Interior point algorithms for general linear complementarity problems

Marianna Nagy

PhD Thesis

Supervisor: Tibor Illés
Associate Professor, PhD

Eötvös Loránd University of Sciences
Institute of Mathematics

Doctoral School: Mathematics
Director: Miklós Laczkovich

Doctoral Program: Applied Mathematics
Director: György Michaletzky

This thesis was written at the Department of Operations Research at the Eötvös Loránd University of Sciences.

Budapest
May, 2009

Contents

Acknowledgements

First of all, I would like to express my gratitude to my supervisor, Tibor Illés, who has introduced me to the theory of interior point methods, and who has provided me guidance and support. He has given me not only professional support, but personal as well. I am greatly indebted to him for the several possibilities which I owe him. I am thankful him, that he made me possible to visit him in Glasgow, hereby also helped me to finish my thesis.

I would also like to thank the support of Tamás Terlaky. Our cooperation and the experience I gained when visiting his faculty in Canada with Tibor all contributed to the results presented in this thesis.

Thanks are also due to Zsolt Csizmadia, who was a great PhD-mate, I have learnt a lot through him. I am grateful to him for a lot of useful comments which have improved the distinctness of the thesis.

I thank the research fellowship of MOL Plc. – the Hungarian Oil Company, and that I had opportunity to take part in a real life project, which gave me a lot of important experience.

Finally, I am indebted to my Mother,and to my sister, Veronika for their constant and unconditional love and endless encouragement. And also thanks for my partner, Péter Eisenberg his emotional support and that he made me laugh during at the hard periods.

I dedicate this thesis to the memory of my Father, who first opened the world of mathematics to me.

Notation

Scalars and indices are denoted by lowercase Latin letters, vectors by lowercase boldface Latin letters, matrices by capital Latin letters, and finally sets by capital calligraphic letters.

$\mathbb{R}^n_+$	the set of n dimensional positive vectors
$\mathbb{R}^n_\oplus$	the set of n dimensional nonnegative vectors
$\mathcal{I}$	the index set, namely $\mathcal{I} = \{1, 2, \ldots, n\}$
$\|\cdot\|$	the Euclidean norm
$\|\cdot\|_\infty$	the infinity norm
$\emptyset$	the empty set
$int\,\mathcal{H}$	the interior of the set $\mathcal{H}$
$\mathcal{B}_\gamma(\mathbf{w})$	the closed ball of radius γ at $\mathbf{w}$, namely $\{\mathbf{z} \,:\, \|\mathbf{z} - \mathbf{w}\| \leq \gamma\}$
$\mathcal{F}_P$	the feasible set of the LCP
$\mathcal{F}_D$	the feasible set of the dual LCP
$\mathcal{F}^+$	the set of feasible interior points of the LCP
$\mathcal{F}^*$	the solution set of the LCP
$\mathbf{x}, x_i$	vectors are thick letters, while scalars are normal letters
$\mathbf{x} > \mathbf{0}$	all coordinates of the vector $\mathbf{x}$ is positive
$\mathbf{x}\,\mathbf{s}$	the componentwise product (Hadamard product) of vectors $\mathbf{x}$ and $\mathbf{s}$
$\mathbf{x}^\alpha$	n-dimensional vector whose ith component is x_i^α
$\mathbf{x}^T\mathbf{s}$	the scalar product of two vectors
M	the coefficient matrix of the LCP, $M \in \mathbb{R}^{n\times n}$
X	the diagonal matrix from the vector $\mathbf{x}$, so $X = diag(\mathbf{x})$
I	identity matrix of size $n \times n$
$\mathbf{e}$	the vector of ones with the appropriate size
$\mathbf{0}$	the vector of zeros with the appropriate size
$\mathcal{I}_+(\mathbf{x})$	$= \{1 \leq i \leq n :\ x_i(Mx)_i > 0\}$
$\mathcal{I}_-(\mathbf{x})$	$= \{1 \leq i \ \leq n :\ x_i(Mx)_i < 0\}$
$\Delta\mathbf{x}$, $\Delta\mathbf{s}$	the Newton directions
$\big(\mathbf{x}(\theta), \mathbf{s}(\theta)\big)$	the new point after a step with step lenght θ (in interior point methods)

Abbreviations

SS	class of skew-symmetric matrices
PD	class of positive definite matrices
PSD	class of positive semidefinite matrices
CS	class of column sufficient matrices
RS	class of row sufficient matrices
LCP	linear complementarity problem
DLCP	dual linear complementarity problem
IPM	interior point method
CPP	central path problem
EP	existentially polynomial time (theorem)
LS	modified long-step path following interior point algorithm
PC	modified predictor corrector interior point algorithm

Chapter 1

Introduction

The linear complementarity problem (LCP) is still one of the intensely studied area of mathematical programming. Several books (see e.g. [11, 51, 61]) and more than a thousand articles have been published on the subject of LCPs. This is due to not only theoretical results but the important and wide range of practical applications in engineering, economics and finance.

The complementarity condition first appeared in the optimality conditions of the continuous nonlinear programming problem given by Karush in 1939 [49]. The Karush-Kuhn-Tucker optimality conditions of the linear programming and quadratic programming problems are LCP problems (see Section 1.2.1 and 1.2.2). This fact provided the motivation for studying LCP in the early time.

In 1963 the paper of Lemke and Howson [53] gave a new impulse to the research of LCPs. They showed that the Nash equilibria of a bimatrix game are the same as the solutions of an appropriate LCP (see Section 1.2.3). Furthermore, they developed a pivot method to solve the generated LCP, this is the well known Lemke algorithm. Later Cottle and Dantzig gave the unified format of the linear and quadratic programming problems and the bimatrix games as LCPs in 1968. From that time on the research of LCPs have been more and more vigorous and fruitful. A lot of theoretical question have been examined, for example the existence and uniqueness of solutions, the connectivity of the solution set and there have been several generalizations of the problem. According to the properties of LCPs several types of matrix classes have been defined. From the point of applications, not only mechanical and economical equilibrium problems have been modeled as LCPs, but for instance contact mechanics problems, network design problems, optimal control problems, optimal stopping problems, convex hulls in the plane or optimal invariant capital stock problem – mentioning only a few of those.

Consider the *linear complementarity problem* in the standard form: find vectors $\mathbf{x}, \mathbf{s} \in$

$\mathbb{R}^n$, which satisfy the constraints

$$-M\mathbf{x}+\mathbf{s}=\mathbf{q}, \quad \mathbf{x}\,\mathbf{s}=\mathbf{0}, \quad \mathbf{x},\,\mathbf{s}\geq\mathbf{0}, \tag{1.1}$$

where $M\in\mathbb{R}^{n\times n}$ and $\mathbf{q}\in\mathbb{R}^n$.

It is easy to see that the conditions can be divided into three groups. The first set of constraints are the linear equations, which describe the connection between variables. The complementarity conditions (here we take the Hadamard product of the two vectors, i.e., componentwise product, see Notations) are the second type. Finally, there are nonnegativity restrictions on the variables. According to these three types of conditions, the following three sets are introduced:
the *feasible set* [1]

$$\mathcal{F}_P:=\left\{(\mathbf{x},\mathbf{s})\in\mathbb{R}^{2n}_{\oplus}:-M\mathbf{x}+\mathbf{s}=\mathbf{q}\right\},$$

the *set of feasible interior points* [2]

$$\mathcal{F}^+:=\left\{(\mathbf{x},\mathbf{s})\in\mathbb{R}^{2n}_{+}:-M\mathbf{x}+\mathbf{s}=\mathbf{q}\right\},$$

the *set of complementarity points*, namely the *solution set of LCP*

$$\mathcal{F}^*:=\left\{(\mathbf{x},\mathbf{s})\in\mathcal{F}:\mathbf{x}\,\mathbf{s}=\mathbf{0}\right\}.$$

The feasibility problem, namely the problem of computing an element of the feasible set is relatively easy problem, as we need to solve a linear system, which is computable in polynomial time, for example with interior point algorithms. The difficulty arises from the complementarity condition, as a consequence of which, the LCP is a nonlinear problem. More precisely, the LCP (even if the coefficient matrix M is restricted to be negative definite or negative semidefinite) belongs to the class of $\mathbb{NP}$-complete problems, since the feasibility problem of linear equations with binary variables can be formulated as an LCP problem [8].

There are several algorithms to solve LCPs, but each of them requires some kind of special properties of the coefficient matrix M for finiteness, efficiency and reliability. There are two main approaches to solve LCPs, an algebraic and an analytical one. The first one is the so called pivot algorithms, which use the well known pivoting technique, and generate a point in each iteration, which satisfies the linear equations and (almost) the complementarity conditions, These algorithms try to set the nonnegativity of the solution through the iterates. If the matrix M belongs to a suitable class (for example in the case of Lemke algorithm in the

[1] Here the subscript P refers to the primal problem, because later we will deal with the dual problem of LCP, too.

