (4.15)$$

Hence, the backtracking procedure for evaluating γ_k in Step 3 is finitely terminating.

Proof 1. Since

$$\nabla F_{\phi_{\mu}}(x, y) = \begin{bmatrix} M & -I \\ \nabla_x \Phi_{\mu}(x, y) & \nabla_y \Phi_{\mu}(x, y) \end{bmatrix}, \quad (4.16)$$

where

$$\nabla_x \Phi_{\mu}(x, y) = \text{diag} \left(\frac{\partial \phi_{\mu}(x_i, y_i)}{\partial a} \right), \quad \text{and} \quad \nabla_y \Phi_{\mu}(x, y) = \text{diag} \left(\frac{\partial \phi_{\mu}(x_i, y_i)}{\partial b} \right),$$

with

$$\frac{\partial \psi_{\mu}(a, b)}{\partial a} = 1 - \frac{a - b}{\sqrt{(a - b)^2 + 4\mu^2}} \quad \text{and} \quad \frac{\partial \psi_{\mu}(a, b)}{\partial b} = 1 + \frac{a - b}{\sqrt{(a - b)^2 + 4\mu^2}}.$$

By Lemma 2.1.3, both $\nabla_x \Phi_{\mu}(x, y)$ and $\nabla_y \Psi_{\mu}(x, y)$ are positive diagonal matrices whenever $\mu > 0$. Therefore, if

$$\nabla F_{\psi_{\mu}}(x, y) \begin{bmatrix} u \\ v \end{bmatrix} = 0,$$

then

$$\begin{aligned} Mu - v &= 0, \\ \nabla_x \Phi_\mu(x, y)u + \nabla_y \Phi_\mu(x, y)v &= 0. \end{aligned}$$

It follows that

$$(\nabla_x \Phi_\mu(x, y) + \nabla_y \Phi_\mu(x, y)M)u = 0.$$

Since $\nabla_x \Phi_\mu(x, y)$ and $\nabla_y \Phi_\mu(x, y)$ are positive diagonal matrices and M is a P_0 matrix, $(\nabla_x \Phi_\mu(x, y) + \nabla_y \Phi_\mu(x, y)M)$ is a P matrix, so $u = 0$ and therefore $v = 0$. This establishes Part 1.

2. Let $(\Delta x^k, \Delta y^k)$ be chosen to satisfy the Newton equation (4.11). It follows from lemma 4.2.1 and Assumption (B) that

$$\begin{aligned} & \|\Phi_{\mu_k}(x^k + \lambda \Delta x^k, y^k + \lambda \Delta y^k)\|_\infty \\ & \leq (1 - \lambda) \|\Phi_{\mu_k}(x^k, y^k)\|_\infty + \frac{\lambda^2}{\mu_k} \|(\Delta x^k, \Delta y^k)\|_\infty^2 \\ & \leq (1 - \lambda) \|\Phi_{\mu_k}(x^k, y^k)\|_\infty + \frac{\lambda^2}{\mu_k} C^2 \|\Phi_{\mu_k}(x^k, y^k)\|_\infty^2 \\ & \leq (1 - \lambda) \|\Phi_{\mu_k}(x^k, y^k)\|_\infty + \frac{1}{\mu_k} \lambda^2 C^2 \beta \mu_k \|\Phi_{\mu_k}(x^k, y^k)\|_\infty \\ & = (1 - \lambda) \|\Phi_{\mu_k}(x^k, y^k)\|_\infty + \lambda^2 C^2 \beta \|\Phi_{\mu_k}(x^k, y^k)\|_\infty \\ & \leq (1 - \sigma_1 \lambda) \|\Phi_{\mu_k}(x^k, y^k)\|_\infty \quad \text{for all } \lambda \in [0, \bar{\lambda}], \end{aligned}$$

Therefore $\lambda_k \geq \bar{\lambda}$ with $\bar{\lambda} = \alpha_1 \bar{\lambda}$.

3. We consider two cases. If $\|\Phi_{\mu_k}(x^k, y^k)\|_\infty = 0$, then $x^{k+1} = x^k$ and $y^{k+1} = y^k$.

Thus

$$\begin{aligned} \frac{\|\Phi_{(1-\gamma)\mu_k}(x^{k+1}, y^{k+1})\|_\infty}{(1-\gamma)\mu_k} &= \frac{\|\Phi_{(1-\gamma)\mu_k}(x^k, y^k)\|_\infty}{(1-\gamma)\mu_k} \\ &\leq \frac{\Phi_{\mu_k}(x^k, y^k) + 2\gamma\mu_k}{(1-\gamma)\mu_k} \\ &= \frac{2\gamma\mu_k}{(1-\gamma)\mu_k} \\ &\leq \beta \quad \text{for all } \gamma \in [0, \tilde{\gamma}]. \end{aligned}$$

Next consider the case $\|\Phi_{\mu_k}(x^k, y^k)\|_\infty \neq 0$. From part 2, we have

$$\begin{aligned}
\frac{\|\Phi_{(1-\gamma)\mu_k}(x^{k+1}, y^{k+1})\|_\infty}{(1-\gamma)\mu_k} &\leq \frac{\|\Phi_{\mu_k}(x^{k+1}, y^{k+1})\|_\infty + 2\gamma\mu_k}{(1-\gamma)\mu_k} \\
&\leq \frac{(1-\sigma_1\bar{\lambda})\|\Phi_{\mu_k}(x^k, y^k)\|_\infty + 2\gamma\mu_k}{(1-\gamma)\mu_k} \\
&\leq \frac{(1-\sigma_1\bar{\lambda})\beta\mu_k + 2\gamma\mu_k}{(1-\gamma)\mu_k} \\
&= \frac{(1-\sigma_1\bar{\lambda})\beta + 2\gamma}{1-\gamma} \\
&\leq \beta \quad \text{for all } \gamma \in [0, \bar{\gamma}].
\end{aligned}$$

Therefore $\gamma_k \geq \bar{\gamma}$ with $\bar{\gamma} = \min\{1, \sigma_2^{-1}\alpha_2\bar{\gamma}\}$. \square

4.3 Global Linear Convergence

We now state and prove the global linear convergence result for the algorithm described in the previous section.

Theorem 4.3.1 *Suppose that M is a P_0 and Assumption (B) holds. Let (x^k, y^k, μ_k) be the sequence generated by the algorithm of the preceding section. Then*

(i) For $k = 0, 1, \dots$,

$$Mx^k - y^k + q = 0, \quad (4.17)$$

$$(x^k, y^k) \in \mathcal{N}_s(\beta, \mu_k), \quad (4.18)$$

$$(1 - \sigma_2\gamma_{k-1}) \dots (1 - \sigma_2\gamma_0)\mu_0 = \mu_k. \quad (4.19)$$

(ii) For all $k \geq 0$, we have

$$\gamma_k \geq \bar{\gamma} := \min\left\{1, \frac{\sigma_2^{-1}\alpha_2\sigma_1\bar{\lambda}\beta}{2 + \beta}\right\}, \quad (4.20)$$

where

$$\bar{\lambda} = \alpha_1 \min\left\{1, \frac{1 - \sigma_1}{\beta C^2}\right\},$$

and C is the constant defined in (3.21). Therefore, μ_k converges to 0 at a global linear rate.

(iii) The sequence $\{(x^k, y^k)\}$ is bounded and converges to a solution of $LCP(q, M)$.

Proof (i) We establish (4.17)–(4.19) by induction on k . Clearly these relations hold for $k = 0$. Now assume that they hold for some $k > 0$. By Theorem 4.2.2, the algorithm is well defined and so (4.18) and (4.19) hold with k replaced by $k + 1$. Since (4.11) is satisfied for all k with $Mx^0 - y^0 + q = 0$, we have that $Mx^k - y^k + q = 0$ for all k , and so, in particular, it is true when k replaced by $k + 1$. Hence, by induction, (4.17)–(4.19) hold for all k .

(ii) Follows from Theorem 4.2.2.

(iii) By Assumption (B) and part (ii), we have

$$\begin{aligned} \|(x^{k+1}, y^{k+1}) - (x^k, y^k)\|_\infty &= \lambda_k \|(\Delta x^k, \Delta y^k)\|_\infty \\ &\leq C \|\Phi_{\mu_k}(x^k, y^k)\|_\infty \\ &\leq C\beta\mu_k \\ &\leq C\beta(1 - \sigma_2\bar{\gamma})^k. \end{aligned}$$

Therefore $\{(x^k, y^k)\}$ is a Cauchy sequence and therefore it is bounded and converges to a point (x^*, y^*) . It follows from $(x^k, y^k) \in \mathcal{N}_s(\beta, \mu_k)$ that $(x^*, y^*) \in S$. \square

4.4 A Condition-Based Complexity Bound for LCP with a P -Matrix

In this section, we show that if M is a P -matrix, the constant C defined in (3.21) can be estimated in terms of the fundamental quantity of P -matrix. This fundamental quantity is then used to establish a condition-based complexity bound for the non-interior path following method for LCP with a P -Matrix.

It is well known that a matrix $M \in \mathbb{R}^{n \times n}$ is a P -matrix if and only if for every $x \in \mathbb{R}^n$ and $x \neq 0$,

$$\max_{1 \leq i \leq n} x_i (Mx)_i > 0.$$

In [52], Mathias and Pang introduce a fundamental quantity associated with a P -matrix.

$$l(M) := \min_{\|x\|_\infty=1} (\max_{1 \leq i \leq n} x_i (Mx)_i), \quad (4.21)$$

and show how this quantity is useful in deriving error bounds for the linear complementarity problem with a P -matrix. It is easy to see that $l(M)$ is well defined, finite, and positive. Moreover for any $x \in \mathbb{R}^n$,

$$l(M) \|x\|_\infty^2 \leq \max_{1 \leq i \leq n} x_i (Mx)_i. \quad (4.22)$$

In this section, we give a complexity bound in terms of the fundamental quantity $l(M)$. An important tool used in the complexity analysis is the principal pivotal transform of M .

Let α be a subset of $\{1, \dots, n\}$ and $\bar{\alpha}$ the complement set of α , that is $\bar{\alpha} = \{1, \dots, n\} \setminus \alpha$. By means of a principal rearrangement, we may assume that $M_{\alpha\alpha}$ is a leading principal submatrix of M . The *principal pivotal transform* of M with respect to the index set α , denoted by $\mathcal{P}_\alpha(M)$, is the matrix

$$\begin{bmatrix} M_{\alpha\alpha}^{-1} & -M_{\alpha\alpha}^{-1} M_{\alpha\bar{\alpha}} \\ M_{\bar{\alpha}\alpha} M_{\alpha\alpha}^{-1} & M_{\bar{\alpha}\bar{\alpha}} - M_{\bar{\alpha}\alpha} M_{\alpha\alpha}^{-1} M_{\alpha\bar{\alpha}} \end{bmatrix}. \quad (4.23)$$

The matrix

$$M_{\bar{\alpha}\bar{\alpha}} - M_{\bar{\alpha}\alpha} M_{\alpha\alpha}^{-1} M_{\alpha\bar{\alpha}} \quad (4.24)$$

is an instance of a *Schur complement*, and is denoted by $(M/M_{\alpha\alpha})$. The following properties of the principal pivotal transform $\mathcal{P}_\alpha(M)$ can be found in Cottle, Pang and Stone [20].

Lemma 4.4.1 *Let α be a subset of $\{1, \dots, n\}$. Suppose that the principal submatrix $M_{\alpha\alpha}$ is nonsingular and*

$$\begin{bmatrix} y_\alpha \\ y_{\bar{\alpha}} \end{bmatrix} = \begin{bmatrix} M_{\alpha\alpha} & M_{\alpha\bar{\alpha}} \\ M_{\bar{\alpha}\alpha} & M_{\bar{\alpha}\bar{\alpha}} \end{bmatrix} \begin{bmatrix} x_\alpha \\ x_{\bar{\alpha}} \end{bmatrix}, \quad (4.25)$$

then

$$\begin{bmatrix} x_\alpha \\ y_{\bar{\alpha}} \end{bmatrix} = \mathcal{P}_\alpha(M) \begin{bmatrix} y_\alpha \\ x_{\bar{\alpha}} \end{bmatrix}. \quad (4.26)$$

Lemma 4.4.2 *Let α be any subset of $\{1, \dots, n\}$. If M is a P -matrix, then so is $\mathcal{P}_\alpha(M)$.*

Let

$$\ell := \min\{l(\mathcal{P}_\alpha(M)) \mid \alpha \subseteq \{1, \dots, n\}\}. \quad (4.27)$$

Since there are only finite number of principal pivotal transforms of M , we have $\ell > 0$ and

$$\ell \|x\|_\infty^2 \leq \max_{1 \leq i \leq n} x_i (\mathcal{P}_\alpha(M)x)_i \quad \text{for all } x \in \mathbb{R}^n \text{ and } \alpha \subseteq \{1, \dots, n\}. \quad (4.28)$$

Lemma 4.4.3 *Let $(\Delta x^k, \Delta y^k)$ be the Newton direction (4.11) generated by the Algorithm in Section 4.2. Suppose that M is a P -matrix, then*

$$\|(\Delta x^k, \Delta y^k)\|_\infty \leq \left(1 + \frac{1}{\ell}\right) \|\Phi_{\mu_k}(x^k, y^k)\|_\infty, \quad (4.29)$$

where ℓ is defined in (4.27).