[2] We will consider this set at the interior point methods.

copositive plus, while in the case of criss-cross algorithms in the sufficient matrix class), then the algorithm terminates in a finite number of steps if a proper index rule is used that avoids cycling. But pivot methods are not polynomial like in linear programming. The analytical approach leads to algorithms which do not give an exact solution in finite steps, but only converge in limit. They satisfy the linear equalities and nonnegativity of the variables, and iterate to satisfy the complementarity conditions. These methods are less sensitive to numerical errors and more efficient for large scale problems than pivot algorithms. Analytical approach for LCPs are, for instance, the different types of splitting methods (the coefficient matrix is divided into two matrices and the LCP is transformed into a fixed point problem), the damped-Newton method and different type of interior point methods. This thesis deals with the last group of the foregoing algorithms, namely with the interior point methods.

We have already mentioned that there are several applications which lead to an LCP, therefore there is a real demand for an efficient algorithm solving LCPs. In general applications, we do not know whether the matrix belongs to one of the above mentioned suitable matrix classes or not, and the checking of those properties is also an $\mathbb{NP}$-complete problem (see Section 2). Furthermore, in most cases the matrix of a general application does not inherit such special properties. Due to these facts, generally we can not expect a polynomial time algorithm for solving LCP problems. Therefore, our aim is not to solve the LCP in all cases, but to construct an efficient, polynomial time algorithm which provides some kind of information about the given LCP problem. If everything turns out well, we solve the problem or the dual problem proving the unsolvability of the original problem. If we are less lucky, we get a polynomial sized certificate that the matrix of the problem does not hold a given special property (see Section 6).

1.1 Structure of the thesis

The rest of this introductory chapter presents some applications, that can be formulated as an LCP. The first two examples are theoretical: the linear and the quadratic programming problem. As we have already mentioned, in the 1940s this fact induced the research on the LCP. Furthermore, several efficient methods for quadratic programming problems are based on the LCP formulation. After that, we touch three other, in our days still important applications. The bimatrix games are well known from game theory and were first converted to an LCP by Lemke and Howson [53]. This connection between bimatrix games and LCPs has theoretical importance, too, as it gives a constructive tool to determine an equilibrium point, and has initiated a new approach of equilibrium theory. The fourth instance is the optimal stopping problem of Markov chain, which is a classical problem in stochastic control. It has

a wide range of applications and it is a basic element of lots of general applied probability models. The last problem is an economy equilibrium problem, more precisely a special class of the famous Arrow-Debreu exchange market equilibrium problem. Walras entered the history of economics with his general equilibrium theory. " Walras first formulated the state of the economic system at any point of time as the solution of a system of simultaneous equations representing the demand for goods by consumers, the supply of goods by producers, and the equilibrium condition that supply equal demand on every market." wrote Arrow and Debreu in [2]. The idea of general equilibrium can not be evaded – not because the markets are always thought to be in balance, but it gives us an excellent reference point. It is not by chance, that Kenneth Arrow, the remarkable talent of 20th century economics, the first precise mathematical establisher of this theory got one of the first Nobel prizes of this science in 1972 almost 100 years after Walras.

" Together with Gerhard Debreu, he produced in 1954 a very abstract model, based on mathematical set theory, which opened up fresh possibilities of making interesting analyses. For example, he and Debreu were the first to be able to demonstrate, in a mathematically stringent manner, the conditions which must be fulfilled if a neoclassical general equilibrium system is to have a unique and economically meaningful solution. By introducing a new technique for dealing with the theory of decision-making under conditions of uncertainty and risk, and by incorporating this theory in the general equilibrium theory, Arrow has also achieved results of great theoretical and practical interest." [4]

Debreu got the Nobel Prize in 1983 "for having incorporated new analytical methods into economic theory and for his rigorous reformulation of the theory of general equilibrium" [57].

Here we briefly describe the Arrow-Debreu exchange market equilibrium with Leontief utility functions and the related LCP formulation, which was introduced by Ye in [85]. Ye and his colleagues studied the Arrow-Debreu economy equilibrium with different utility functions and they presented not only algorithmic complexity results, but in some cases an algorithm (based on interior point method) for computing the equilibrium prices, as well [7, 9, 17, 85, 86, 88]. For the mentioned problem they provide a homotopy based interior point path following algorithm and a fully polynomial-time approximation scheme, and reported computational results [88]. This gave us the idea to test our modified interior point methods on this problem, that is, on the LCP formulation of it. Unfortunately, in the case of LCPs there are no test sets of real-life problems like NETLIB for linear programming yet. Therefore, the efficiency of algorithms for LCPs can not be really compared on a universal basis.

Let us note, that this list of LCP applications is far from the full spectrum (for example the Karush-Kuhn-Tucker system of several nonlinear problems also leads to an LCP). We

would only like to present some instances to demonstrate how important and necessary it is for applications to develop an efficient algorithm to solve LCP problems.

In Chapter 2 we deal with some matrix classes related to the LCP, which are important for our purposes, that is, which have some kind of connection to interior point methods. These are the $\mathcal{P}$, $\mathcal{P}_0$, $\mathcal{P}_*(\kappa)$, $\mathcal{P}_*$ and sufficient matrix classes. An LCP has exactly one solution for all right hand side vector $\mathbf{q}$, if and only if the coefficient matrix M is a $\mathcal{P}$-matrix. The $\mathcal{P}_0$ matrix class is a generalization of the $\mathcal{P}$ class. Kojima et al. [51] introduced the matrix class $\mathcal{P}_*(\kappa)$, which is the widest class where interior point methods are polynomial, however the complexity depends on the parameter κ, too (it is also a polynomial dependence). The union of the sets $\mathcal{P}_*(\kappa)$ for all nonnegative κ is the $\mathcal{P}_*$ class. Väliaho [80] proved that it is the same as the sufficient matrix class defined by Cottle et al. [12], which is the widest matrix class where the finiteness of the criss-cross algorithm with minimal index rule can be proved [18]. Meanwhile, Csizmadia and Illés claimed that the finiteness of the criss-cross algorithm holds also with other, more general index rules [15, 16]. The sufficient matrix class, i.e., the set of $\mathcal{P}_*$-matrices, includes the $\mathcal{P}$ class and it is a subset of the $\mathcal{P}_0$ matrix class.

The basis of this chapter is manuscript [39], which is expanded with some interesting observations. After the introduction of the major properties of the mentioned matrix classes, we turn our attention to the handicap of a matrix, that is, the smallest value of κ with which the matrix is a $\mathcal{P}_*(\kappa)$. This parameter has cardinal importance at interior point methods, as it appears for example in the dependence of complexity on κ. Furthermore, a finite handicap of a matrix means that the matrix is a $\mathcal{P}_*$-matrix, so the LCP problem can be solved with an interior point algorithm. Unfortunately, there is no known polynomial time algorithm to determine the handicap of a matrix. The set of sufficient matrices is a nonconvex, neither closed, nor open set. Besides the lack of good characterization of the sufficient matrix class it has nice properties as well, for example, it is a cone and it is invariant under special row and column scaling, and under principal pivoting. We visit the question of complexity regarding matrix classes,too. The decision problem related to copositive, $\mathcal{P}$ and $\mathcal{P}_*$ matrix classes, namely whether the given matrix belongs to the matrix class or not, is co-$\mathbb{NP}$-complete. We close this chapter with some important results about the solution set of the LCP and the matrix classes.

Chapter 3 summarizes the basic theory of interior point methods. In the first part we collect the main results from the manuscript of Illés et al. [38], which establish the interior point theory of the LCP for $\mathcal{P}_*(\kappa)$ matrices without using the implicit function theorem. They showed the existence and uniqueness of the central path, the guideline of interior point

methods, and proved that it converges to a maximally complementary solution of the problem. After introducing the Newton system, the linear relaxation of the central path problem, we give a short list of centrality measures and the most frequently used neighbourhoods of the central path. We close the brief review of interior point method theory with the general sketch of algorithms and a few estimations of Newton directions which are useful at the complexity analysis of interior point methods. The initial interior point is always a crucial question of interior point methods. There are two methodologies to solve it. One is to apply an infeasible method and the other is to use the embedding technique. We discuss the second one. Kojima et al. [51] showed that with an appropriate embedding the special property of the coefficient matrix can be preserved. The results considered in this chapter will be used in Chapter 5 and 6.

The first three chapters mainly summarize known results (only the second one contains some of our own observations). These provide a basis for the second part of the dissertation, where we present our results. In Chapter 4 the dual of the LCP is introduced which was developed in a general form for oriented matroids by Fukuda and Terlaky in [26]. Csizmadia and Illés examined it for LCPs, related to the criss-cross algorithm in [16], we will use this form of the dual. Probably because of the general form of the dual by Fukuda et al., the result of this chapter escaped the attention of most researchers. We show that the dual LCP can be solved in polynomial time if the matrix is sufficient, and give an EP type theorem based on this complexity result. The achievements of this chapter have been published in paper [43].

Chapter 5 deals with one of the most remarkable interior point methods, the Mizuno–Todd–Ye predictor-corrector algorithm. The basis of this chapter is the paper [65] by Potra. He examined this algorithm with a wide neighbourhood of the central path for skew-symmetric and positive semidefinite LCPs. We generalize it for LCPs with $\mathcal{P}_*(\kappa)$-matrices and show that the algorithm preserves its nice property, that is, if the two used neighbourhoods are well chosen, then after each predictor step we can return to the smaller neighbourhood of the central path with only one corrector step. However, the neighbourhoods depend on the parameter κ, which means smaller and smaller suitable neighbourhoods for larger κ. Therefore, this algorithm is not well suited for solving practical problems.

Our main results are stated in Chapter 6. Here we take a further step in the generalization of interior point methods. As we have already pointed out, usually we do not know anything about the matrix of a real life problem, moreover in most cases it is not a $\mathcal{P}_*(\kappa)$-matrix.

Therefore, we construct modified interior point methods (a long step path-following, an affine scaling and a predictor-corrector), which can handle any LCPs. These algorithms either solve the problem or its dual (in the latter case proving that the problem has no solution), or give a polynomial certificate that the matrix is not a $\mathcal{P}_*(\tilde{\kappa})$-matrix with an a priori chosen but an arbitrary $\tilde{\kappa}$. These results have a theoretical side, too. They give a constructive proof of an EP type theorem. A similar result was proved by Csizmadia and Illés [16] based on the criss-cross algorithm, but our modified interior point methods are still polynomial. On the other hand, the criss-cross algorithm solves LCPs for sufficient matrices, namely, with arbitrary large κ, which does not need to be fixed a priori. This chapter is based on papers [42, 44].

There are some further research directions. For example, although these algorithms do not solve an LCP problem in all cases, our preliminary computational experiences show that a successful run depends on the initial point. Thus, a randomized, multistart algorithm can help. Another question is how we can use the information of a run which does not give a solution of an LCP or its dual.

We close the thesis with Further questions and Summary.

1.2 Applications

In this section we review five problems which can be formulated as LCP problems, the linear programming, the quadratic programming with linear constraints, the bimatrix game, the optimal stopping of Markov chain and a special class of the Arrow-Debreu exchange market equilibrium problem. In each case the coefficient matrix of the LCP problem has a special structure, however, generally only the first problem can be solved in polynomial time. Unfortunately the LCP algorithms can not take advantage of these types of structure.

1.2.1 Linear programming

Consider the *linear programming primal* (P) and *dual* (D) problem in the standard form: find a vector $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$ such that

$$\left.\begin{array}{rcl} \min \mathbf{c}^T\mathbf{x} & & \\ A\mathbf{x} & \geq & \mathbf{b} \\ \mathbf{x} & \geq & \mathbf{0} \end{array}\right\} \quad (P) \qquad\qquad \left.\begin{array}{rcl} \max \mathbf{b}^T\mathbf{y} & & \\ A^T\mathbf{y} & \leq & \mathbf{c} \\ \mathbf{y} & \geq & \mathbf{0} \end{array}\right\} \quad (D)$$

where $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$ and $\mathbf{c} \in \mathbb{R}^n$.

Then the well known optimality conditions are the following:

$$\begin{array}{rrrcll} -A\,\mathbf{x} + & & \mathbf{z} & = -\mathbf{b}, & \mathbf{x} \geq \mathbf{0}, & \mathbf{z} \geq \mathbf{0}, \\ A^T\mathbf{y} + & & \mathbf{s} & = \mathbf{c}, & \mathbf{y} \geq \mathbf{0}, & \mathbf{s} \geq \mathbf{0}, \\ & \mathbf{x}^T\mathbf{s} + & \mathbf{y}^T\mathbf{z} & = 0. & & \end{array} \tag{1.2}$$

The last condition is equivalent to $\mathbf{x}\,\mathbf{s} = \mathbf{0}$ and $\mathbf{y}\,\mathbf{z} = \mathbf{0}$ using the nonnegativity of variables $\mathbf{x}$, $\mathbf{s}$, $\mathbf{y}$, $\mathbf{z}$. Therefore, the optimality conditions are in the LCP form (1.1) with

$$M = \begin{bmatrix} O & A \\ -A^T & O \end{bmatrix} \in \mathbb{R}^{(m+n)\times(m+n)}, \qquad \mathbf{q} = \begin{pmatrix} -\mathbf{b} \\ \mathbf{c} \end{pmatrix} \in \mathbb{R}^{m+n}.$$

One can see, that here the matrix M is skew-symmetric , that is, $M^T = -M$. The LCP problem for skew-symmetric matrices can be solved in polynomial time, for example, with interior point methods.

In the theory of interior point methods, there is another transformation of the (P)-(D) pair to the LCP form. This is also based on the optimality conditions, but in a slightly different way as in (1.2), because the last condition is replaced by $\mathbf{b}^T\mathbf{y} - \mathbf{c}^T\mathbf{x} \geq \mathbf{0}$ according to the Weak duality theorem. After homogenization, we get the Goldman-Tucker system, which is an LCP problem (1.1) with

$$M = \begin{bmatrix} O & A & -\mathbf{b} \\ -A^T & O & \mathbf{c} \\ \mathbf{b}^T & -\mathbf{c}^T & 0 \end{bmatrix} \in \mathbb{R}^{(m+n+1)\times(m+n+1)}, \qquad \mathbf{q} = \mathbf{0} \in \mathbb{R}^{m+n+1}.$$

Note that this matrix is also skew-symmetric. It is easy to check that the latter LCP always has a solution, the all zero vector. The Goldman-Tucker theorem [29] describes the connection between the solution of the primal and dual linear programming problem and the LCP (for details see e.g. [70]).

1.2.2 Quadratic programming

Consider the *quadratic programming problem* (QP): find a vector $\mathbf{x} \in \mathbb{R}^n$ such that

$$\left.\begin{array}{rrcl} \min\ \dfrac{1}{2}\,\mathbf{x}^T Q\,\mathbf{x} & + & \mathbf{c}^T\mathbf{x} & \\ A\mathbf{x} & \geq & \mathbf{b} & \\ \mathbf{x} & \geq & \mathbf{0} & \end{array}\right\}$$

where $Q \in \mathbb{R}^{n\times n}$ is a symmetric matrix, $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$.

The optimality conditions according to the Karush-Kuhn-Tucker (KKT) conditions:

$$\begin{aligned} \mathbf{u} &= \mathbf{c} + Q\mathbf{x} - A^T\mathbf{y} \geq \mathbf{0}, & \mathbf{x} \geq \mathbf{0}, & & \mathbf{x}^T\mathbf{u} = 0, \\ \mathbf{v} &= -\mathbf{b} + A\mathbf{x} \geq \mathbf{0}, & \mathbf{y} \geq \mathbf{0}, & & \mathbf{y}^T\mathbf{v} = 0. \end{aligned}$$

It can be reformulated as an LCP problem (1.1), where

$$M = \begin{bmatrix} Q & -A^T \\ A & O \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)} \qquad \mathbf{q} = \begin{pmatrix} \mathbf{c} \\ -\mathbf{b} \end{pmatrix} \in \mathbb{R}^{n+m}$$

It is obvious that the matrix M has a special structure again. This matrix is bisymmetric, as it is the sum of a symmetric and a skew-symmetric matrix.

It is a well known result, that if the matrix Q is a positive semidefinite matrix, than $\mathbf{x}$ is a solution of QP if and only if there exists a vector $\mathbf{y}$ such that $(\mathbf{x}, \mathbf{y})$ is a KKT point (see e.g. [3]). The QP problem is an $\mathbb{NP}$-complete problem [28, 72, 82], but if the matrix Q is a positive semidefinite matrix, there are polynomial algorithms, for example interior point methods.

1.2.3 Bimatrix games

There are two players. Each of them has a finite set of strategies, $\mathcal{I} = \{1, \ldots, n\}$ and $\mathcal{J} = \{1, \ldots, m\}$. The matrix $A \in \mathbb{R}^{n\times m}$ is the payoff matrix of the first player, and $B \in \mathbb{R}^{n\times m}$ is the payoff matrix of the second player – it means that if they play the $i \in \mathcal{I}$ and $j \in \mathcal{J}$ strategies respectively, then the first player pays a_{ij} amount of money and the second player pays b_{ij}. We can define mixed strategies $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$, which mean the probability distribution of selecting the strategies, therefore $\mathbf{x}, \mathbf{y} \geq \mathbf{0}$ and $\mathbf{e}^T\mathbf{x} = 1$, $\mathbf{e}^T\mathbf{y} = 1$. In this case, the expected costs of the game are $\mathbf{x}^T A\mathbf{y}$ and $\mathbf{x}^T B\mathbf{y}$. Each player would like to minimize his expense.