Proof The Newton direction $(\Delta x^k, \Delta y^k)$ satisfies the following equations

$$\begin{aligned} M\Delta x^k - \Delta y^k &= 0, \\ D_1\Delta x^k + D_2\Delta y^k &= -\Phi_{\mu_k}(x^k, y^k), \end{aligned}$$

where

$$D_1 = \text{diag} \left(1 - \frac{x_i^k - y_i^k}{\sqrt{(x_i^k - y_i^k)^2 + 4\mu_k^2}} \right) \quad \text{and} \quad D_2 = \text{diag} \left(1 + \frac{x_i^k - y_i^k}{\sqrt{(x_i^k - y_i^k)^2 + 4\mu_k^2}} \right).$$

Let

$$\alpha = \{i | x_i^k - y_i^k \leq 0\}.$$

By changing the role of $(\Delta x^k)_\alpha$ and $(\Delta y^k)_\alpha$, we may assume without loss of generality that $(\Delta x^k, \Delta y^k)$ satisfies

$$\tilde{M}\Delta x^k - \Delta y^k = 0, \quad (4.30)$$

$$D_1\Delta x^k + D_2\Delta y^k = -\Phi_{\mu_k}(x^k, y^k), \quad (4.31)$$

where $\tilde{M} = \mathcal{P}_\alpha(M)$ is the principal pivotal transform of M with respect to the index set α , and $x_i^k - y_i^k \geq 0$ for $i = 1, \dots, n$. Therefore

$$1 + \frac{x_i^k - y_i^k}{\sqrt{(x_i^k - y_i^k)^2 + 4\mu_k^2}} \geq 1 \quad \text{for } i = 1, \dots, n. \quad (4.32)$$

From equation (4.30), $\Delta y^k = \tilde{M}\Delta x^k$. Substituting Δy^k into (4.31), we have

$$D_1\Delta x^k + D_2\tilde{M}\Delta x^k = -\Phi_{\mu_k}(x^k, y^k).$$

Multiplying D_2^{-1} on both side of the equation yields

$$D_2^{-1}D_1\Delta x^k + \tilde{M}\Delta x^k = -D_2^{-1}\Phi_{\mu_k}(x^k, y^k). \quad (4.33)$$

Therefore

$$\begin{aligned} \ell \|\Delta x^k\|_\infty^2 &\leq \max_{1 \leq i \leq n} (\Delta x^k)_i (\tilde{M}\Delta x^k)_i \\ &\leq \max_{1 \leq i \leq n} (\Delta x^k)_i (D_2^{-1}D_1\Delta x^k + \tilde{M}\Delta x^k)_i \\ &\leq \max_{1 \leq i \leq n} |(\Delta x^k)_i| \max_{1 \leq i \leq n} |(D_2^{-1}\Phi_{\mu_k}(x^k, y^k))_i| \\ &= \|\Delta x^k\|_\infty \|D_2^{-1}\Phi_{\mu_k}(x^k, y^k)\|_\infty \\ &\leq \|\Delta x^k\|_\infty \|D_2^{-1}\|_\infty \|\Phi_{\mu_k}(x^k, y^k)\|_\infty \\ &\leq \|\Delta x^k\|_\infty \|\Phi_{\mu_k}(x^k, y^k)\|_\infty, \end{aligned}$$

and so

$$\|\Delta x^k\|_\infty \leq \frac{1}{\ell} \|\Phi_{\mu_k}(x^k, y^k)\|_\infty. \quad (4.34)$$

By (4.31) and (4.34), we have

$$\begin{aligned}
\|\Delta y^k\|_\infty &= \|-D_2^{-1}\Phi_{\mu_k}(x^k, y^k) - D_2^{-1}D_1\Delta x^k\| \\
&\leq \|D_2^{-1}\|_\infty \|\Phi_{\mu_k}(x^k, y^k)\|_\infty + \|D_2^{-1}\|_\infty \|D_1\|_\infty \|\Delta x^k\|_\infty \\
&\leq \|\Phi_{\mu_k}(x^k, y^k)\|_\infty + \|\Delta x^k\|_\infty \\
&\leq \left(1 + \frac{1}{\ell}\right) \|\Phi_{\mu_k}(x^k, y^k)\|_\infty.
\end{aligned} \tag{4.35}$$

Combining (4.34) and (4.35), we obtain

$$\|(\Delta x^k, \Delta y^k)\|_\infty \leq \left(1 + \frac{1}{\ell}\right) \|\Phi_{\mu_k}(x^k, y^k)\|_\infty.$$

□

Since our primary interest is to establish a complexity bound and examine how this bound depends on ℓ when ℓ is small, we may assume that $\ell < 1$. For simplicity, we further assume that $\sigma_2 = 1$ in the Algorithm. Then $\bar{\gamma}$ defined in (4.20) has the form

$$\bar{\gamma} = \frac{\alpha_2 \sigma_1 \bar{\lambda} \beta}{2 + \beta},$$

where

$$\bar{\lambda} = \frac{\alpha_1(1 - \sigma_1)\ell^2}{2\beta}.$$

Theorem 4.4.4 *Assume that M is a P -matrix and μ_0 is chosen such that $\mu_0 \leq \frac{1}{1+2\beta}$. Given $\epsilon > 0$, the algorithm of Section 4.2 finds a solution in the set*

$$\{(x, y) : Mx - y + q = 0, \|\min\{x, y\}\|_\infty \leq \epsilon\},$$

in

$$\mathcal{O}\left(\frac{\log |\epsilon|}{\ell^2}\right) \tag{4.36}$$

steps, where ℓ is defined in (4.27).

Proof It follows from Corollary 4.1.4 and Theorem 4.3.1 that

$$\begin{aligned}
\|\min\{x^k, Mx^k + q\}\|_\infty &= \|\min\{x^k, y^k\}\|_\infty \\
&\leq \mu_k + 2\|\Phi_{\mu_k}(x^k, y^k)\|_\infty \\
&\leq (1 + 2\beta)\mu_k \\
&= (1 + 2\beta)(1 - \gamma_{k-1}) \dots (1 - \gamma_0)\mu_0 \\
&\leq (1 + 2\beta)\mu_0 \left(1 - \frac{\alpha_2 \sigma_1 \bar{\lambda} \beta}{2 + \beta}\right)^k \\
&\leq \left(1 - \frac{\alpha_2 \sigma_1 \bar{\lambda} \beta}{2 + \beta}\right)^k \\
&= (1 - \delta \ell^2)^k,
\end{aligned}$$

where $\delta = \frac{\alpha_1 \alpha_2 \sigma_1 (1 - \sigma_1)}{2(2 + \beta)}$.

From a standard argument in complexity analysis, we have

$$k = \mathcal{O}\left(\frac{\log |\epsilon|}{\ell^2}\right).$$

□

4.5 A Condition-Based Complexity for LCP with a Symmetric Positive Definite Matrix

In this section, we establish a complexity bound for the non-interior path following method for LCP with a symmetric positive definite matrix. In this section, we make a blanket assumption that M is symmetric positive definite. Let $\lambda_1(M), \lambda_2(M), \dots, \lambda_n(M)$ denote the eigenvalues of M . We adopt the convention that $\lambda_{\min}(M) = \lambda_n(M) \leq \dots \leq \lambda_2(M) \leq \lambda_1(M) = \lambda_{\max}(M)$. One of the most elegant results in matrix theory is the Cauchy Interlacing Theorem which states that the eigenvalues of a real symmetric matrix are interlaced by the eigenvalues of any principal submatrix.

Theorem 4.5.1 (*Cauchy Interlacing Theorem*) *Let M be an $n \times n$ symmetric matrix*

and have the partitioned form

$$M = \begin{bmatrix} M_{\alpha\alpha} & M_{\alpha\bar{\alpha}} \\ M_{\bar{\alpha}\alpha} & M_{\bar{\alpha}\bar{\alpha}} \end{bmatrix}, \quad (4.37)$$

where $M_{\alpha\alpha} \in \mathbb{R}^{r \times r}$. Then

$$\lambda_{i+n-r}(M) \leq \lambda_i(M_{\alpha\alpha}) \leq \lambda_i(M), \quad \text{for } i = 1, 2, \dots, r. \quad (4.38)$$

The following theorem proved by Smith [73] shows that the eigenvalues of a Schur complement of a symmetric positive semidefinite matrix interlace the eigenvalues of the matrix itself.

Theorem 4.5.2 [73] *Let M be a $n \times n$ symmetric positive semidefinite matrix and have the partitioned form*

$$M = \begin{bmatrix} M_{\alpha\alpha} & M_{\alpha\bar{\alpha}} \\ M_{\bar{\alpha}\alpha} & M_{\bar{\alpha}\bar{\alpha}} \end{bmatrix}, \quad (4.39)$$

where $M_{\alpha\alpha} \in \mathbb{R}^{r \times r}$. Then

$$\lambda_{i+r}(M) \leq \lambda_i(M/M_{\alpha\alpha}) \leq \lambda_i(M), \quad \text{for } i = 1, 2, \dots, n-r. \quad (4.40)$$

Note that when M is symmetric and α is any subset of $\{1, \dots, n\}$, the principal pivotal transform $\mathcal{P}_\alpha(M)$ is skew symmetric, and

$$\widetilde{\mathcal{P}_\alpha(M)} := \frac{\mathcal{P}_\alpha(M) + \mathcal{P}_\alpha(M)^T}{2} = \begin{bmatrix} M_{\alpha\alpha}^{-1} & 0 \\ 0 & M_{\bar{\alpha}\bar{\alpha}} - M_{\bar{\alpha}\alpha} M_{\alpha\alpha}^{-1} M_{\alpha\bar{\alpha}} \end{bmatrix} \quad (4.41)$$

is symmetric. It is easy to see that if M is symmetric positive definite, then so is $\widetilde{\mathcal{P}_\alpha(M)}$.

Using Cauchy Interlacing Theorem and Theorem 4.5.2, we establish a lower bound for $\lambda_{\min}(\widetilde{\mathcal{P}_\alpha(M)})$.

Theorem 4.5.3 *Let M be symmetric positive definite. Then*

$$\lambda_{\min}(\widetilde{\mathcal{P}_\alpha(M)}) \geq \min\{\lambda_{\min}(M), 1/\lambda_{\max}(M)\}, \quad (4.42)$$

for any $\alpha \subseteq \{1, \dots, n\}$.

Proof It follows from (4.41) that

$$\lambda_{\min}(\widetilde{\mathcal{P}_\alpha(M)}) = \min\{\lambda_{\min}(M_{\alpha\alpha}^{-1}), \lambda_{\min}(M/M_{\alpha\alpha})\}.$$

By Cauchy Interlacing Theorem, we have

$$\lambda_{\min}(M_{\alpha\alpha}^{-1}) \geq 1/\lambda_{\max}(M),$$

and by Theorem 4.5.2, we have

$$\lambda_{\min}(M/M_{\alpha\alpha}) \geq \lambda_{\min}(M).$$

Therefore

$$\lambda_{\min}(\widetilde{\mathcal{P}_\alpha(M)}) \geq \min\{\lambda_{\min}(M), 1/\lambda_{\max}(M)\}.$$

□

Lemma 4.5.4 *Let $(\Delta x^k, \Delta y^k)$ be the Newton direction (4.11) generated by the Algorithm in Section 4.2. Suppose that M is symmetric positive definite. Then*

$$\|(\Delta x^k, \Delta y^k)\|_\infty \leq \left(1 + \frac{\sqrt{n}}{\min\{\lambda_{\min}(M), 1/\lambda_{\max}(M)\}}\right) \|\Phi_{\mu_k}(x^k, y^k)\|_\infty. \quad (4.43)$$

Proof The notations used here are the same as that used in Lemma 4.4.3. By (4.33), we have

$$D_2^{-1} D_1 \Delta x^k + \mathcal{P}_\alpha(M) \Delta x^k = -D_2^{-1} \Phi_{\mu_k}(x^k, y^k), \quad (4.44)$$

and so

$$\begin{aligned} & (\min\{\lambda_{\min}(M), 1/\lambda_{\max}(M)\}) \|\Delta x^k\|_2^2 \\ & \leq \lambda_{\min}(\widetilde{\mathcal{P}_\alpha(M)}) \|\Delta x^k\|_2^2 \\ & \leq (\Delta x^k)^T \widetilde{\mathcal{P}_\alpha(M)} \Delta x^k \\ & = (\Delta x^k)^T \mathcal{P}_\alpha(M) \Delta x^k \\ & \leq (\Delta x^k)^T (D_2^{-1} D_1 + \mathcal{P}_\alpha(M)) \Delta x^k \end{aligned}$$

$$\begin{aligned}
&\leq |(\Delta x^k)^T D_2^{-1} \Phi_{\mu_k}(x^k, y^k)| \\
&\leq \|\Delta x^k\|_2 \|D_2^{-1} \Phi_{\mu_k}(x^k, y^k)\|_2 \\
&\leq \sqrt{n} \|\Delta x^k\|_2 \|D_2^{-1} \Phi_{\mu_k}(x^k, y^k)\|_\infty \\
&\leq \sqrt{n} \|\Delta x^k\|_2 \|D_2^{-1}\|_\infty \|\Phi_{\mu_k}(x^k, y^k)\|_\infty \\
&\leq \sqrt{n} \|\Delta x^k\|_2 \|\Phi_{\mu_k}(x^k, y^k)\|_\infty.
\end{aligned}$$

Therefore

$$\|\Delta x^k\|_\infty \leq \|\Delta x^k\|_2 \leq \frac{\sqrt{n}}{\min\{\lambda_{\min}(M), 1/\lambda_{\max}(M)\}} \|\Phi_{\mu_k}(x^k, y^k)\|_\infty.$$

The rest of the proof is similar to that of Lemma 4.4.3. \square

Theorem 4.5.5 *Assume that M is a symmetric positive definite matrix and μ_0 is chosen such that $\mu_0 \leq \frac{1}{1+2\beta}$. Given $\epsilon > 0$, the algorithm of Section 4.2 finds a solution in the set*

$$\{(x, y) : Mx - y + q = 0, \|\min\{x, y\}\|_\infty \leq \epsilon\},$$

in

$$\mathcal{O}((\max\{1/\lambda_{\min}(M), \lambda_{\max}(M)\})^2 n \log |\epsilon|) \quad (4.45)$$

steps, where $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$ are the smallest and the largest eigenvalues of M respectively.