A pair of mixed strategies $(\mathbf{x}^*, \mathbf{y}^*)$ is a *Nash equilibrium* (i.e., optimal to both players unless they cooperate) if

$$(\mathbf{x}^*)^T A\,\mathbf{y}^* \leq \mathbf{x}^T A\,\mathbf{y}^* \qquad \text{for all } \mathbf{x} \geq \mathbf{0} \quad \text{and} \quad \mathbf{e}^T\mathbf{x} = 1, \tag{1.3}$$

$$(\mathbf{x}^*)^T B\,\mathbf{y}^* \leq (\mathbf{x}^*)^T B\,\mathbf{y} \qquad \text{for all } \mathbf{y} \geq \mathbf{0} \quad \text{and} \quad \mathbf{e}^T\mathbf{y} = 1. \tag{1.4}$$

The condition (1.3) is equivalent to

$$(\mathbf{x}^*)^T A\,\mathbf{y}^* \leq \mathbf{a}_i\,\mathbf{y}^* \qquad \text{for all } \; i = 1 \ldots n,$$

where $\mathbf{a}_i$ is the ith row of the matrix A. This is in the vector form

$$\Big((\mathbf{x}^*)^T A\,\mathbf{y}^*\Big)\,\mathbf{e} \leq A\,\mathbf{y}^*. \tag{1.5}$$

Similarly, the condition (1.4) is equivalent to

$$\left((\mathbf{x}^*)^T B\,\mathbf{y}^*\right)\mathbf{e} \le B^T\mathbf{x}^*. \tag{1.6}$$

Without loss of generality we can assume, that A and B are positive matrices (each entry is positive), because adding a positive number to each entry of the payoff matrix does not change the Nash equilibrium points. Therefore, we can assume that $(\mathbf{x}^*)^T A\,\mathbf{y}^* > 0$ and $(\mathbf{x}^*)^T B\,\mathbf{y}^* > 0$. Let $\bar{\mathbf{x}} = \dfrac{\mathbf{x}^*}{(\mathbf{x}^*)^T B\,\mathbf{y}^*}$ and $\bar{\mathbf{y}} = \dfrac{\mathbf{y}^*}{(\mathbf{x}^*)^T A\,\mathbf{y}^*}$. With these notation the inequalities (1.5) and (1.6) are the following:

$$\mathbf{e} \le A\,\bar{\mathbf{y}} \qquad \text{and} \qquad \mathbf{e} \le B^T\bar{\mathbf{x}}.$$

Furthermore,

$$\mathbf{e}^T\mathbf{x}^* = 1 = (\mathbf{x}^*)^T A\,\frac{\mathbf{y}^*}{(\mathbf{x}^*)^T A\,\mathbf{y}^*} = (\mathbf{x}^*)^T A\,\bar{\mathbf{y}},$$

which becomes the following using the new notation:

$$\mathbf{e}^T\bar{\mathbf{x}} = \bar{\mathbf{x}}^T A\,\bar{\mathbf{y}}.$$

Similarly, for the mixed strategy of the other player we get

$$\mathbf{e}^T\bar{\mathbf{y}} = \bar{\mathbf{x}}^T B\,\bar{\mathbf{y}}.$$

Summing up the above conditions, we get the according LCP:

$$\left.\begin{array}{llll} \mathbf{u} = -\mathbf{e} + A\,\bar{\mathbf{y}} \ge \mathbf{0}, & \bar{\mathbf{x}} \ge \mathbf{0}, & \bar{\mathbf{x}}^T\mathbf{u} = 0 \\ \mathbf{v} = -\mathbf{e} + B^T\bar{\mathbf{x}} \ge \mathbf{0}, & \bar{\mathbf{y}} \ge \mathbf{0}, & \bar{\mathbf{y}}^T\mathbf{v} = 0 \end{array}\right\}, \tag{1.7}$$

namely

$$M = \begin{bmatrix} O & A \\ B^T & O \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}, \qquad \mathbf{q} = -\mathbf{e} \in \mathbb{R}^{n+m}.$$

Note that the matrix M is a nonnegative matrix with zero diagonal elements.

Based on the above, the Nash equilibrium points and the solutions of the LCP (1.7) can be assigned in the following sense

- Let $(\mathbf{x}^*, \mathbf{y}^*)$ be a Nash equilibrium, then
  $\bar{\mathbf{x}} = \dfrac{\mathbf{x}^*}{(\mathbf{x}^*)^T B\,\mathbf{y}^*}$ and $\bar{\mathbf{y}} = \dfrac{\mathbf{y}^*}{(\mathbf{x}^*)^T A\,\mathbf{y}^*}$ is a solution of problem (1.7).
- Let $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ be a solution of the system (1.7), then
  $\mathbf{x}^* = \dfrac{\bar{\mathbf{x}}}{\mathbf{e}^T\bar{\mathbf{x}}}$ and $\mathbf{y}^* = \dfrac{\bar{\mathbf{y}}}{\mathbf{e}^T\bar{\mathbf{y}}}$ is a Nash equilibrium.

Let us remark that if the game is a zero-sum game, namely the sum of the payoffs is always are zero ($A = -B$), then the problem can be formulated as a linear programming problem.

1.2.4 Optimal stopping of Markov chain

Let us consider the *Markov chain* problem: we have a finite state space $\mathcal{E} = \{1, 2, \ldots, n\}$ and a transition probability matrix P. In each step there are two opportunities: either we stop, in this case the payoff is given by r_i ($i \in \mathcal{E}$), or we continue the process to the next state according to the matrix P. The aim is to maximize the expected payoff.

Let v_i be the stationary optimal expected payoff if the process starts at the initial state $i \in \mathcal{E}$. Then

$$\mathbf{v} = \max(P\mathbf{v}, \mathbf{r}).$$

Considering the properties of the maximum, it is equivalent to:

$$\mathbf{v} \geq P\mathbf{v}, \quad \mathbf{v} \geq \mathbf{r}, \quad (\mathbf{v} - \mathbf{r})^T(\mathbf{v} - P\mathbf{v}) = 0.$$

This is in LCP form:

$$\mathbf{u} = \mathbf{v} - \mathbf{r}, \quad M = I - P, \quad \mathbf{q} = (I - P)\mathbf{r}.$$

In this case the matrix M has a nonnegative diagonal, but all off-diagonal elements are nonpositive.

1.2.5 Economy equilibrium problem

In this section we consider the Arrow-Debreu exchange market equilibrium problem which is a fundamental model used in general equilibrium theory of economics. First Léon Walras formulated a model in 1874 [84], then Arrow and Debreu built up the precise mathematical background, and gave an axiomatic description of the economy equilibrium in 1954 [2]. In this problem, there are m traders and n goods on the market. Each trader i has an initial endowment of commodities $\mathbf{w}^i = (w_{i1}, \ldots, w_{in}) \in \mathbb{R}^n_\oplus$. They sell it at a given price $\mathbf{p} \in \mathbb{R}^n_\oplus$ and then use the income to buy a bundle of goods $\mathbf{x}^i = (x_{i1}, \ldots, x_{in}) \in \mathbb{R}^n_\oplus$. Each trader i has a utility function u_i, which describes his preferences for the different bundle of commodities, and he maximizes his individual utility function subject to the budget constraint $\mathbf{p}^T\mathbf{x}^i \leq \mathbf{p}^T\mathbf{w}^i$. Let us denote by $\mathbf{x}^i(\mathbf{p})$ a maximizer vector, which is the demand of trader i at price $\mathbf{p}$.

The vector of prices $\mathbf{p}$ is an equilibrium for the exchange economy, if there is a bundle of goods $\mathbf{x}^i(\mathbf{p})$ (so a maximizer of the utility function u_i subject to the budget constraint) for all traders i, such that $\sum_{i=1}^m x_{ij}(\mathbf{p}) \leq \sum_{i=1}^m w_{ij}$ for all goods. In other words, the question is whether there are such prices of goods, where the demand $\sum_i x_{ij}(\mathbf{p})$ does not exceed the supply $\sum_i w_{ij}$ for all good j, namely whether the price could be set for goods in such a way that each trader can maximize his utility function individually. This was the question of Walras.

From a market point of view, the equilibrium problem can be composed as an aggregated model by the following notations. For good j at price $\mathbf{p}$ denotes $d_j(\mathbf{p}) = \sum_{i=1}^{m} x_{ij}(\mathbf{p})$ the market demand and $z_j(\mathbf{p}) = d_j(\mathbf{p}) - \sum_{i=1}^{m} w_{ij}$ the market excess demand. Then vector $\mathbf{d}(\mathbf{p}) = (d_1(\mathbf{p}), \ldots, d_n(\mathbf{p}))$ is the market demand and $\mathbf{z}(\mathbf{p}) = (z_1(\mathbf{p}), \ldots, z_n(\mathbf{p}))$ is the market excess demand.

The market satisfies the Walras' Law if for any price $\mathbf{p}$ we have $\mathbf{p}^T\mathbf{z}(\mathbf{p}) = 0$. We say that $\mathbf{z}(\mathbf{p})$ is well defined, if each trader has an optimal bundle of goods, namely if a vector $\mathbf{x}(\mathbf{p})$ exists. In this way, a vector of prices $\mathbf{p} \in \mathbb{R}^n_{\oplus}$ is an equilibrium if $\mathbf{z}(\mathbf{p})$ is well defined and $\mathbf{z}(\mathbf{p}) \leq \mathbf{0}$.

There are some special utility functions in literature, for example linear, Leontief, Cobb-Dougles and CES functions (for a good summary wee e.g. [30]). Later on we will only deal with the Leontief utility function.

In 1954 Arrow and Debreu gave an answer for Walras' question [2]. They proved that under mild conditions and if the utility functions are concave, such equilibrium exists. However, they did not provide any algorithm to compute an equilibrium of a market. Fisher was the first to present an algorithm to determine equilibrium prices, however his model was a special type of Walras' model. There players are divided into two sets: producers and consumers. Producers sell their goods for money and consumers have money to buy goods and maximize their utility functions. We get the Walras' model if money is also considered as a commodity. On the other hand in the Arrow-Debreu model each trader is both a producer and a consumer.

There are a lot of special cases when some algorithms and complexity results are presented, but there is no known method to compute an equilibrium of a market in the general case.

From now on we will consider a special class of the Arrow-Debreu model. Let us assume that each trader enters the market with exactly one good (so $n = m$) and has exactly one unit of it. Therefore, if the price vector $\mathbf{p}$ is given, then the budget of trader i is p_i, thus the optimal strategy of the trader i is determined by the following optimization problem:

$$\left.\begin{aligned} \max u_i(\mathbf{x}^i) \\ \mathbf{p}^T\mathbf{x}^i \leq p_i, \\ \mathbf{x}^i \geq \mathbf{0}. \end{aligned}\right\} \tag{1.8}$$

Let $\mathbf{x}^i(\mathbf{p})$ be an optimal solution of problem (1.8). Then the vector $\mathbf{p}$ is an Arrow-Debreu price equilibrium if for each trader i there is an $\mathbf{x}^i(\mathbf{p})$ optimal solution of the system (1.8) such that

$$\sum_{i=1}^{n} \mathbf{x}^i(\mathbf{p}) = \mathbf{e},$$

namely the demand of every good is exactly one unit, so the demand equals the supply.

In the remainder of this section we will be concerned with the Leontief exchange economy equilibrium problem, when the utility functions are Leontief functions defined in the following way:

$$u_i(\mathbf{x}^i) = \min_j \left\{ \frac{x_{ij}}{a_{ij}} : \ a_{ij} > 0 \right\},$$

where $A = (a_{ij}) \in \mathbb{R}^{n\times n}_{\oplus}$ is the Leontief coefficient matrix.

Furthermore, we assume that every trader likes at least one commodity, so the matrix A has no all-zero row.

Eisenberg and Gale [20, 21, 27] gave a convex programming formulation of the Fisher equilibrium problem with Leontief utility functions and proved that an equilibrium price vector is an optimal Lagrangian multiplier of this convex programming problem:

$$\left.\begin{array}{rl} \max \sum\limits_{i=1}^{n} w_i \ \log u_i & \\ A^T\mathbf{u} & \leq \mathbf{e} \\ \mathbf{u} & \geq \mathbf{0}, \end{array}\right\} \tag{1.9}$$

where in the Arrow-Debreu model the initial budget w_i of trader i is not given and will be p_i, i.e., the price of his good. Furthermore, u_i represents the utility value of trader i and A is the Leontief matrix.

Therefore, we search such weights w_i that an optimal Lagrangian multiplier vector $\mathbf{p}$ of (1.9) equals $\mathbf{w}$. This means that $\mathbf{p}$ is a solution of the following system[3]:

$$\left.\begin{array}{rl} UA\mathbf{p} & = \mathbf{p} \\ P(\mathbf{e} - A^T\mathbf{u}) & = \mathbf{0} \\ A^T\mathbf{u} & \leq \mathbf{e} \\ \mathbf{u}, \mathbf{p} & \geq 0 \\ \mathbf{p} & \neq \mathbf{0} \end{array}\right\} \tag{1.10}$$

where U and P are diagonal matrices whose diagonal is $\mathbf{u}$ and $\mathbf{p}$, respectively. Under the previous assumption that the matrix A has no all-zero row, this system always has a solution. Remember – as we have already mentioned – that every Arrow-Debreu equilibrium price vector satisfies the system (1.10). However, the implication can not be reversed, a solution of the system (1.10) may not be an equilibrium of the Arrow-Debreu model.

We close this section with the LCP problem, which is equivalent to the pairing Arrow-Debreu model with the Leontief utility. This LCP was presented by Ye in [85] based on his following result:

[3]This is the KKT system of (1.9) using the equality $\mathbf{w} = \mathbf{p}$. Since $\mathbf{w}$ is the vector of initial budgets, it is a meaningful condition, that $\mathbf{p} = \mathbf{w} \neq \mathbf{0}$.

Theorem 1.1 *Let $\mathcal{B} \subset \{1, 2, \ldots, n\}$, $\mathcal{N} = \{1, 2, \ldots, n\} \setminus \mathcal{B}$, A_{BB} be irreducible*[4]*, and $\mathbf{u}_B$ satisfy the linear system*

$$A_{BB}^T \mathbf{u}_B = \mathbf{e}, \quad A_{BN}^T \mathbf{u}_B \leq \mathbf{e}, \quad \textit{and} \quad \mathbf{u}_B > \mathbf{0}. \tag{1.11}$$

Then the (right) Perron-Frobenius eigenvector[5] *$\mathbf{p}_B$ of $U_B A_{BB}$ together with $\mathbf{p}_N = \mathbf{0}$ will be a solution of the system (1.10). And the converse is also true. Moreover, there is always a rational solution for every such $\mathcal{B}$, that is, the entries of price vector are rational numbers, if the entries of A are rational. Furthermore, the size (bit-length) of the solution is bounded by the size (bit-length) of A.*

This theorem is a good characterization of the solutions of the system (1.10). It provides a combinatorial algorithm to solve the problem (1.10), however it is not suitable in practice, because we possibly need to examine 2^n subset $\mathcal{B}$ of $\mathcal{I}$ to see whether the matrix $A_{\mathcal{BB}}$ is irreducible and the system (1.11) has a solution or not. Even so Theorem 1.1 establishes an algorithm to compute the Arrow-Debreu equilibrium if A is a positive matrix. Search a nontrivial solution $\mathbf{u} \neq \mathbf{0}$ of the following LCP problem

$$\left.\begin{aligned} A^T\mathbf{u} + \mathbf{v} &= \mathbf{e} \\ \mathbf{uv} &= \mathbf{0} \\ \mathbf{u}, \mathbf{v} &\geq \mathbf{0} \end{aligned}\right\} \tag{1.12}$$

Then take the support of the nontrivial solution $\mathbf{u}$ as $\mathcal{B}$, that is, $\mathcal{B} = \{i : \ u_i > 0\}$. According to Theorem 1.1, the (right) Perron-Frobenius eigenvector of $U_B A_{BB}$ is an Arrow-Debreu equilibrium.

We will return to this problem at the end of the dissertation. In Chapter 6 we will review the computation experiences of our modified interior point algorithm on the LCP (1.12).

[4] A matrix $A \in \mathbb{R}^{n \times n}$ is irreducible if and only if for any partition $\mathcal{I} = \mathcal{J} \cup \mathcal{K}$ there exists $j \in \mathcal{J}$ and $k \in \mathcal{K}$ such that $a_{jk} \neq 0$.

[5] Let $A \in \mathbb{R}^{n \times n}$ be an irreducible matrix with positive entries. Then there is a positive real eigenvalue λ of A such that $\rho(A) = \lambda$, where $\rho(A)$ is the spectral radius of the matrix. Furthermore, λ is simple, i.e., it is a simple root of the characteristic polynomial of A. The (right) eigenvector associated with the eigenvalue λ is called the (right) Perron-Frobenius eigenvector of the matrix A [35].

Chapter 2

Matrix classes and the linear complementarity problem

The aim of this chapter is to collect the main results in connection with some matrix classes related to LCPs. We consider not only the results which are used later in this thesis, but also present a general idea showing the difficulties and the nice properties as well. The family of matrix classes related to LCPs is really huge and diversified. Cottle wrote a survey paper [10], a well arranged guide for 65 matrix classes that appear in the literature of LCPs. We treat with seven of those: $\mathcal{P}$, $\mathcal{P}_0$, $\mathcal{P}_*(\kappa)$, $\mathcal{P}_*$, row sufficient, column sufficient and the sufficient matrix classes. The $\mathcal{P}_*(\kappa)$-matrices, defined by Kojima et al. [51] in 1991, are in our focus, because this is the widest class where interior point methods are polynomial, however, the complexity depends on the parameter κ, too (it is also a polynomial dependence). The union of the $\mathcal{P}_*(\kappa)$-matrices for all nonnegative κ is the $\mathcal{P}_*$ matrix class. Almost at the same time, in 1989 sufficient matrices were introduced by Cottle et al. [12]. Later, Väliaho proved that the $\mathcal{P}_*$ and the sufficient matrix classes are the same [80].