Proof Follows the proof of Theorem 4.4.4. \square

Remark

The condition number for symmetric positive definite matrix is $\lambda_{\max}(M)/\lambda_{\min}(M)$.

The number $\max\{1/\lambda_{\min}(M), \lambda_{\max}(M)\}$ gives a better bound in the sense that

$$\max\{1/\lambda_{\min}(M), \lambda_{\max}(M)\} \leq \frac{\lambda_{\max}(M)}{\lambda_{\min}(M)}$$

if $\lambda_{\max}(M) \geq 1$ and $\lambda_{\min}(M) \leq 1$.

Chapter 5

A POLYNOMIAL TIME INFEASIBLE INTERIOR POINT ALGORITHM FOR MONOTONE LCP

In this chapter, we establish the polynomial complexity of an interior point path following algorithm for LCP. The proposed algorithm can be viewed as an interior point variation on the non-interior path following algorithms studied in the previous two chapters. The algorithm has the same best polynomial-time complexity as is exhibited by the standard short-step interior point path following algorithm. The results of this chapter represent a first step toward understanding the relationship between interior and non-interior path following methods and provide a spring-board for discovering the polynomial complexity of the non-interior path following algorithms for LCP.

To compare the proposed algorithm with standard interior point algorithms, we use a scaled version of the smoothed Fischer-Burmerster function ψ and the CHKS smoothing function ϕ

$$\bar{\psi}_\mu(a, b) = \frac{a+b}{\sqrt{2}} - \sqrt{\frac{a^2+b^2}{2} + \mu}, \quad (5.1)$$

$$\bar{\phi}_\mu(a, b) = \frac{a+b}{2} - \sqrt{\frac{(a-b)^2}{4} + \mu}, \quad (5.2)$$

$$\theta_\mu(a, b) = ab - \mu. \quad (5.3)$$

It is easy to show that $\bar{\psi}_\mu(a, b) = 0$ (or $\bar{\phi}_\mu(a, b) = 0$) if and only if $0 \leq a$, $0 \leq b$, and $\theta_\mu(a, b) = 0$.

Using $\bar{\psi}_\mu$, $\bar{\phi}_\mu$ and θ_μ as building blocks, one defines the functions

$$\bar{\Psi}_\mu(x, y) = \begin{bmatrix} \bar{\psi}_\mu(x_1, y_1) \\ \dots \\ \bar{\psi}_\mu(x_n, y_n) \end{bmatrix}. \quad (5.4)$$

$$\bar{\Phi}_\mu(x, y) = \begin{bmatrix} \bar{\phi}_\mu(x_1, y_1) \\ \dots \\ \bar{\phi}_\mu(x_n, y_n) \end{bmatrix}. \quad (5.5)$$

$$\Theta_\mu(x, y) = \begin{bmatrix} \theta_\mu(x_1, y_1) \\ \dots \\ \theta_\mu(x_n, y_n) \end{bmatrix}. \quad (5.6)$$

The plan of this chapter follows. In section 5.1, we introduce a rescaled Newton direction and compare this direction with the one used in interior point literature. An interior point method based on the rescaled Newton direction is proposed in section 5.2. In section 5.3, we establish the global linear convergence and a polynomial complexity bound for the proposed interior point method for monotone LCP. If the method is initiated at an interior point that is also feasible with respect to the affine constraints, then the complexity bound is $O(\sqrt{n}L)$; otherwise, the complexity bound is $O(nL)$. Finally, in section 5.4, we point out the connection between the proposed algorithm and a framework developed by Mizuno.

5.1 The Rescaled Newton Directions

The first step in our analysis is to rescale the Newton step to yield iterates comparable to those of a standard interior point strategy. By analogy with the infeasible interior point strategies, at iteration k we compute a Newton step $(\Delta x^k, \Delta y^k)$ based on the equations

$$M\Delta x - \Delta y = -\gamma_k(Mx^k - y^k + q) \quad (5.7)$$

$$\nabla_x \bar{\Psi}_{\mu_k}(x^k, y^k) \Delta x + \nabla_y \bar{\Psi}_{\mu_k}(x^k, y^k) \Delta y = -\bar{\Psi}_{\mu_k}(x^k, y^k), \quad (5.8)$$

where $0 < \gamma_k < 1$.

$$\nabla_x \bar{\Psi}_{\mu_k}(x^k, y^k) := \text{diag} \left(\frac{1}{\sqrt{2}} - \frac{x_i^k}{2\sqrt{\frac{(x_i^k)^2 + (y_i^k)^2}{2} + \mu_k}} \right),$$

and

$$\nabla_y \bar{\Psi}_{\mu_k}(x^k, y^k) := \text{diag} \left(\frac{1}{\sqrt{2}} - \frac{y_i^k}{2\sqrt{\frac{(x_i^k)^2 + (y_i^k)^2}{2} + \mu_k}} \right).$$

Observe that if (x, y) is on the central path, then

$$\sqrt{\frac{x_i^2 + y_i^2}{2} + \mu} = \frac{x_i + y_i}{\sqrt{2}}, \text{ for } i = 1, \dots, n.$$

By replacing the expression $\sqrt{\frac{(x_i^k)^2 + (y_i^k)^2}{2} + \mu_k}$ in the definitions of the diagonal matrices $\nabla_x \bar{\Psi}_{\mu_k}(x^k, y^k)$ and $\nabla_y \bar{\Psi}_{\mu_k}(x^k, y^k)$ by the expression $\frac{x_i^k + y_i^k}{\sqrt{2}}$ and then multiplying (5.8) through by the diagonal matrix $\text{diag}(\sqrt{2}(x_i^k + y_i^k))$, we obtain the rescaled Newton equations

$$M \Delta x - \Delta y = -\gamma_k (M x^k - y^k + q) \quad (5.9)$$

$$Y^k \Delta x + X^k \Delta y = -2\hat{\Psi}_{\mu_k}(x^k, y^k). \quad (5.10)$$

where, for the sake of convenience, we define

$$\hat{\Psi}_{\mu}(x, y) := \text{diag} \left(\frac{x_i + y_i}{\sqrt{2}} \right) \bar{\Psi}_{\mu}(x, y).$$

The only difference between these rescaled Newton equations and the Newton equations used in a standard interior point path following strategy occurs in equation (5.10) where $2\hat{\Psi}_{\mu_k}(x^k, y^k)$ replaces the usual term $\Theta_{\mu_k}(x^k, y^k)$. The pattern of our development should now be clear. After a few identities and inequalities relating the functions $\hat{\Psi}_{\mu}(x, y)$ and $\Theta_{\mu}(x, y)$ have been established, a convergence theory and complexity analysis can be developed which is based on standard techniques from the theory of interior point path following methods. The necessary identities and inequalities are given in the next lemma.

Lemma 5.1.1 For $a, b, \mu \in \mathbb{R}$ satisfying $a > 0, b > 0, \mu > 0$, we have

$$\hat{\psi}_\mu(a, b) = \frac{(a+b)\theta_\mu(a, b)}{(a+b) + \sqrt{a^2 + b^2 + 2\mu}}, \quad (5.11)$$

$$\bar{\psi}_\mu^2(a, b) = 2\hat{\psi}_\mu(a, b) - \theta_\mu(a, b), \text{ and} \quad (5.12)$$

$$|\hat{\psi}_\mu(a, b)| \leq |\theta_\mu(a, b)|, \quad (5.13)$$

where $\hat{\psi}_\mu(a, b) := \frac{a+b}{\sqrt{2}}\bar{\psi}_\mu(a, b)$. In addition, given $0 < \beta < 1, 0 < \mu$, and $(x, y) \in \mathbb{R}_{++}^n \times \mathbb{R}_{++}^n$ satisfying

$$\|\Theta_\mu(x, y)\| \leq \beta\mu, \quad (5.14)$$

we have

$$\left\| 2\hat{\Psi}_\mu(x, y) - \Theta_\mu(x, y) \right\| \leq \frac{\beta^2\mu}{2(1-\beta)}. \quad (5.15)$$

Remarks 1. The identity (5.11) is due to B. Chen [6].

2. Inequality (5.13) implies that for $\mu > 0$ and $\beta > 0$, the condition (5.14) (used in standard interior point methods to define the neighborhood of the central path) implies that the condition

$$\left\| \hat{\Psi}_\mu(x, y) \right\| \leq \beta\mu$$

is also satisfied.

3. The identities (5.11) and (5.12), the inequalities (5.13) and (5.15), and the second remark remain valid with the expressions $\left[(a+b) + \sqrt{a^2 + b^2 + 2\mu} \right]$, $\hat{\psi}_\mu$, and $\hat{\Psi}_\mu$ replaced by $\left[(a+b) + \sqrt{(a-b)^2 + 4\mu} \right]$, $\hat{\phi}_\mu$ and $\hat{\Phi}_\mu$, respectively, where $\hat{\phi}_\mu(a, b) := \frac{a+b}{2}\bar{\phi}_\mu(a, b)$ and $\hat{\Phi}_\mu(x, y) := \text{diag} \left(\frac{x_i+y_i}{2} \right) \bar{\Phi}_\mu(x, y)$.
4. The bound (5.15) shows that the values $2\hat{\Psi}_{\mu_k}(x^k, y^k)$ approach the values $\Theta_{\mu_k}(x^k, y^k)$ used in the standard interior point methods as μ_k approaches 0. This partially explains why the interior point method based on the rescaled Newton direction studied in the next section has the same best polynomial-time complexity as the standard short step path-following interior point methods.

Proof For $a, b, \mu \in \mathbb{R}$ satisfying $a > 0, b > 0, \mu > 0$, the identity (5.11) is easily derived. The identity (5.12) and the inequality (5.13) follow readily from (5.11).

In order to see the bound (5.15), note that for any $(x, y) \in \mathbb{R}_{++}^n \times \mathbb{R}_{++}^n$ satisfying (5.14), we have

$$x_i y_i \geq (1 - \beta)\mu \quad \text{for } i = 1, 2, \dots, n,$$

and so

$$\frac{(x_i + y_i)^2}{2} = \frac{(x_i - y_i)^2 + 4x_i y_i}{2} \geq 2x_i y_i \geq 2(1 - \beta)\mu \quad \text{for } i = 1, 2, \dots, n. \quad (5.16)$$

It now follows from the identity (5.12), the inequality (5.13), and (5.16) that

$$\begin{aligned} \left\| 2\hat{\Psi}_\mu(x, y) - \Theta_\mu(x, y) \right\| &\leq \left\| \bar{\Psi}_\mu(x, y) \right\|^2 \leq \frac{\left\| \hat{\Psi}_\mu(x, y) \right\|^2}{2(1 - \beta)\mu} \leq \frac{\left\| \Theta_\mu(x, y) \right\|^2}{2(1 - \beta)\mu} \\ &\leq \frac{\beta^2 \mu^2}{2(1 - \beta)\mu} = \frac{\beta^2 \mu}{2(1 - \beta)}. \end{aligned}$$

5.2 The Algorithm

We present an algorithm based on the interior point algorithm proposed by Tseng [76].

An Infeasible Interior-Point Algorithm

Choose any $(\beta_1, \beta_2) \in \mathbb{R}^2$ satisfying

$$0 < \beta_1 < \beta_2 < 1, \quad \frac{2\beta_1}{1 - \beta_1} < \beta_2, \quad \frac{\beta_1^2}{2(1 - \beta_1)} + 2\beta_1\beta_2 + \beta_2^2(1 - \beta_1) < \beta_1, \quad (5.17)$$

and any $(x^0, y^0, \mu_0) \in \mathbb{R}_{++}^{2n+1}$ satisfying $\|\Theta_{\mu_0}(x^0, y^0)\| \leq \beta_1 \mu_0$. Let

$$\eta_1 = \frac{\beta_1 - \left[\frac{\beta_1^2}{2(1 - \beta_1)} + 2\beta_1\beta_2 + \beta_2^2(1 - \beta_1) \right]}{\sqrt{n} + \beta_1}. \quad (5.18)$$

For $k = 0, 1, \dots$, compute $(x^{k+1}, y^{k+1}, \mu_{k+1})$ from (x^k, y^k, μ_k) according to

$$x^{k+1} = x^k + \Delta x^k, \quad y^{k+1} = y^k + \Delta y^k, \quad \mu_{k+1} = (1 - \gamma_k)\mu_k, \quad (5.19)$$

where γ^k is the largest $\gamma \in (0, \eta_1]$ satisfying

$$\left\| 2\hat{\Psi}_{\mu_k}(x^k, y^k) + \gamma X^k(Mx^k - y^k + q) \right\| \leq \beta_2(\mu_k - \|\Theta_{\mu_k}(x^k, y^k)\|), \quad (5.20)$$

and $(\Delta x^k, \Delta y^k)$ is the unique vector in \mathbb{R}^{2n} satisfying

$$M\Delta x^k - \Delta y^k = -\gamma_k(Mx^k - y^k + q), \quad (5.21)$$

$$Y^k\Delta x^k + X^k\Delta y^k = -2\hat{\Psi}_{\mu_k}(x^k, y^k). \quad (5.22)$$

Remarks 1. To implement the algorithm using the function ϕ_μ , begin by selecting the parameters β_1 and β_2 so that

$$0 < \beta_1 < \beta_2 < 1, \quad \frac{2\beta_1}{1-\beta_1} < \beta_2, \quad \frac{\beta_1^2}{(1-\beta_1)} + 2\beta_1\beta_2 + \beta_2^2(1-\beta_1) < \beta_1. \quad (5.23)$$

Then set

$$\eta_1 = \frac{\beta_1 - \left[\frac{\beta_1^2}{(1-\beta_1)} + 2\beta_1\beta_2 + \beta_2^2(1-\beta_1) \right]}{\sqrt{n} + \beta_1} \quad (5.24)$$

and replace the function $\hat{\Psi}_{\mu_k}$ in (5.20) and (5.22) by the function $\hat{\Phi}_{\mu_k}$.

2. The set of pairs (β_1, β_2) satisfying either (5.17) or (5.23) is non-empty. In both cases, it follows that $\eta_1 > 0$. For a choice of β_1 and β_2 satisfying both (5.17) and (5.23), take $\beta_1 = 0.09, \beta_2 = 0.2$.

The following theorem shows that if the algorithm is initiated in the positive orthant, then it is well-defined and the iterates remain both in the positive orthant and the set $\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : \|\Theta_\mu(x, y)\| \leq \beta_1\mu\}$ for decreasing values of μ .

Theorem 5.2.1 Fix any $(\beta_1, \beta_2) \in \mathbb{R}^2$ satisfying (5.17). Let η_1 be given by (5.18). Suppose that $(x^k, y^k, \mu_k) \in \mathbb{R}_{++}^{2n+1}$ satisfies $\|\Theta_{\mu_k}(x^k, y^k)\| \leq \beta_1\mu_k$ and $(\Delta x^k, \Delta y^k)$ satisfies (5.21) and (5.22), with γ_k being the largest $\gamma \in (0, \eta_1]$ satisfying (5.20), then $\gamma_k > 0$ exists and

$$(x^k + \Delta x^k, y^k + \Delta y^k) > 0, \quad (5.25)$$

$$\|\Theta_{(1-\gamma_k)\mu_k}(x^k + \Delta x^k, y^k + \Delta y^k)\| \leq \beta_1(1-\gamma_k)\mu_k. \quad (5.26)$$

Proof For the sake of simplicity, denote $(x, y, \mu) = (x^k, y^k, \mu_k)$, $(\Delta x, \Delta y) = (\Delta x^k, \Delta y^k)$ and $\gamma = \gamma_k$ respectively. We first establish that $\gamma > 0$ exists. By Lemma 5.1.1, $\|2\hat{\Psi}_\mu(x, y)\| \leq 2\|\Theta_\mu(x, y)\| \leq 2\beta_1\mu$. By the choice of β_1 and β_2 in (5.17), we know that $2\beta_1 < \beta_2(1 - \beta_1)$. Therefore,

$$\|2\hat{\Psi}_\mu(x, y)\| < \beta_2\mu(1 - \beta_1) \leq \beta_2(\mu - \|\Theta_\mu(x, y)\|),$$

which implies that $\gamma > 0$ exists since a strict inequality holds in (5.20) when $\gamma = 0$.

Next set $r = 2\hat{\Psi}_\mu(x, y)$, $s = Mx - y + q$, and $z = X^{-1}\Delta x$ and $\beta = \left\|\frac{\Theta_\mu(x, y)}{\mu}\right\|$. Then the system (5.21) and (5.22) can be rewritten as

$$\begin{aligned} MXz - \Delta y &= -\gamma s, \\ YXz + X\Delta y &= -r. \end{aligned}$$

It follows that

$$(YX + XMX)z = -r - \gamma Xs.$$

Since M is positive semidefinite, we have

$$z^T YXz \leq z^T (YX + XMX)z = z^T (-r - \gamma Xs) \leq \|z\| \|r + \gamma Xs\|, \quad (5.27)$$

which implies that

$$\|z\| \leq \frac{\|r + \gamma Xs\|}{\min_i x_i y_i} \leq \frac{\|r + \gamma Xs\|}{\mu(1 - \beta)}, \quad (5.28)$$

where the second inequality follows from the inequality

$$Xy \geq (1 - \beta)\mu e, \quad (5.29)$$

which is itself a consequence of the relation $\|\Theta_\mu(x, y)\| = \|Xy - \mu e\| \leq \beta\mu$. By combining (5.28) with (5.20), we find that $\|z\| \leq \beta_2 < 1$. Thus, in particular, $e + z > 0$. Let $x' = x + \Delta x$ and $y' = y + \Delta y$. Hence $x' = x + Xz = X(e + z) > 0$, since $x > 0$. From Lemma 5.1.1 and (5.22), for each $i = 1, \dots, n$, we have

$$|(x_i + (\Delta x)_i)(y_i + (\Delta y)_i) - \mu|$$

$$\begin{aligned}
&= |x_i y_i - \mu + [x_i(\Delta y)_i + y_i(\Delta x)_i] + (\Delta x)_i(\Delta y)_i| \\
&= |2\hat{\psi}_\mu(x_i, y_i) - \frac{\hat{\psi}_\mu^2(x_i, y_i)}{\left(\frac{x_i+y_i}{\sqrt{2}}\right)^2} - 2\hat{\psi}_\mu(x_i, y_i) + (\Delta x)_i(\Delta y)_i| \\
&\leq \frac{\hat{\psi}_\mu^2(x_i, y_i)}{\frac{(x_i+y_i)^2}{2}} + |(\Delta x)_i(\Delta y)_i| \\
&\leq \frac{\hat{\psi}_\mu^2(x_i, y_i)}{2(1-\beta)\mu} + |(\Delta x)_i(\Delta y)_i|, \tag{5.30}
\end{aligned}$$

where (5.30) follows from the fact that $\frac{(x_i+y_i)^2}{2} \geq 2x_i y_i \geq 2(1-\beta)\mu$. Therefore,

$$\begin{aligned}
\|X'y' - \mu e\| &\leq \frac{1}{2(1-\beta)\mu} \left\| \begin{pmatrix} \hat{\psi}_\mu^2(x_1, y_1) \\ \dots \\ \hat{\psi}_\mu^2(x_n, y_n) \end{pmatrix} \right\| + \|ZX\Delta y\| \quad (\text{by (5.30)}) \\
&\leq \frac{1}{2(1-\beta)\mu} \left\| \begin{pmatrix} \hat{\psi}_\mu^2(x_1, y_1) \\ \dots \\ \hat{\psi}_\mu^2(x_n, y_n) \end{pmatrix} \right\|_1 + \|Z(-2\hat{\Psi}_\mu(x, y) - Y\Delta x)\| \quad (\text{by (5.22)}) \\
&\leq \frac{1}{2(1-\beta)\mu} \|\hat{\Psi}_\mu(x, y)\|^2 + 2\|Z\hat{\Psi}_\mu(x, y)\| + \|ZYXz\| \\
&\leq \frac{1}{2(1-\beta)\mu} \|\hat{\Psi}_\mu(x, y)\|^2 + 2\|Z\hat{\Psi}_\mu(x, y)\|_1 + \|ZYXz\|_1 \\
&\leq \frac{1}{2(1-\beta)\mu} \|\hat{\Psi}_\mu(x, y)\|^2 + 2\|z\| \|\hat{\Psi}_\mu(x, y)\| + z^T Y X z \\
&\leq \frac{(\beta\mu)^2}{2(1-\beta)\mu} + 2\beta\beta_2\mu + \|z\| \|r + \gamma Xs\| \quad (\text{Lemma 5.1.1 and (5.27)}) \\
&\leq \frac{(\beta\mu)^2}{2(1-\beta)\mu} + 2\beta\beta_2\mu + \beta_2\beta_2\mu(1-\beta) \quad (\text{by (5.20)}) \\
&\leq \frac{\beta_1^2\mu}{2(1-\beta_1)} + 2\beta_1\beta_2\mu + \beta_2^2(1-\beta_1)\mu, \tag{5.31}
\end{aligned}$$

where (5.31) follows from the fact that $\beta \leq \beta_1$ and

$$\begin{aligned}
2\beta\beta_2\mu + \beta_2^2(1-\beta)\mu &= 2\beta\beta_2\mu + \beta_2^2\mu - \beta_2^2\beta\mu \\
&= \beta(2\beta_2 - \beta_2^2)\mu + \beta_2^2\mu \\
&\leq \beta_1(2\beta_2 - \beta_2^2)\mu + \beta_2^2\mu \\
&= 2\beta_1\beta_2\mu + \beta_2^2(1-\beta_1)\mu.
\end{aligned}$$

Therefore, by (5.17) and (5.31), $\|X'y' - \mu e\| \leq \beta_1 \mu$. It follows from $x' > 0$ and $\beta_1 < 1$ that $y' > 0$. The triangle inequality, (5.31), and the inequality $\gamma < \eta_1$ now imply that

$$\begin{aligned} \frac{\|X'y' - (1-\gamma)\mu e\|}{(1-\gamma)\mu} &\leq \frac{\|X'y' - \mu e\|}{(1-\gamma)\mu} + \frac{\gamma\sqrt{n}}{1-\gamma} \\ &\leq \frac{\frac{\beta_1^2}{2(1-\beta_1)} + 2\beta_1\beta_2 + \beta_2^2(1-\beta_1)}{1-\gamma} + \frac{\gamma\sqrt{n}}{1-\gamma} \\ &\leq \frac{\beta_1 - \eta_1(\sqrt{n} + \beta_1)}{1-\eta_1} + \frac{\eta_1\sqrt{n}}{1-\eta_1} = \beta_1. \end{aligned}$$

5.3 Global Linear Convergence and Polynomial Complexity

The following global linear convergence result is patterned on [76, Theorem 3.1].

Lemma 5.3.1 [57, Lemma 3.3] *Assume that $S \neq \emptyset$. For any $\mu \in [0, 1]$, any $(x^*, y^*) \in S$ and any $(x, y) \in R^{2n}$ satisfying $(x, y) \geq 0$ and $Mx - y + q = \mu(Mx^0 - y^0 + q)$, we have*

$$\mu(x^T y^0 + y^T x^0) \leq x^T y + \mu((x^0)^T y^0 + (x^*)^T y^0 + (x^0)^T y^*), \quad (5.32)$$

$$(1-\mu)(x^T y^* + y^T x^*) \leq x^T y + \mu((x^0)^T y^0 + (x^*)^T y^0 + (x^0)^T y^*). \quad (5.33)$$

Theorem 5.3.2 *Let S denote the set of solutions to LCP:*

$$S := \{(x, y) : 0 \leq x, 0 \leq y, y = Mx + q, \text{ and } x^T y = 0\},$$

and let β_1, β_2, η_1 and $\{(x^k, y^k, \mu_k, \gamma_k)\}_{k=0,1,\dots}$ be generated by the Algorithm of Section 5.2. Then

$$0 < (x^k, y^k), \quad (5.34)$$

$$\beta_1 \mu_k \geq \|\Theta_{\mu_k}(x^k, y^k)\|, \quad \text{and} \quad (5.35)$$

$$\frac{\mu_k}{\mu_0}(Mx^0 - y^0 + q) = Mx^k - y^k + q, \quad (5.36)$$

for all k , where for $k > 0$

$$\mu_k = (1 - \gamma_{k-1}) \dots (1 - \gamma_0) \mu_0. \quad (5.37)$$

Moreover, the sequence $\{(x^k, y^k)\}$ is bounded if and only if the solution set S is non-empty, in which case, for any $(x^*, y^*) \in S$, we have $\gamma_k \geq \min\{\eta_1, \eta_2\}$ for all k , where

$$\eta_2 = \begin{cases} \frac{[\beta_2(1-\beta_1)-2\beta_1]\mu^0 \min_i y_i^0}{[(1+\beta_1)n\mu^0+(x^0)^T y^0+(x^*)^T y^0+(x^0)^T y^*] \|Mx^0-y^0+q\|_\infty} & \text{if } Mx^0 - y^0 + q \neq 0, \\ \infty & \text{if } Mx^0 - y^0 + q = 0. \end{cases} \quad (5.38)$$

Proof It is easily seen that (5.34), (5.35) and (5.36) hold for $k = 0$. Assume that (5.34), (5.35) and (5.36) hold for some $k \geq 0$. We first show that γ_k is well defined. Since when $\gamma = 0$, by Proposition 5.1.1, we have

$$\|\bar{\Psi}_{\mu_k}(x^k, y^k)\| \leq \|\Theta_{\mu_k}(x^k, y^k)\| = \mu_k \alpha_k,$$

here

$$\alpha_k := \left\| \frac{\Theta_{\mu_k}(x^k, y^k)}{\mu_k} \right\| = \left\| \frac{X^k y^k - \mu_k e}{\mu_k} \right\| \leq \beta_1. \quad (5.39)$$

It follows from $\beta_2 > \frac{2\beta_1}{1-\beta_1} \geq \frac{2\alpha_k}{1-\alpha_k}$ that $2\alpha_k < \beta_2(1 - \alpha_k)$. So

$$\begin{aligned} \|2\bar{\Psi}_{\mu_k}(x^k, y^k)\| &\leq 2\mu_k \alpha_k \\ &< \mu_k \beta_2 (1 - \alpha_k) \\ &= \beta_2 (\mu_k - \|X^k y^k - \mu_k e\|). \end{aligned}$$

Therefore γ_k is well defined.