The sufficient matrix class is between matrix classes $\mathcal{P}$ and $\mathcal{P}_0$, more precisely, the sufficient matrix class includes $\mathcal{P}$-matrices and it is in the $\mathcal{P}_0$ matrix class. The matrix class $\mathcal{P}$ was defined independently from LCPs by Fiedler and Ptak in 1962 [23]. A few years later it first appeared in connection with the LCP in the Ph.D. thesis of Cottle [13]. The $\mathcal{P}_0$-matrix was introduced as a generalization of positive semidefinite matrices by Fiedler and Ptak [24]. Let us note here, that through the thesis we consider positive semidefinite matrices without the assumption of symmetry.

The basis of this chapter is the manuscript [39], which is expanded with some interesting observations. The interested reader can find some more details, for example in the book of Cottle et al. [11], in the book of Kojima et al. [51] and in the papers [6, 10, 23, 24].

At the beginning of this chapter let us introduce some more notations which will be used only in this chapter.

Let $\mathcal{J}, \mathcal{K} \subseteq \mathcal{I}$ and $A \in \mathbb{R}^{n\times n}$, then $A_{\mathcal{J}\mathcal{K}}$ is the submatrix of A whose rows and columns are in $\mathcal{J}$ and $\mathcal{K}$, respectively, i.e., $A_{\mathcal{J}\mathcal{K}} = (a_{jk})_{j\in\mathcal{J},k\in\mathcal{K}}$. We will call $A_{\mathcal{J}\mathcal{J}}$ *principal submatrix*. The *principal minors* of the matrix A are the determinants of principal submatrices, namely the numbers $\det(A_{\mathcal{J}\mathcal{J}})$, where $\mathcal{J} \subseteq \mathcal{I}$.

A *principal pivotal transformation* of the matrix $A = \begin{pmatrix} A_{\mathcal{J}\mathcal{J}} & A_{\mathcal{J}\mathcal{K}} \\ A_{\mathcal{K}\mathcal{J}} & A_{\mathcal{K}\mathcal{K}} \end{pmatrix}$ (where $\mathcal{J}\cup\mathcal{K} = \mathcal{I}$) for nonsingular $A_{\mathcal{J}\mathcal{J}}$ is the matrix

$$\begin{pmatrix} A_{\mathcal{J}\mathcal{J}}^{-1} & -A_{\mathcal{J}\mathcal{J}}^{-1} A_{\mathcal{J}\mathcal{K}} \\ A_{\mathcal{K}\mathcal{J}} A_{\mathcal{J}\mathcal{J}}^{-1} & A_{\mathcal{K}\mathcal{K}} - A_{\mathcal{K}\mathcal{K}} A_{\mathcal{J}\mathcal{J}}^{-1} A_{\mathcal{J}\mathcal{K}} \end{pmatrix}.$$

The following characteristic of the matrix M was introduced by Kojima et al. for $\mathcal{P}$-matrices:

$$\gamma(M) = \min_{\|\mathbf{x}\|_2=1} \max_i x_i \left(Mx\right)_i.$$

2.1 $\mathcal{P}$-matrices

Definition 2.1 *A matrix $M \in \mathbb{R}^{n\times n}$ is a* $\mathcal{P}$-matrix*, if all of its principal minors are positive.*

The set of $\mathcal{P}$-matrices is denoted by $\mathcal{P}$. We will use a similar notation for other following matrix classes, as well.

The following lemma summarizes different characterizations of $\mathcal{P}$-matrices. (The first five statements are the classical results of Fiedler and Ptak [23], for other proofs see [1, 11, 31, 58, 59].)

Lemma 2.2 *The following properties for a matrix M are equivalent:*

1. *M is a $\mathcal{P}$-matrix.*

2. *For every nonzero $\mathbf{x} \in \mathbb{R}^n$ there is an index i such that $x_i[Mx]_i > 0$.*
   Or reformulated: If $x_i[Mx]_i \leq 0$ for every i, then $\mathbf{x} = 0$.
   Or reformulated: $\gamma(M) > 0$.

3. *For every nonzero $\mathbf{x} \in \mathbb{R}^n$ there exists a diagonal matrix $D_{\mathbf{x}}$ with a positive diagonal such that $\mathbf{x}^T D_{\mathbf{x}} M\mathbf{x} > 0$.*

4. *For every nonzero $\mathbf{x} \in \mathbb{R}^n$ there exists a diagonal matrix $H_{\mathbf{x}}$ with a nonnegative diagonal such that $\mathbf{x}^T H_{\mathbf{x}} M\mathbf{x} > 0$.*

5. *Every real eigenvalue of M, as well as of each principal submatrix, is positive.*

6. *M^{-1} is a $\mathcal{P}$-matrix.*

7. *Every principal submatrix of M is a $\mathcal{P}$-matrix.*

8. *There is a vector $\mathbf{x} \geq \mathbf{0}$ such that $M\mathbf{x} > \mathbf{0}$ and M, as well as every principal pivotal transformation of M, satisfies the condition that the rows corresponding to nonpositive diagonal entries are nonpositive.*

9. *All diagonal elements of M and all its principal pivotal transformations are positive.*

10. *For all diagonal matrices D with nonnegative diagonal elements $M + D$ is a $\mathcal{P}$-matrix.*

11. $\det(I - \Lambda + \Lambda M) > 0$ *for all diagonal matrices Λ with nonnegative and less than one diagonal elements (i.e., $0 \leq \Lambda \leq I$).*

12. $I - \Lambda + \Lambda M \in \mathcal{P}$ *for all diagonal matrices $0 \leq \Lambda \leq I$.*

13. *For each $\mathcal{J} \subseteq \mathcal{I}$, $(E^{\mathcal{J}} M E^{\mathcal{J}})\mathbf{x} > \mathbf{0}$ has a solution $\mathbf{x} > \mathbf{0}$. Here $E^{\mathcal{J}}$ is the $n \times n$ diagonal matrix with $(E^{\mathcal{J}})_{jj} = -1$ for $j \in \mathcal{J}$ and $(E^{\mathcal{J}})_{kk} = 1$ for $k \notin \mathcal{J}$.*

A natural generalization of $\mathcal{P}$-matrices is the $\mathcal{P}_0$ matrix class. It may be considered as the closure of the $\mathcal{P}$ matrix class.

Definition 2.3 *A matrix $M \in \mathbb{R}^{n \times n}$ is a* $\mathcal{P}_0$-matrix, *if all of its principal minors are nonnegative.*

A $\mathcal{P}_0$-matrix M is said to be *adequate*, if for each $\mathcal{J} \subseteq \mathcal{I}$ the following two equivalences hold:

$\det(M_{\mathcal{J}\mathcal{J}}) = 0$ if and only if the rows of $M_{\mathcal{J}\mathcal{I}}$ are linearly dependent and

$\det(M_{\mathcal{J}\mathcal{J}}) = 0$ if and only if the columns of $M_{\mathcal{I}\mathcal{J}}$ are linearly dependent.

Alternatively, M is adequate if $XZ\mathbf{x} \leq 0$ implies $Z\mathbf{x} = \mathbf{0}$ for $Z = M$ and $Z = M^T$, too.

Lemma 2.4 ([12]) *If M is nonsingular, then it is a $\mathcal{P}$-matrix if and only if it is adequate.*

Analogously, the corresponding results for P_0-matrices can be proved (the first four due to [24], for other proofs see [11, 31, 51]). One can see the analogy of the first six statements with the appropriate statements of Lemma 2.2. Statement seven again confirms that the $\mathcal{P}_0$ class is in some sense the closure of the $\mathcal{P}$ class. The last statement is very important related to interior point algorithms, because this property ensures the existence and uniqueness of the search direction, the Newton direction (see Chapter 3).

Lemma 2.5 *The following properties for a matrix M are equivalent:*

1. *M is a $\mathcal{P}_0$-matrix.*
2. *For every nonzero $\mathbf{x} \in \mathbb{R}^n$ there is an index i such that $x_i \neq 0$ and $x_i[Mx]_i \geq 0$.*
3. *For every nonzero $\mathbf{x} \in \mathbb{R}^n$ there exists a diagonal matrix $H_\mathbf{x}$ with a nonnegative diagonal such that $\mathbf{x}^T H_\mathbf{x} \mathbf{x} > 0$ and $\mathbf{x}^T H_\mathbf{x} M \mathbf{x} \geq 0$.*
4. *Every real eigenvalue of M, as well as of each principal submatrix, is nonnegative.*
5. *$\det(I - \Lambda + \Lambda M) \geq 0$ for all diagonal matrices $0 \leq \Lambda \leq I$.*
6. *$I - \Lambda + \Lambda M \in \mathcal{P}_0$ for all diagonal matrices $0 \leq \Lambda \leq I$.*
7. *$M + \varepsilon I$ is a $\mathcal{P}$-matrix for every $\varepsilon > 0$.*
8. *The matrix $\begin{pmatrix} Y & X \\ -M & I \end{pmatrix}$ is nonsingular for any positive diagonal matrices Y and X.*

2.2 $\mathcal{P}_*$-matrices

Hereafter we deal with the subclasses of the $\mathcal{P}_0$ matrix class. We start with three well known matrix classes.

Definition 2.6 *A matrix $M \in \mathbb{R}^{n \times n}$ belongs to the class of* positive definite matrices *(PD), if $\mathbf{x}^T M \mathbf{x} > 0$ holds for all $\mathbf{x} \in \mathbb{R}^n \backslash \{\mathbf{0}\}$. Likewise, $M \in \mathbb{R}^{n \times n}$ belongs to the class of* positive semidefinite matrices *(PSD) if $\mathbf{x}^T M \mathbf{x} \geq 0$ holds for all $\mathbf{x} \in \mathbb{R}^n$.*
Furthermore, $M \in \mathbb{R}^{n \times n}$ is a skew-symmetric matrix[1] *(SS), if $\mathbf{x}^T M \mathbf{x} = 0$ for all $\mathbf{x} \in \mathbb{R}^n$.*

We remark that this differs from the usual definition of PD and PSD in linear algebra, as we do not ask for symmetry. So it is not possible to conclude that all eigenvalues of PSD-matrices are real and nonnegative. Let us note that in the subset of all symmetric matrices $\mathcal{P}$ and PD as well as $\mathcal{P}_0$ and PSD coincide.

The $P_*(\kappa)$-matrices were introduced by Kojima, Megiddo, Noma and Yoshise [51], and can also be considered as a generalization of positive semidefinite matrices.

[1] Sometimes the skew-symmetric matrix is called antisymmetric according to its other definition: a matrix M is skew-symmetric, if $M^T = -M$.

Definition 2.7 *Let $\kappa \geq 0$ be a nonnegative number. A matrix $M \in \mathbb{R}^{n\times n}$ is called* $\mathcal{P}_*(\kappa)$-matrix *if*

$$(1+4\kappa)\sum_{i\in\mathcal{I}_+(\mathbf{x})} x_i(Mx)_i + \sum_{i\in\mathcal{I}_-(\mathbf{x})} x_i(Mx)_i \geq 0, \quad \textit{for all } \mathbf{x}\in\mathbb{R}^n, \tag{2.1}$$

where $\mathcal{I}_+(\mathbf{x}) = \{1 \leq i \leq n : x_i(Mx)_i > 0\}$ *and* $\mathcal{I}_-(\mathbf{x}) = \{1 \leq i \leq n : x_i(Mx)_i < 0\}$.

The nonnegative real number κ denotes the weight needed to be used at the positive terms so that the weighted 'scalar product' be nonnegative for each vector $\mathbf{x} \in \mathbb{R}^n$. Therefore, naturally, the $\mathcal{P}_*(0)$ is the positive semidefinite matrix class (n.b. we set aside the symmetry of the matrix M).

It is easy to see, that $\mathcal{P}_*(\kappa_1) \subseteq \mathcal{P}_*(\kappa_2)$ if $\kappa_1 \leq \kappa_2$. Therefore, the smallest κ for which the matrix M is $\mathcal{P}_*(\kappa)$ is specific, it is called the *handicap* and is denoted by $\widehat{\kappa}(M)$. The inequality in the definition of $\mathcal{P}_*(\kappa)$-matrices gives the following lower bound on κ for any vector $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{x}^T M\mathbf{x} < 0$ (in other points $\kappa = 0$ is a proper choice):

$$\widehat{\kappa}(M) \geq \kappa_M(\mathbf{x}) := -\frac{1}{4}\frac{\mathbf{x}^T M\mathbf{x}}{\sum_{i\in\mathcal{I}_+} x_i(Mx)_i},$$

furthermore,

$$\widehat{\kappa}(M) = \begin{cases} 0 & \text{if } M \in PSD \\ \frac{1}{4}\sup\left\{\kappa_M(\mathbf{x}) :\ \mathbf{x}^T M\mathbf{x} < 0\right\} & \text{otherwise.} \end{cases}$$

Hereafter we only write $\widehat{\kappa}$ and $\kappa(\mathbf{x})$ if it is not ambiguous (mostly they will be analyzed for the matrix of the LCP M).

Unfortunately, in the general case function $\kappa(\mathbf{x})$ is not continuous. Figure 2.1 shows the function $\kappa(x, y)$ for the matrix $\left(\begin{smallmatrix} 1 & -1 \\ 3 & 0 \end{smallmatrix}\right)$. There is a discontinuity at the line $x = 0$. This matrix is a $\mathcal{P}_*$ matrix, and its handicap is 0.5 (it is easy to determine by the test of Väliaho, see later).

Definition 2.8 *A matrix $M \in \mathbb{R}^{n\times n}$ is called a* $\mathcal{P}_*$-matrix *if it is a $\mathcal{P}_*(\kappa)$-matrix for some $\kappa \geq 0$, i.e.,*

$$\mathcal{P}_* = \bigcup_{\kappa\geq 0} \mathcal{P}_*(\kappa).$$

The $\mathcal{P}_*$ is the matrix class, where we can warrant, that interior point methods solve the LCP in polynomial time. However the complexity depends on κ, too and we need to know the handicap of the matrix a priori or at least an upper bound on it (see Chapter 5 and for further results Chapter 6).

Almost in the same time, the sufficient matrix class was introduced by Cottle, Pang and Venkateswaran [12]. This is the widest class where the finiteness of the criss-cross algorithm can be proved.

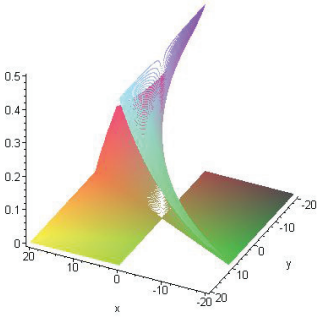


Figure 2.1: The function $\kappa(x, y)$ is not continuous.

Definition 2.9 *A matrix $M \in \mathbb{R}^{n\times n}$ is called a* column sufficient matrix *(CS) if for all* $\mathbf{x} \in \mathbb{R}^n$

$$X(M\mathbf{x}) \leq 0 \quad \textit{implies} \quad X(M\mathbf{x}) = 0,$$

and row sufficient *(RS) if M^T is column sufficient. The matrix M is* sufficient *if it is both row and column sufficient.*[2]

Kojima et al. [51] proved that a $\mathcal{P}_*$-matrix is column sufficient and Guu and Cottle [32] substantiated that it is row sufficient, too. Therefore, each $\mathcal{P}_*$-matrix is sufficient. Väliaho proved the other direction of inclusion [80], so the class of $\mathcal{P}_*$-matrices is equal to the class of sufficient matrices.

We collect some properties of sufficient matrices. Most of them is important in respect to pivot algorithms, for example the sign structure of the matrix has a crucial role in most finiteness proofs. The most essential properties of sufficient matrices in connection with interior point methods will be discussed in Chapter 3.

1. A matrix M is sufficient if and only if

[2]Let us notice that the implication in the definition of sufficient matrix is very similar to that in definition of adequate matrix. Based on this similarity, Lemma 2.4 and the regularity of a $\mathcal{P}$-matrix (see statement 6 of Lemma 2.2), it is easy to see that every $\mathcal{P}$-matrix is sufficient, too.

(a) every principal 2×2 submatrix of M and each of its principal pivotal transformations are sufficient [32].

(b) for every principal pivotal transformation $\overline{M}$ of M: $\overline{m}_{ii} \geq 0$ for all i, furthermore, if $\overline{m}_{ii} = 0$ and $[\,\overline{m}_{ij} = 0$ or $\overline{m}_{ji} = 0\,]$ then $\overline{m}_{ji} = 0$ and $\overline{m}_{ij} = 0$ [32].

(c) $I - \Lambda + \Lambda M$ is sufficient for all diagonal matrices $0 \leq \Lambda \leq I$ [31].

(d) every principal submatrix of order $r + 1$ of M is sufficient, where $r < n$ is the rank of M [79].

2. If M is a sufficient matrix, then

(a) $m_{ii} = 0$ implies $m_{ij} = m_{ji} = 0$ or $m_{ij}\, m_{ji} < 0$ for each $j \neq i$ [81].

(b) the rows $j \in \mathcal{J} \subseteq \mathcal{I}$ are linearly independent if and only if the columns $j \in \mathcal{J}$ are linearly independent [79].

3. The matrix M is sufficient, if (one of the following statements holds)

(a) every principal submatrix of order $n - 1$ of M is sufficient and $\det M > 0$ [79].

(b) every principal submatrix of order r of M is a $\mathcal{P}$-matrix, where $r < n$ is the rank of M [79].

4. If M with rank $r < n$ is such that every principal submatrix of order r is sufficient, then M is sufficient if and only if for every $\mathcal{J} \subseteq \mathcal{I}$ with $|\mathcal{J}\,| = r$

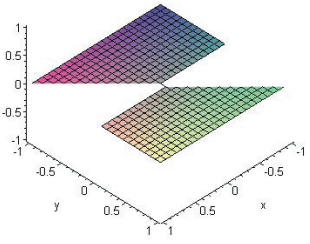
$$\det M_{\mathcal{J}\mathcal{J}} = 0 \Rightarrow \text{the rows and columns } j \in \mathcal{J} \text{ of } M \text{ are linearly dependent [79].}$$

5. $\mathcal{P}_1 \subset \mathcal{P}_*$. Here $\mathcal{P}_1$ denotes the set of all matrices whose principal minors are all positive except one, which is zero [79].