By Theorem 5.2.1, (5.34) and (5.35) hold when k is replaced by $k + 1$. It follows from (5.21) that

$$\begin{aligned} M(x^k + \Delta x^k) - (y^k + \Delta y^k) + q &= (Mx^k - y^k + q) + (M\Delta x^k - \Delta y^k) \\ &= (Mx^k - y^k + q) - \gamma_k(Mx^k - y^k + q) \\ &= (1 - \gamma_k)(Mx^k - y^k + q) \\ &= (1 - \gamma_k) \frac{\mu_k}{\mu_0} (Mx^0 - y^0 + q) \\ &= \frac{\mu_{k+1}}{\mu_0} (Mx^0 - y^0 + q). \end{aligned}$$

Therefore (5.36) holds when k is replaced by $k + 1$. By induction, (5.34), (5.35) and (5.36) hold for all $k \geq 0$.

For any k , we have from $\|\frac{\Theta_{\mu_k}(x^k, y^k)}{\mu_k}\| \leq \beta_1$ that $X^k y^k \leq (1 + \beta_1)\mu_k e$. So

$$(x^k)^T y^k \leq (1 + \beta_1)n\mu_k. \quad (5.40)$$

It follows from triangle inequality and Lemma 5.1.1 that for any $\gamma \in [0, \infty)$, we have

$$\begin{aligned} & \frac{\|2\bar{\Psi}_{\mu_k}(x^k, y^k) + \gamma X^k(Mx^k - y^k + q)\|}{\mu_k} \\ & \leq \frac{2\|X^k y^k - \mu_k e\|}{\mu_k} + \frac{\gamma\|X^k(Mx^k - y^k + q)\|}{\mu_k} \\ & = 2\alpha_k + \frac{\gamma\|X^k(Mx^k - y^k + q)\|}{\mu_k}, \end{aligned}$$

here α_k is defined by (5.39). Let $\bar{\gamma}_k$ be the largest γ for which the lefthand side is below $\beta_2(1 - \frac{\|X^k y^k - \mu_k e\|}{\mu_k})$ (so $\gamma_k = \min\{\eta_1, \bar{\gamma}_k\}$). Then $\bar{\gamma}_k$ must exceed the largest γ for which the righthand side is below $\beta_2(1 - \frac{\|X^k y^k - \mu_k e\|}{\mu_k})$. This yields

$$\begin{aligned} \bar{\gamma}_k & \geq \frac{[\beta_2(1 - \alpha_k) - 2\alpha_k]\mu_k}{\|X^k(Mx^k - y^k + q)\|} \\ & \geq \frac{[\beta_2(1 - \beta_1) - 2\beta_1]\mu_k}{\|x^k\|_1 \|Mx^k - y^k + q\|_\infty} \\ & = \frac{[\beta_2(1 - \beta_1) - 2\beta_1]\mu_0}{\|x^k\|_1 \|Mx^0 - y^0 + q\|_\infty}, \end{aligned}$$

here the second inequality follows from $\alpha_k \leq \beta_1$ and the equality follows from $Mx^k - y^k + q = \frac{\mu_k}{\mu_0}(Mx^0 - y^0 + q)$.

Assume that $\{(x^k, y^k)\}$ is bounded. The the above inequalities yield that $\bar{\gamma}_k \geq \eta$ for some constant $\eta > 0$ and so $\gamma_k = \min\{\eta_1, \bar{\gamma}_k\} \geq \min\{\eta_1, \eta\}$ for all k . By (5.37), $\{\mu_k\} \rightarrow 0$. Since $(x^k, y^k) > 0$, $Mx^k - y^k + q = \frac{\mu_k}{\mu_0}(Mx^0 - y^0 + q)$ and (5.40) holds, this implies that any cluster point of $\{(x^k, y^k)\}$ is in S , so $S \neq \emptyset$.

Assume that $S \neq \emptyset$ and fix any $(x^*, y^*) \in S$. By (5.34-5.36) and Lemma 5.3.1, we have

$$\|x^k\|_1(\min_i y_i^0) + \|y^k\|_1(\min_i x_i^0)$$

$$\begin{aligned}
&\leq (x^k)^T y^0 + (y^k)^T x^0 \\
&\leq (x^k)^T y^k \mu_0 / \mu_k + (x^0)^T y^0 + (x^*)^T y^0 + (x^0)^T y^* \\
&\leq (1 + \beta_1) n \mu_0 + (x^0)^T y^0 + (x^*)^T y^0 + (x^0)^T y^*,
\end{aligned}$$

where the first inequality follows from the nonnegativity of (x^k, y^k) and $(x^0, y^0) > 0$, and the last inequality follows from (5.40). Since $(x^0, y^0) > 0$, this shows that $\{(x^k, y^k)\}$ is bounded. We also have

$$\begin{aligned}
\bar{\gamma}_k &\geq \frac{[\beta_2(1 - \beta_1) - 2\beta_1]\mu_0}{\|x^k\|_1 \|Mx^0 - y^0 + q\|_\infty} \\
&\geq \frac{[\beta_2(1 - \beta_1) - 2\beta_1]\mu_0 \min_i y_i^0}{[(1 + \beta_1)n\mu_0 + (x^0)^T y^0 + (x^*)^T y^0 + (x^0)^T y^*] \|Mx^0 - y^0 + q\|_\infty}.
\end{aligned}$$

The righthand side is exactly η_2 , so $\gamma_k = \min\{\eta_1, \bar{\gamma}_k\} \geq \min\{\eta_1, \eta_2\}$.

Thus, if S is nonempty, the Algorithm of Section 5.2 forces μ_k to zero at a global linear rate with the convergence ratio less than $1 - \min\{\eta_1, \eta_2\}$. Therefore, by standard results in the interior point literature (e.g., see [44]), one can find an element of S in $O((\min\{\eta_1, \eta_2\})^{-1}L)$ iterations, where L denotes the size of the binary encoding of the problem. It is easily seen that $\eta_1^{-1} = O(\sqrt{n})$, so it only remains to estimate η_2^{-1} . In the case where (x^0, y^0, μ_0) is chosen so that η_2^{-1} is $O(\sqrt{n})$ (such as when $Mx^0 - y^0 + q = 0$), the iteration count is $O(\sqrt{n}L)$. In the case where (x^0, y^0, μ_0) is the standard choice

$$\begin{aligned}
x^0 &= \rho_p e, \quad y^0 = \rho_d e, \quad \mu_0 = \rho_p \rho_d, \\
\rho_p &\geq \frac{\|x^*\|_1}{n}, \quad \rho_d \geq \max\left\{\frac{\|y^*\|_1}{n}, \|\rho_p M e + q\|_\infty\right\},
\end{aligned}$$

where (x^*, y^*) is any element of S , the formula (5.38) yields

$$\eta_2^{-1} \leq \frac{3(4 + \beta_1)n}{\beta_2(1 - \beta_1) - 2\beta_1},$$

so the iteration count is $O(nL)$.

5.4 The Connection to Standard Interior Point Methods

The interior point path following method studied in this chapter is essentially a variation on standard interior point methods wherein the right hand side in the Newton equations is perturbed in a very special way. For this reason, it is possible to analyze the algorithm within the framework developed by Mizuno. In [55, 56], Mizuno proposed a class of feasible interior point algorithms for monotone LCP which are based on the search direction $(\Delta x^k, \Delta y^k)$ satisfying the following equations

$$M\Delta x - \Delta y = 0, \quad (5.41)$$

$$Y^k\Delta x + X^k\Delta y = v^k - \sigma X^k y^k. \quad (5.42)$$

where $v^k \in \mathbb{R}_{++}^n$ and $\sigma > 0$. By adjusting the choice of the sequences $\{v^k\}$ with $\sigma = 1$, Mizuno is able to construct both path following and potential reduction methods and thereby provides a unifying framework within which a number of interior point methods can be studied. In order for this program to work, one must first show that the sequence $\{v^k\}$ satisfies the following three properties:

(a) $v^k > 0$ for $k = 0, 1, \dots$

(b) the sequence $\{v^k\}$ is an α -sequence for some $\alpha \geq 0$, that is, $v^{k+1} \in \mathcal{N}(v^k, \alpha)$ for all $k = 0, 1, 2, \dots$, where $\mathcal{N}(v, \alpha) = \{u \in \mathbb{R}^n : \|V^{-0.5}(v - u)\| \leq \alpha\sqrt{v_{\min}}\}$, with $V = \text{diag}(v)$, and

(c) there is an iteration index $m = O(\sqrt{n}L)$ such that $0 \leq v^m \leq 2^{-2L+1}e$.

The algorithm of Section 3 can be cast within Mizuno's framework. To see this, define

$$v^k = 2X^k y^k - 2\hat{\Psi}_{\mu_k}(x^k, y^k). \quad (5.43)$$

With this definition, the Newton equations (5.41) and (5.42) are identical to the equations (5.9) and (5.10) when $\sigma = 2$. If one now assumes that $\beta \in (0, \frac{1}{2}]$ and

$(\mu_k - \mu_{k+1})/\mu_k = O(1/\sqrt{n})$, it can be shown that the sequence $\{v^k\}$ defined by 5.43 satisfies the conditions (a) and (b). The condition $\beta \in (0, \frac{1}{2}]$ can be enforced during the initialization phase of the algorithm. It is used to show that

$$\frac{1}{2}|\theta_\mu(x_i, y_i)| \leq |\hat{\psi}_\mu(x_i, y_i)| \leq |\theta_\mu(x_i, y_i)|, \text{ for } i = 1, 2, \dots, n$$

whenever $(x, y) \in \mathbb{R}_{++}^n \times \mathbb{R}_{++}^n$ and

$$\|\Theta_\mu(x, y)\| \leq \beta\mu,$$

which in turn shows that condition (a) is satisfied. The bounds $\min\{\eta_1, \eta_2\} \leq \gamma_k \leq \eta_1$ (Theorem 5.3.2 and (5.18)) show that $O(1/\sqrt{n}) = \gamma_k$ if (x^0, y^0, μ_0) is chosen so that $\eta_2 = O(1/\sqrt{n})$ (for example, when $y^0 = Mx^0 + q$). This in turn implies that the condition $(\mu_k - \mu_{k+1})/\mu_k = O(1/\sqrt{n})$ is also satisfied. Finally, condition (c) can be verified using the complexity result established in this paper. This connection to Mizuno's work should provide a basis for developing a deeper understanding of the relationship between standard path following methods, potential reduction methods, and the path following method proposed in this chapter.

Chapter 6

NUMERICAL EXPERIMENTS

In this chapter, we present some preliminary numerical results on several implementations of non-interior path-following methods for LCP. The implementations are based on the algorithms studied in Chapters 3 and 4. The experiments are designed to test the effectiveness of the search directions introduced section 3.1, the sensitivity of the algorithms when the matrix in the LCP has high rank-deficiency, and the behavior of the algorithms when the feasibility to the affine constraints is not enforced. The plan of this chapter is as follows.

In section 6.1, we list the test problems used in our numerical experiments. In section 6.2, we present the numerical results for an implementation of the predictor-corrector algorithm introduced in section 3.2. In section 6.3, we give the numerical results on the algorithm studied in section 4.2. Both algorithms are feasible algorithms in the sense that every iterate (x^k, y^k) satisfies the affine constraints $Mx^k - y^k + q = 0$. The behaviors of an infeasible implementation of the second algorithm are examined in section 6.4. Finally in section 6.5, we draw a few conclusions from these numerical experiments.

6.1 The Test Problems

Example 6.1.1 (*Murty 1988*): n variables,

$$M = \begin{pmatrix} 1 & 2 & 2 & \dots & 2 \\ 0 & 1 & 2 & \dots & 2 \\ 0 & 0 & 1 & \dots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}, \quad (6.1)$$

$$q = (-1, \dots, -1)^T.$$

This is a standard test problem, for which both Lemke's complementarity pivot algorithm and Cottle and Danzig's principal pivoting method are known to run in exponential time. The solution is $x^* = (0, \dots, 0, 1)^T$, $y^* = (1, \dots, 1, 0)^T$. The matrix in this example is a P -matrix.

Example 6.1.2 (*Fathi 1979*): n variables,

$$M = \begin{pmatrix} 1 & 2 & 2 & \dots & 2 \\ 2 & 5 & 6 & \dots & 6 \\ 2 & 6 & 9 & \dots & 10 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 2 & 6 & 9 & \dots & 4(n-1) + 1 \end{pmatrix}, \quad (6.2)$$

$$q = (-1, \dots, -1)^T.$$

This is another standard test problem. Again, the complementarity pivot algorithm and the principal pivoting method are known to run in exponential time. The solution is $x^* = (1, 0, \dots, 0)^T$, $y^* = (0, 1, \dots, 1)^T$. The matrix M of this example is positive definite.

Example 6.1.3 (*Harker and Pang 1990*)

The matrix M is computed as follows: Let $A, B \in \mathbb{R}^{n \times n}$ and $q, d \in \mathbb{R}^n$ be randomly generated such that $a_{ij}, b_{ij} \in (-5, 5)$, $q_i \in (-500, 500)$, $d_i \in (0.0, 0.3)$ and that B is skew-symmetric. Define $M = A^T A + B + \text{diag}(d)$. Then M is a positive definite matrix.

Example 6.1.4 (*Harker and Pang “hard examples” 1990*)

In this example, M is computed in the same way as in the previous example. however, $q \in \mathbb{R}^n$ is randomly generated with entries $q_i \in (-500, 0)$.

Example 6.1.5

Given n and $k < n$, the matrix M is computed as follows: Let $A \in \mathbb{R}^{k \times n}$ and $B \in \mathbb{R}^{n \times n}$ be randomly generated such that $a_{ij}, b_{ij} \in (-5, 5)$ and that B is skew-symmetric. Define $M = A^T A + B$. Let $x, y \in \mathbb{R}^n$ be randomly generated such that $x(i), y(i) \in (0, 10)$ and $x(i) * y(i) = 0$ for $i = 1, \dots, n$. Set $q = Mx - y$. Then M is positive semi-definite and the LCP has a known solution (x, y) . This problem is used to test the sensitivity of the algorithm to the rank-deficiency of the matrix M . The matrix M tends to be more singular when a smaller k is chosen.

6.2 An Implementation Based on the Combined Direction

The algorithm implemented in this section is based the on the predictor-corrector algorithm studied in Chapter 3, but only the corrector step is used. As we pointed out in Chapter 3, the global linear convergence depends only on the corrector step and is independent of whether or not the predictor step is implemented on any given iteration. The search direction used in the corrector step is a combination of predictor direction and centering direction with a weighting parameter σ_k . A larger $\sigma_k \in (0, 1)$ means a higher weight has been put on the predictor direction, while a smaller $\sigma_k \in (0, 1)$ means a higher weight on the centering direction. The implemented

algorithm dynamically adjust the weighting parameter σ_k depending on how good is the parameter in the previous step.

Algorithm 1: An Implementation Using the Combined Direction

Step 0: (Initialization)

Choose $x^0 \in \mathbb{R}^n$, set $y^0 = Mx^0 + q$, and let $\mu_0 > 0$ be such that $\Phi(x^0, y^0, \mu_0) < 0$.

Choose $\beta > 0$ so that $\|\Phi(x^0, y^0, \mu_0)\|_\infty \leq \beta\mu_0$. We now have $(x^0, y^0) \in \mathcal{N}(\beta, \mu_0)$.

Choose σ_0, ϵ and α from $(0, 1)$.

Step 1: (The Corrector Step)

Let $(\Delta x^k, \Delta y^k, \Delta \mu_k)$ solve the equation

$$M\Delta x^k - \Delta y^k = 0,$$

$$\nabla_x \Phi(x^k, y^k, \mu_k) \Delta x^k + \nabla_y \Phi(x^k, y^k, \mu_k) \Delta y^k = -\Phi(x^k, y^k, \mu_k) + \sigma_k \mu_k \nabla_\mu \Phi(x^k, y^k, \mu_k). \quad (6.3)$$

and let $\lambda_k = \alpha^{l_k}$ be the maximum of the value $1, \alpha, \alpha^2, \dots$, such that

$$\|\Phi(x^k + \lambda_k \Delta x^k, y^k + \lambda_k \Delta y^k, (1 - \sigma_k \lambda_k) \mu_k)\|_\infty \leq (1 - \sigma_k \lambda_k) \beta \mu_k. \quad (6.4)$$

Set

$$x^{k+1} = x^k + \lambda_k \Delta x^k, \quad y^{k+1} = y^k + \lambda_k \Delta y^k, \quad \mu_{k+1} = (1 - \sigma_k \lambda_k) \mu_k, \quad (6.5)$$

$$\sigma_{k+1} := \begin{cases} 1 - (1 - \sigma_k)^3 & \text{if } l_k = 0, \\ \sigma_k & \text{if } l_k = 1, \\ 0.5\sigma_k & \text{if } l_k > 1, \end{cases} \quad (6.6)$$

and $k := k + 1$. If $\|\min\{x^k, y^k\}\|_\infty \leq \epsilon$, stop; otherwise, return to Step 1.

In our implementation, we choose $\epsilon = 10^{-6}$, $\sigma_0 = 0.01$ and $\alpha = 0.6$. Given x^0 , let $y^0 = Mx^0 + q$ and

$$\beta = \frac{\|\Phi(x^0, y^0, \mu_0)\|_\infty}{\mu_0}.$$

We now report the numerical results of the algorithm on the LCP test problems listed in section 6.1.

Example 6.1.1

We take $x^0 = (1, \dots, 1)^T$ and $\mu_0 = 10^{-6}$ as our starting point. The numerical results for this test problem can be found in Table 1.

Example 6.1.2

We take $x^0 = (1, \dots, 1)^T$ and $\mu_0 = 10^{-6}$ as our starting point. The numerical results for this test problem can be found in Table 2.

Example 6.1.3

Ten problems are generated for each of the dimensions $n = 50, 100, 150, 200$. The maximum, average, and minimum number of iterations needed by the algorithm are summarized in Table 3. In all runs, the starting point is chosen to be $x^0 = (0, \dots, 0)^T$ and $\mu_0 = 10^{-6}$.

Example 6.1.4

Ten problems are generated for each of the dimensions $n = 50, 100, 150, 200$. The maximum, average, and minimum number of iterations needed by the algorithm are summarized in Table 4. In all the test runs, the starting point is chosen to be $x^0 = (0, \dots, 0)^T$ and $\mu_0 = 10^{-6}$.

Example 6.1.5

Choose $n = 100$. Ten problems are generated for each of the choices $k = 90, 80, 70, 60, 50, 40, 30, 20, 10$. The maximum, average, and minimum number of iterations needed by the algorithm are summarized in Table 5. In all runs, the starting point is chosen to be $x^0 = (0, \dots, 0)^T$ and $\mu_0 = 1$.

From these results, we observe that the algorithm performs extremely well when M is positive definite matrix or a P -matrix. The algorithm perform well when the rank deficiency of M is not too large. With the increase on the singularity of M , more iterations are needed for the algorithm. When $k < 20$, the algorithm tends to stall and fails to produce a solution. This suggests that the combined direction is a good direction when the matrix is a positive definite matrix or a P -matrix and is a poor direction when M has high rank-deficiency.

6.3 An Implementation Based on the Centering Direction

The algorithm implemented in this section is based the on the algorithm studied in section 4.2. The algorithm uses a centering direction in each iteration.

Algorithm 2: An Implementation Using Centering Direction

Step 0 (Initialization)

Let $\mu_0 > 0$, $\beta > 0$, and $(x^0, y^0) \in \mathbb{R}^{2n}$ be given so that $(x^0, y^0) \in \mathcal{N}_s(\beta, \mu_0)$, and choose $\sigma_i \in (0, 1]$, $\alpha_i \in (0, 1)$ for $i = 1, 2$ and $\epsilon \in (0, 1)$.

Step 1 (Computation of the Newton Direction)

Let $(\Delta x^k, \Delta y^k)$ solve the equation

$$F_{\phi_{\mu_k}}(x^k, y^k) + \nabla F_{\phi_{\mu_k}}(x^k, y^k)^T \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = 0. \quad (6.7)$$

Step 2 (Backtracking Line Search)

If $\Phi_{\mu_k}(x^k, y^k) = 0$, set $(x^{k+1}, y^{k+1}) = (x^k, y^k)$; otherwise, let λ_k be the maximum of the values $1, \alpha_1, \alpha_1^2, \dots$ such that

$$\|\Phi_{\mu_k}(x^k + \lambda_k \Delta x^k, y^k + \lambda_k \Delta y^k)\|_{\infty} \leq (1 - \sigma_1 \lambda_k) \|\Phi_{\mu_k}(x^k, y^k)\|_{\infty}, \quad (6.8)$$

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

Step 3 (Update the Continuation Parameter)

Let γ_k be the maximum of the values $1, \alpha_2, \alpha_2^2, \dots$ such that

$$\|\Phi_{(1-\sigma_2\gamma_k)\mu_k}(x^{k+1}, y^{k+1})\|_\infty \leq \beta(1 - \sigma_2\gamma_k)\mu_k, \quad (6.9)$$

and set $\mu_{k+1} := (1 - \sigma_2\gamma_k)\mu_k$, $k = k + 1$. If $\|\min\{x^k, y^k\}\|_\infty \leq \epsilon$, stop; otherwise, return to Step 1.

Remarks 1. In our implementation, if (6.8) does not hold for $\lambda_k = 1, \alpha_1, \alpha_1^2, \alpha_1^3, \alpha_1^4$, we set $\lambda_k = \alpha_1^4$. Similarly, if (6.9) does not hold for $\gamma_k = 1, \alpha_2$, we set $\mu_{k+1} = \mu_k$ in Step 3.

2. In our implementation, we choose $\epsilon = 10^{-6}$, $\sigma_1 = 10^{-4}$, $\sigma_2 = 0.99999$, $\alpha_1 = 0.7$ and $\alpha_2 = 0.7$. Given x^0 , let $y^0 = Mx^0 + q$ and

$$\beta = \frac{\|\Phi(x^0, y^0, \mu_0)\|_\infty}{\mu_0}.$$

We now report the numerical results of the algorithm on the LCP test problems listed in section 6.1.

Example 6.1.1

We take $x^0 = (1, \dots, 1)^T$ and $\mu_0 = 10^{-6}$ as our starting point. The numerical results for this test problem can be found in Table 1.

Example 6.1.2

We take $x^0 = (1, \dots, 1)^T$ and $\mu_0 = 10^{-6}$ as our starting point. The numerical results for this test problem can be found in Table 2.

Example 6.1.3

Ten problems are generated for each of the dimensions $n = 50, 100, 150, 200$. The maximum, average, and minimum number of iterations needed by the algorithm are summarized in Table 3. In all runs, the starting point is chosen to be $x^0 = (0, \dots, 0)^T$ and $\mu_0 = 10^{-6}$.

Example 6.1.4

Ten problems are generated for each of the dimensions $n = 50, 100, 150, 200$. The maximum, average, and minimum number of iterations needed by the algorithm are summarized in Table 4. In all the test runs, the starting point is chosen to be $x^0 = (0, \dots, 0)^T$ and $\mu_0 = 10^{-6}$.

Example 6.1.5

Choose $n = 100$. Ten problems are generated for each of the choices $k = 90, 80, 70, 60, 50, 40, 30, 20, 10$. The maximum, average, and minimum number of iterations needed by the algorithm are summarized in Table 5. In all runs, σ_2 is chosen to be 0.7 and the starting point is chosen to be $x^0 = (0, \dots, 0)^T$. We choose $\mu_0 = 1$ for $k = 90, 80, 70, 60, 50$ and $\mu_0 = 10^3$ for $k = 40, 30, 20, 10$.

From these results, we observe that the algorithm performs well when M is positive definite matrix or a P -matrix. Unlike Algorithm 1, the algorithm is very robust even when M has high rank-deficiency.

6.4 An Infeasible Implementation Based on the Centering Direction

The algorithm implemented in this section is an infeasible version of algorithm studied in section 4.2. The iterates $\{(x^k, y^k)\}$ generated by the algorithm are not required to satisfy the affine constraints $Mx^k - y^k + q = 0$.

Algorithm 3 : An Infeasible Algorithm Using the Centering Direction

Step 0: (Initialization) Let $\mu_0 > 0$, $\beta > 0$, and $(x^0, y^0) \in \mathbb{R}^{2n}$ be given so that

$$\|Mx^0 - y^0 + q\|_\infty + \|\Phi(x^0, y^0, \mu_0)\|_\infty \leq \beta\mu_0.$$

Choose $\sigma_1, \sigma_2 \in (0, 1)$, $\alpha_1, \alpha_2 \in (0, 1)$ and $\epsilon \in (0, 1)$.

Step 1: Let $(\Delta x^k, \Delta y^k)$ solve the equation

$$\begin{aligned} M\Delta x^k - \Delta y^k &= -(Mx^k - y^k + q), \\ \nabla_x \Phi(x^k, y^k, \mu_k)\Delta x^k + \nabla_y \Phi(x^k, y^k, \mu_k)\Delta y^k &= -\Phi(x^k, y^k, \mu_k). \end{aligned} \quad (6.10)$$

Step 2: If $\|Mx^k - y^k + q\|_\infty + \|\Phi(x^k, y^k, \mu_k)\|_\infty = 0$, set $(x^{k+1}, y^{k+1}) = (x^k, y^k)$; otherwise, let $\lambda_k = \alpha_1^{s_k}$, where s_k is the smallest nonnegative integer s satisfying the following inequality

$$\begin{aligned} (1 - \lambda_k) \|Mx^k - y^k + q\|_\infty + \|\Phi_{\mu_k}(x^k + \lambda_k \Delta x^k, y^k + \lambda_k \Delta y^k)\|_\infty \leq \\ (1 - \sigma_1 \lambda_k) (\|Mx^k - y^k + q\|_\infty + \|\Phi_{\mu_k}(x^k, y^k)\|_\infty), \end{aligned} \quad (6.11)$$

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

Step 3: Let $\gamma_k = \alpha_2^{t_k}$, where t_k is the smallest nonnegative integer t satisfying the following inequality

$$\|Mx^{k+1} - y^{k+1} + q\|_\infty + \|\Phi_{(1-\sigma_2\gamma_k)\mu_k}(x^{k+1}, y^{k+1})\|_\infty \leq \beta(1 - \sigma_2\gamma_k)\mu_k, \quad (6.12)$$

Set $\mu_{k+1} = (1 - \sigma_2\gamma_k)\mu_k$ and $k := k+1$. If $\|Mx^k - y^k - q\|_\infty + \|\min\{x^k, y^k\}\|_\infty \leq \epsilon$, stop; otherwise, return to Step 1.