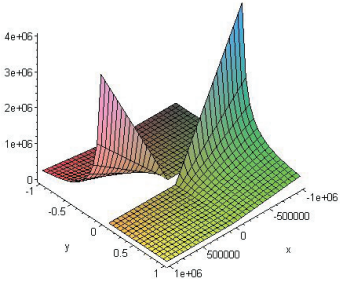
Further properties can be found about $\mathcal{P}_*$ and $\mathcal{P}_*(\kappa)$-matrices in connection with interior point methods in Chapter 3.

The lack of sufficiency of a matrix M means, that $\widehat{\kappa}$ is not finite. It can occur in two ways – there is a point $\mathbf{x}$ where $\kappa(\mathbf{x})$ is not defined (because the set $\mathcal{I}_+$ is empty), or there is a sequence $\{\mathbf{x}_k\}$ such that $\kappa(\mathbf{x}_k)$ tends to infinity. The first case means that the matrix is not column sufficient, because $X(M\mathbf{x}) \lneq \mathbf{0}$. The second case does not always occur when the matrix is not row sufficient (for example if $M = -I$, then $\kappa(\mathbf{x})$ is not defined for all $\mathbf{x} \in \mathbb{R}^n$). However, if the matrix is column sufficient then there exists a sequence $\{\mathbf{x}_k\}$ such that $\lim_{k\to\infty} \kappa(\mathbf{x}_k) = \infty$.

One can see on Figure 2.2 that the function $\kappa(x, y)$ is not defined for the matrix $\left(\begin{smallmatrix} 0 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ over the set $\{(x, y) : \ y(x + y) > 0\}$. In the other picture the second phenomena is illustrated. If the point (x, y) tends to the line $x = 0$ (instead of the point $(0, 0)$), then the function tends to infinity (for example $\kappa(-1, 1/n) = n/4$).



Matrix $\left(\begin{smallmatrix} 0 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ is row sufficient, but not column sufficient.



Matrix $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ is column sufficient, but not row sufficient.

Figure 2.2: The function $\kappa(x, y)$ for two non sufficient matrices.

Väliaho developed two tests. One to decide whether a matrix is sufficient [79] and another to determine the handicap value of sufficient matrices [81]. Unfortunately, both methods are exponential, and there is no known polynomial algorithm for these problems. Tseng proved that the decision problem whether a matrix is column sufficient is co-$\mathbb{NP}$-complete [78], therefore a polynomial algorithm can not be expected. It is an open question whether there is a polynomial time algorithm to compute the handicap of a sufficient matrix.

The classification of Väliaho [79] for 2×2-matrices $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$:

$$\begin{aligned} M \in \mathcal{P} \quad &\Leftrightarrow \quad a > 0,\ d > 0,\ ad - bc > 0, \\ M \in PSD \quad &\Leftrightarrow \quad (a \geq 0,\ d \geq 0,\ (b + c)^2 \leq 4ad, \\ M \in \mathcal{P}_* \quad &\Leftrightarrow \quad (a \geq 0,\ d \geq 0,\ (ad - bc > 0 \vee \\ &\qquad \vee\ (ad - bc = 0 \wedge ((a = 0 \vee d = 0) \Rightarrow b = 0, c = 0)))). \end{aligned}$$

The handicap of 2×2-matrices $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is the following (Väliaho [81]):

$$\begin{aligned} M \in PSD \quad &\Rightarrow \quad \widehat{\kappa} = 0 \quad \text{(by definition)}, \\ M \in \mathcal{P}_* \setminus PSD \quad &\Rightarrow \quad 1 + 4\widehat{\kappa} = \frac{\max\{b^2, c^2\}}{(\sqrt{ad} + \sqrt{ad - bc})^2}, \\ M \in \mathcal{P}_* \setminus (\mathcal{P} \cup PSD) \quad &\Rightarrow \quad 1 + 4\widehat{\kappa} = \max\left\{ \left|\frac{b}{c}\right|, \left|\frac{c}{b}\right| \right\}. \end{aligned}$$

The following properties are known for the handicap of a matrix [80, 81]:

1. If $M \in \mathcal{P} \setminus PD$, then there exists an $\tilde{\mathbf{x}} \neq \mathbf{0}$ such that $\widehat{\kappa}(M) = \kappa_M(\tilde{\mathbf{x}})$, namely $\widehat{\kappa}(M) = \max\{\kappa_M(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}$.
   If $M \notin \mathcal{P}$, then this is not true, not even in dimension 2 (see Figure 2.1).

2. $\widehat{\kappa}(DMD) = \widehat{\kappa}(M)$ for any diagonal matrix D with nonzero elements of the same size as M.

3. $\widehat{\kappa}(B) = \widehat{\kappa}(A)$ if B is a principal pivotal transformation of A.

4. $\widehat{\kappa}(M) \geq \widehat{\kappa}(\overline{M})$ for all principal submatrix $\overline{M}$ of matrix M.

5. If $M = \operatorname{diag}(M_1, M_2)$, then $\widehat{\kappa}(M) = \max\{\widehat{\kappa}(M_1), \widehat{\kappa}(M_2)\}$.

6. If $M \in \mathcal{P}_*$ and D is a nonnegative diagonal matrix of the same size as M, then $\widehat{\kappa}(M + D) \leq \widehat{\kappa}(M)$.

7. If $M \in \mathcal{P}_*$ and D is a nonnegative diagonal matrix of the same size as M, then $\widehat{\kappa}\left(\left(\begin{smallmatrix} M & I \\ -I & D \end{smallmatrix}\right)\right) = \widehat{\kappa}(M)$.
   As a corollary, for nonnegative scalars $d \geq 0$, $\widehat{\kappa}\left(\left(\begin{smallmatrix} M & -e_1^T \\ e_1^T & d \end{smallmatrix}\right)\right) = \widehat{\kappa}(M)$, where $\mathbf{e}_1$ is the first unit vector.

8. Let $M \in \mathcal{P}_*$ with $m_{ij} = 0$ for $1 \leq i, j \leq n-1$. Then $1 + 4\widehat{\kappa}(M) = \frac{\max_i |m_{ni}/m_{in}|}{\min_i |m_{ni}/m_{in}|}$ (define $0/0$ to 1).

9. Let $M \in \mathcal{P}_*$ and $m_{jk} = m_{jh}$, $m_{kj} = m_{hj}$ for all $j \neq k$, and $m_{kk} \geq m_{hh}$. Then $\widehat{\kappa}(M_{\mathcal{J}\mathcal{J}}) = \widehat{\kappa}(M)$ for $\mathcal{J} = \mathcal{I} \setminus \{k\}$.

10. If $A \in \mathcal{P}_* \setminus \mathcal{P}$, then $\widehat{\kappa}(A) = \max\left\{\widehat{\kappa}(B_{\mathcal{J}\mathcal{J}}) : B \text{ is a principal pivotal transformation of } A \text{ and } \mathcal{J} = \mathcal{I} \setminus \{i\} \text{ for some } i\right\}$.

11. The handicap of a sufficient matrix is the same as the handicap of its transpose, i.e., if $M \in \mathcal{P}_*$, then $\widehat{\kappa}(M^T) = \widehat{\kappa}(M)$.

12. It is a conjecture that $\widehat{\kappa}(M)$ is a continuous function of the entries of M for $M \in \mathcal{P}_*$. This is only proved for 2×2 matrices and for $\mathcal{P}$-matrices at the time this thesis was written.

The characteristic $\gamma(M)$ is defined at the beginning of this chapter. Kojima et al. introduced an expression of $\gamma(M)$, too, $\bar{\gamma}(M) = \sqrt{\gamma(M)\gamma\left(M^{-1}\right)}$. Furthermore, let us denote the smallest eigenvalue of matrix M by $\lambda_{\min}(M)$.

Making use of these characteristics, Kojima et al. introduced global optimization problems which determine upper bounds on the handicap of matrix M, if $M \in \mathcal{P}$. There is no known better estimation, which also shows the difficulty of evaluating the handicap of a matrix even for $\mathcal{P}$-matrices.

Theorem 2.10 ([51]) *Let M be a $\mathcal{P}$-matrix. Furthermore, let*

$$\begin{aligned} \kappa^* &= \max\left\{\frac{-\lambda_{\min}(M)}{4\gamma(M)}, 0\right\} \\ \kappa^{**} &= \frac{1}{4\bar{\gamma}(M)}. \end{aligned}$$

Then $M \in \mathcal{P}_*(\kappa^*) \cap \mathcal{P}_*(\kappa^{**}) = \mathcal{P}_*(\min\{\kappa^*, \kappa^{**}\})$, and therefore $\mathcal{P} \subset \mathcal{P}_*$.

In the general case it can not be stated which of κ^* and κ^{**} is smaller.

2.3 Eigenvalues

In this section we collect a few results about the eigenvalues of the mentioned matrix classes. In the literature, there are statements about eigenvalues of the matrix only for the $\mathcal{P}