Remarks 1. In our implementation, if (6.10) does not hold for $\lambda_k = 1, \alpha_1, \alpha_1^2, \alpha_1^3, \alpha_1^4$, we set $\lambda_k = \alpha_1^4$. Similarly, if (6.11) does not hold for $\gamma_k = 1, \alpha_2$, we set $\mu_{k+1} = \mu_k$ in Step 3.

2. In our implementation, we choose $\epsilon = 10^{-6}$, $\sigma_1 = 10^{-4}$, $\sigma_2 = 0.9999$, $\alpha_1 = 0.7$ and $\alpha_2 = 0.7$. Given x^0 and y^0 , choose

$$\beta = \frac{\|Mx^0 - y^0 + q\|_\infty + \|\Phi(x^0, y^0, \mu_0)\|_\infty}{\mu_0}.$$

We now report the numerical results of the algorithm on the LCP test problems listed in section 6.1.

Example 6.1.1

We take $x^0 = (1, \dots, 1)^T$, $y^0 = (1, \dots, 1)^T$ and $\mu_0 = 1$ as our starting point. The numerical results for this test problem can be found in Table 1.

Example 6.1.2

We take $x^0 = (1, \dots, 1)^T$, $y^0 = (1, \dots, 1)^T$ and $\mu_0 = 10^{-6}$ as our starting point. The numerical results for this test problem can be found in Table 2.

Example 6.1.3

Ten problems are generated for each of the dimensions $n = 50, 100, 150, 200$. The maximum, average, and minimum number of iterations needed by the algorithm are summarized in Table 3. In all runs, the starting point is chosen to be $x^0 = (1, \dots, 1)^T$, $y^0 = (1, \dots, 1)^T$ and $\mu_0 = 10^{-6}$.

Example 6.1.4

Ten problems are generated for each of the dimensions $n = 50, 100, 150, 200$. The maximum, average, and minimum number of iterations needed by the algorithm are summarized in Table 4. In all the test runs, the starting point is chosen to be $x^0 = (1, \dots, 1)^T$, $y^0 = (1, \dots, 1)^T$ and $\mu_0 = 10^{-6}$.

Example 6.1.5

Choose $n = 100$. Ten problems are generated for each of the choices $k = 90, 80, 70, 60, 50, 40, 30, 20, 10$. The maximum, average, and minimum number of iterations needed by the algorithm are summarized in Table 5. In all runs, σ_2 is chosen to be 0.7 and the starting point is chosen to be $x^0 = (1, \dots, 1)^T$, $y^0 = (1, \dots, 1)^T$. We choose $\mu_0 = 1$ for $k = 90, 80, 70, 60, 50, 40, 30$ and $\mu_0 = 10^3$ for $k = 20, 10$.

From these results, we observe that the algorithm performs well when M is positive definite matrix or a P -matrix. The algorithm is the most robust among the three algorithms.

Table 6.1: The number of iterations required for the three algorithms to obtain a solution for Examples 6.1.1.

	n=8	n=16	n=32	n=64	n=128	n=256
Alg. 1	1	1	1	1	1	1
Alg. 2	1	1	1	1	1	1
Alg. 3	2	2	2	2	2	2

Table 6.2: The number of iterations required for the three algorithms to obtain a solution for Example 6.1.2.

	n=8	n=16	n=32	n=64	n=128	n=256
Alg. 1	4	5	6	7	8	8
Alg. 2	4	6	6	6	7	8
Alg. 3	2	2	2	2	2	2

6.5 Conclusions

From the numerical experiments, we make a few observations.

1. All three algorithms perform extremely well when the LCP has a positive definite matrix or a P -matrix. Algorithm 1 is the most efficient among the three algorithms for those types of problems.
2. Algorithms 2 and 3 perform well when the LCP has a high rank-deficient matrix. But Algorithm 1 may fail in this case.
3. Overall, Algorithm 3 is the most robust among the three algorithms.
4. The preliminary numerical results demonstrate that non-interior path-following methods are very promising. However, more extensive numerical experiments

Table 6.3: The number of iterations required for the three algorithms to obtain a solution for Examples 6.1.3.

		Alg. 1	Alg. 2	Alg. 3
n=50	Max.	5	12	10
	Avg.	4.2	8	5
	Min.	2	6	3
n=100	Max.	7	16	9
	Avg.	5.1	11.3	6
	Min.	4	7	4
n=150	Max.	6	15	6
	Avg.	5	12.2	4.9
	Min.	4	9	4
n=200	Max.	7	19	8
	Avg.	5.8	13.4	5.8
	Min.	5	11	5

Table 6.4: The number of iterations required for the three algorithms to obtain a solution for Examples 6.1.4.

		Alg. 1	Alg. 2	Alg. 3
n=50	Max.	8	9	12
	Avg.	6.5	7.3	8.8
	Min.	5	5	6
n=100	Max.	9	13	12
	Avg.	6.9	9.8	9.3
	Min.	5	9	7
n=150	Max.	9	16	16
	Avg.	8.1	11.2	11
	Min.	7	9	8
n=200	Max.	10	15	16
	Avg.	8.7	11.5	11.9
	Min.	7	10	10

Table 6.5: The number of iterations required for the three algorithms to obtain a solution for Examples 6.1.5.

		k=90	k=80	k=70	k=60	k=50	k=40	k=30	k=20	k=10
Alg. 1	Max.	15	18	18	21	34	42	77		
	Avg.	13.1	14.6	15.4	17.2	21.3	31.4	41.6	fail	fail
	Min.	11	12	12	14	14	17	19		
Alg. 2	Max.	12	11	14	22	27	35	43	46	61
	Avg.	10.3	10.3	11.4	14.6	21.2	29.6	32	37.4	42.6
	Min.	8	9	9	12	15	24	23	32	37
Alg. 3	Max.	19	12	12	10	19	22	39	44	46
	Avg.	9	9.4	9.6	9.1	11.2	13	14	39.1	41.2
	Min.	8	8	9	8	9	10	10	35	38

are needed in order to obtain a complete picture of the methods. Additional work is needed to understand how best to choose the parameters and the initial point for the algorithms.

BIBLIOGRAPHY

- [1] S. C. Billups. *Algorithms for Complementarity Problems and Generalized Equations*. PhD thesis, Computer Sciences Department, University of Wisconsin-Madison. 1995.
- [2] J. Burke. Second-order necessary and sufficient conditions for convex composite ndo. *Mathematical Programming*, 38:287—30, 1987.
- [3] J. Burke and S. Xu. A non-interior predictor-corrector path following algorithm for the monotone linear complementarity problem. Technical report, Preprint, Department of Mathematics, University of Washington, Seattle, WA 98195, 1997.
- [4] J. Burke and S. Xu. A non-interior predictor-corrector path-following method for LCP. Technical report, To appear in *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, Edited by M. Fukushima and L. Qi, Kluwer Academic Publishers, 1997.
- [5] J. Burke and S. Xu. The global linear convergence of a non-interior path-following algorithm for linear complementarity problem. To appear in *Mathematics of Operations Research*, 1998.
- [6] B. Chen. Personal communication, January, 1997.
- [7] B. Chen and X. Chen. A global and local super-linear continuation method for $P_0 + R_0$ and monotone NCP. Technical report, Preprint, Department of Management and Systems, Washington State University, Pullman, WA 99164-4736, 1997.

- [8] B. Chen and X. Chen. A global linear and local quadratic continuation method for variational inequalities with box constraints. Preprint. Department of Management and Systems. Washington State University, Pullman, WA 99164-4736, 1997.
- [9] B. Chen and P.T. Harker. A non-interior-point continuation method for linear complementarity problems. *SIAM J. Matrix Anal. Appl.*, 14:1168—1190, 1993.
- [10] B. Chen and P.T. Harker. A continuation method for monotone variational inequalities. *Mathematical Programming*, 69:237—253, 1995.
- [11] B. Chen and P.T. Harker. Smooth approximations to nonlinear complementarity problems. *SIAM J. Optimization*, 7:403—420, 1997.
- [12] B. Chen and N. Xiu. A global linear and local quadratic non-interior continuation method for nonlinear complementarity problems based on Chen-Mangasarian smoothing functions. Preprint. Department of Management and Systems, Washington State University, Pullman, WA 99164-4736, 1997.
- [13] C. Chen and O. L. Mangasarian. A class of smoothing functions for nonlinear and mixed complementarity problems. *Comp. Optim. Appl.*, 5:1168—1190, 1996.
- [14] C. Chen and O.L. Mangasarian. Smoothing methods for convex inequalities and linear complementarity problems. *Mathematical Programming*, 71:51—69, 1995.
- [15] X. Chen, L. Qi, and D. Sun. Global and superlinear convergence of the smoothing newton method and its application to general box constrained variational inequalities. Preprint, School of Mathematics, University of New South Wales, Sydney, Australia, 1996.

- [16] X. Chen and Y. Ye. On homotopy-smoothing methods for variational inequalities. AMR 96/39, Applied Mathematics Report, The University of New South Wales, Sydney, December, 1996.
- [17] R. W. Cottle. Nonlinear programs with positively bounded jacobians. *Journal of the Society for Industrial and Applied Mathematics*, 14:147—158, 1966.
- [18] R. W. Cottle. The principal pivoting method of quadratic programming. In G. B. Dantzig and J. A. F. Veinott, editors, *Mathematics of the Decision Sciences, Part 1*, pages 144—162. American Mathematical Society, Providence, Rhode Island, 1968.
- [19] R. W. Cottle and G. B. Danzig. Complementarity pivot theory of mathematical programming. *Linear Algebra and Its Applications*, 1:103—125, 1968.
- [20] R. W. Cottle, J. S. Pang, and R.E. Stone. *The Linear Complementarity Problem*. Academic Press, New York, New York, 1992.
- [21] G. B. Danzig and R. W. Cottle. Positive (semi-)definite programming. In J. Abadie, editor, *Nonlinear Programming*, pages 55—73. North-Holland, Amsterdam, 1967.
- [22] S.P. Dirkse and M.C. Ferris. MCPLIB: A collection of nonlinear mixed complementarity problems. *Optimization Methods and Software*, 5:319—345, 1995.
- [23] F. Facchinei and J. Soares. Testing a new class of algorithms for nonlinear complementarity problems. In F. Giannessi and A. Maugeri, editors, *Variational Inequalities and Network Equilibrium Problems*, pages 69—83. Plenum Press, New York, 1995.

- [24] F. Facchinei and J. Soares. A new merit function for nonlinear complementarity problems and a related algorithm. To appear in *SIAM J. Optimization*, 1996.
- [25] M.C. Ferris and J.S. Pang. Engineering and economic applications of complementarity problems. *SIAM Review*, 39:669–713, 1997.
- [26] S. Filipowski. On the complexity of solving feasible systems of linear inequalities specified with approximate data. *Math. Programming*, 1995.
- [27] A. Fischer. A special Newton-type optimization method. *Optimization*, 24:269–284, 1992.
- [28] A. Fischer. A Newton-type method for positive semi-definite linear complementarity problems. *J. Optim. Theory Appl.*, 86:585–608, 1995.
- [29] R. M. Freund and J. R. Vera. Some characterization and properties of ‘distance to ill-posedness’ and the condition measure of a conic linear system. Technical report, MIT, 1995.
- [30] M. Fukushima, Z.-Q. Luo, and J.-S. Pang. A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints. Preprint, Department of Mathematical Sciences, Whiting School of Engineering, The Johns Hopkins University, Baltimore, Maryland, 21218–2692, USA, 1996.
- [31] S.A. Gabriel and J.J. Moré. Smoothing of mixed complementarity problems. Technique Report MCS-P541-0995, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, 1995.

- [32] S.A. Gabriel and J.-S. Pang. An inexact NE/SQP method for solving the nonlinear complementarity problem. *Computational Optimization and Applications*, 1:67—91, 1992.
- [33] C. Geiger and C. Kanzow. On the resolution of monotone complementarity problems. *Comp. Optim. Appl.*, 5:155—173, 1996.
- [34] M. S. Gowda and M. A. Tawhid. Existence and limiting behavior of trajectories associated with the p_0 -equation. Technical report, Preprint, Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, Maryland 2125, 1997.
- [35] P. T. Harker and J. S. Pang. Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. *Mathematical Programming, Series B*, 48:161–220, 1990.
- [36] P.T. Harker and B. Xiao. Newton's method for the nonlinear complementarity problem: A B-differentiable equation approach. *Mathematical Programming*, 48:339—358, 1990.
- [37] K. Hotta and A. Yoshise. Global convergence of a class of non-interior-point algorithms using Chen-Harker-Kanzow functions for nonlinear complementarity problems. Discussion Paper Series No. 708, Institute of Policy and Planning Sciences, University of Tsukuba, Tsukuba, Ibaraki 305, Japan, 1996.
- [38] H. Jiang. Smoothed Fischer-Burmeister equation methods for the complementarity problem. Technical report, Report, Department of Mathematics, University of Melbourne, Parkville, Australia, 1997.
- [39] H. Jiang and L. Qi. A new nonsmooth equation approach to nonlinear complementarity problems. *SIAM J. Control and Optimization*, 35:178–193, 1997.

- [40] C. Kanzow. An unconstrained optimization technique for large-scale linearly constrained convex minimization problems. *Computing*, 53:101—117, 1994.
- [41] C. Kanzow. A new approach to continuation methods for complementarity problems with uniform P -functions. Preprint, Institute of Applied Mathematics, University of Hamburg, Bundesstrasse 55 D-20146 Hamburg, Germany, 1996.
- [42] C. Kanzow. Some noninterior continuation methods for linear complementarity problems. *SIAM J. Matrix Anal. Appl.*, 17:851—868, 1996.
- [43] C. Kanzow and H. Jiang. A continuation method for (strongly) monotone variational inequalities. Preprint, Institute of Applied Mathematics, University of Hamburg, Bundesstrasse 55 D-20146 Hamburg, Germany, 1996.
- [44] M. Kojima, N. Meggido, T. Noma, and A. Yoshise. *A unified approach to interior point algorithms for linear complementarity problems*. Springer-Verlag, Berlin, 1991.
- [45] M. Kojima, N. Megiddo, and T. Noma. Homotopy continuation methods for nonlinear complementarity problems. *Mathematics of Operations Research*, 16:754—774, 1991.
- [46] M. Kojima, S. Mizuno, and T. Noma. A new continuation methods for complementarity problems with uniform P -functions. *Mathematical Programming*, 43:107—113, 1989.
- [47] M. Kojima, S. Mizuno, and T. Noma. Limiting behavior of trajectories generated by a continuation methods for monotone complementarity problems. *Mathematics of Operations Research*, 15:662—675, 1990.

- [48] C. E. Lemke and J. T. Howson. Equilibrium points of bimatrix games. *SIAM Journal on Applied Mathematics*, 12:413—423, 1964.
- [49] T. De Luca, F. Facchinei, and C. Kanzow. A semismooth equation approach to the solution of nonlinear complementarity problems. *Mathematical Programming*, 75:407—439, 1996.
- [50] X.-D. Luo and P. Tseng. Conditions for a projection-type error bound for the linear complementarity problem to be global. Technical Report, Department of Mathematics, University of Washington, Seattle, U.S.A, 1993.
- [51] O.L. Mangasarian and J. Ren. New improved error bounds for the linear complementarity problem. *Mathematical Programming*, 66:241—255, 1994.
- [52] R. Mathias and J. S. Pang. Error bounds for the linear complementarity problem with a p -matrix. *Linear Algebra and its Applications*, 132:123—136, 1990.
- [53] L. McLinden. The complementarity problem for maximal monotone multifunctions. In F. Giannessi R. W. Cottle and J.L. Lions, editors, *Variational Inequalities and Complementarity Problems*, pages 251—270. Academic Press, New York, 1980.
- [54] R. Mifflin. Semismooth and semiconvex functions in constrained optimization. *SIAM Journal on Control and Optimization*, 15:957—972, 1977.
- [55] S. Mizuno. An $O(n^3L)$ algorithm using a sequence for a linear complementarity problem. *Journal of the operation research society of Japan*, 33:66—75, 1990.
- [56] S. Mizuno. A new polynomial time method for a linear complementarity problem. *Mathematical Programming*, 56:31—43, 1992.

- [57] S. Mizuno. Polynomiality of the Kojima–Megiddo–Mizuno infeasible interior point algorithm for linear programming. Technical Report 1006, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, 1993.
- [58] N. N. Karmarkar. A new polynomial time algorithm for linear programming. *Combinatorica*, 4:373—395, 1984.
- [59] J.-S. Pang. Newton’s method for B-differentiable equations. *Math. Oper. Res.*, 15:311—341, 1990.
- [60] J.-S. Pang. A B-differentiable equation-based, globally and locally quadratically convergent algorithm for nonlinear programs, complementarity, and variational inequality problems. *Math. Programming*, 51:101—131, 1991.
- [61] J.-S. Pang. Complementarity problems. In R. Horst and P. Pardalos, editors, *Handbook on Global Optimization*, pages 271—338. Kluwer Academic Publishers, Boston, 1994.
- [62] J.-S. Pang and S. A. Gabriel. NE/SQP: A robust algorithm for the non-linear complementarity problem. *Mathematical Programming*, 60:295—338, 1993.
- [63] J.-S. Pang and L. Qi. Nonsmooth equations: motivation and algorithms. *SIAM Journal on Optimization*, 3:443—465, 1993.
- [64] L. Qi and D. Sun. Globally linearly, and globally and locally superlinearly convergent versions of the Hotta–Yoshise non-interior point algorithm for nonlinear complementarity problems. Applied Mathematics Report, School of Mathematics University of New South Wales, Sydney 2052, Australia, May, 1997.
- [65] L. Qi, D. Sun, and G. Zhou. A new look at smoothing newton methods for non-linear complementarity problems and box constrained variational inequal-

- ities. Technical report, Applied Mathematics Report, AMR 97/13, School of Mathematics, University of New South Wales, Sydney 2052, Australia, 1997.
- [66] L. Qi and J. Sun. A nonsmooth version of newton's method. *Mathematical Programming*, 58:353—368, 1993.
- [67] J. Renegar. Some perturbation theory for linear programming. *Math. Programming*, 1994.
- [68] S. M. Robinson. Normal maps induced by linear transformations. *Mathematics of Operations Research*, 17:691—714, 1992.
- [69] R. T. Rockafellar. Linear-quadratic programming and optimal control. *SIAM J. Control Opt.*, 25:781—814, 1987.
- [70] R. T. Rockafellar. Computational schemes for solving large-scale problems in extended linear-quadratic programming. *Mathematical Programming*, 48:447—474, 1990.
- [71] R. T. Rockafellar and R. J-B Wets. A Lagrangian finite generation technique for solving linear-quadratic problems in stochastic programming. *Math. Prog. Studies*, 28:63—93, 1986.
- [72] S. Smale. Algorithms for solving equations. In *Proceedings of the international congress of mathematicians*, pages 172—195. Ameri. Math. Soc., Providence, 1987.
- [73] R. L. Smith. Somt interlacing properties of the schur complement of a hermitian matrix. *Linear Algebra and its Applications*, 177:137—144, 1992.
- [74] D. Sun. Private communication, 1997.

- [75] M. J. Todd and Y. Ye. Approximate farkas lemmas and stopping rules for iterative infeasible interior-point linear programming. Technical report, School of Operations Research and Industrial Engineering, Cornell University, 1994.
- [76] P. Tseng. Simplified analysis of an $o(nL)$ -iteration infeasible predictor-corrector path-following method for monotone LCP. In R.P. Agarwal, editor, *Recent Trends in Optimization Theory and Applications*, pages 423—434. World Scientific Press, Singapore, 1994.
- [77] P. Tseng. Growth behavior of a class of merit functions for the nonlinear complementarity problem. *J. Optim. Theory Appl.*, 89:17—37, 1996.
- [78] P. Tseng. Analysis of a non-interior continuation method based on chen-mangasarian smoothing functions for complementarity problems. Technical report, Department of Mathematics, University of Washington, Seattle, WA 98195, 1997.
- [79] S.J. Wright. A path-following infeasible-interior-point algorithm for linear complementarity problems. *Optimization Methods and Software*, 2:79—106, 1993.
- [80] S.J. Wright. An infeasible-interior-point algorithm for linear complementarity problems. *Mathematical Programming*, 67:29—51, 1994.
- [81] S.J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, 1997.
- [82] S.J. Wright and D. Ralph. A superlinear infeasible-interior-point algorithm for monotone complementarity problems. Technical Report MCS-P344-1292, Argonne National Laboratory, Argonne, Illinois, 1993.
- [83] S. Xu. The global linear convergence of an infeasible non-interior path-following algorithm for complementarity problems with uniform P-functions. Technical

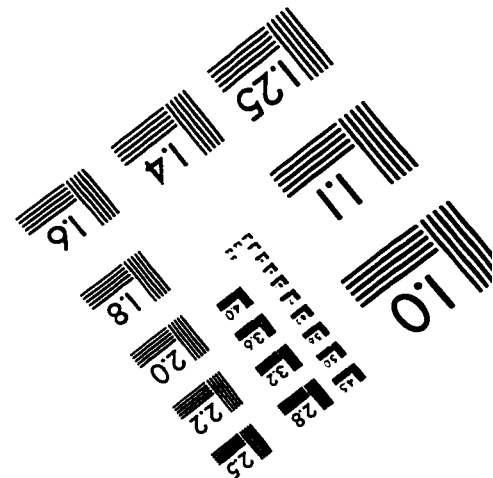
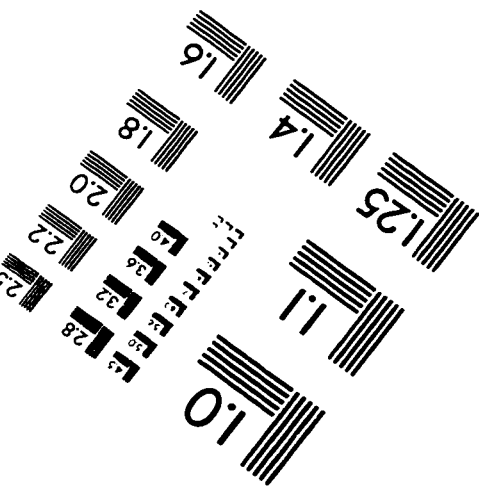
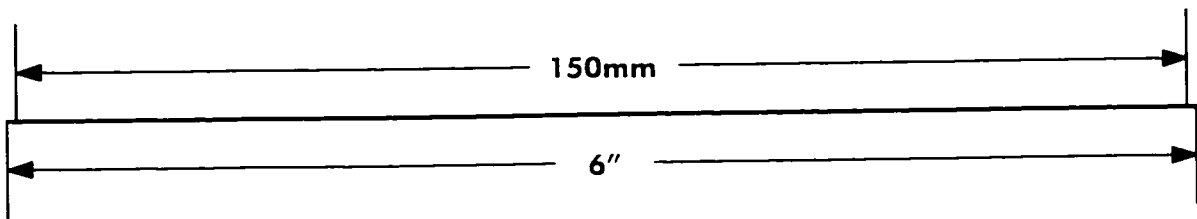
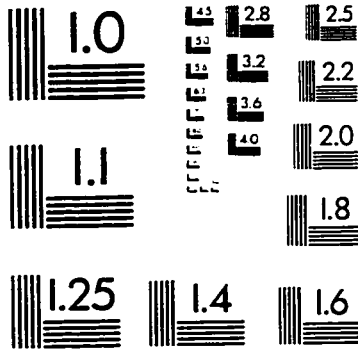
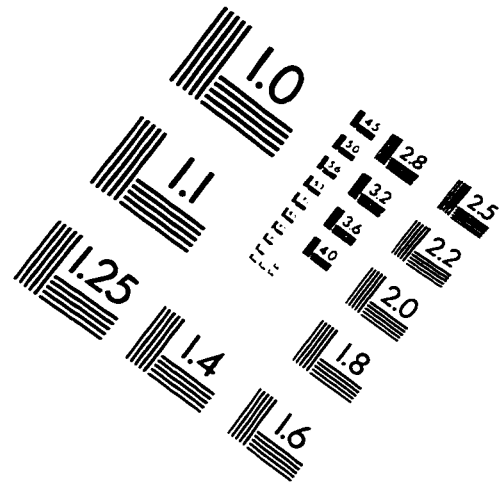
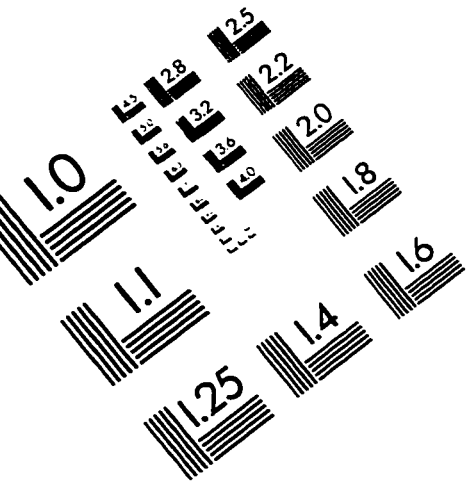
Report, Department of Mathematics, University of Washington, Seattle, WA 98195, December, 1996.

- [84] S. Xu. The global linear convergence and complexity of a non-interior path-following algorithm for monotone lcp based on Chen-Harker-Kanzow-Smale smoothing function. Technical Report, Department of Mathematics, University of Washington, Seattle, WA 98195, February, 1997.
- [85] S. Xu and J.V. Burke. A polynomial time interior-point path-following algorithm for LCP based on Chen-Harker-Kanzow smooth techniques. Technical Report, Department of Mathematics, University of Washington, Seattle, WA 98195, September, 1996.
- [86] Y. Zhang. On the convergence of a class of infeasible interior point algorithms for the horizontal linear complementarity problem. *SIAM J. Optimization*, 4:208—227, 1994.

VITA

Song Xu was born on June 3, 1963 in YingShan, HuBei, China. He received a B.S. and a M.S. in Mathematics from Beijing Normal University in 1985 and 1988 respectively. He was married to Yu Cao in 1988 and blessed with a son, William, in May, 1996. He entered the University of Washington in 1992 and obtained a Ph.D. in Mathematics in August, 1998.

IMAGE EVALUATION TEST TARGET (QA-3)



APPLIED IMAGE, Inc
 1653 East Main Street
 Rochester, NY 14609 USA
 Phone: 716/482-0300
 Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved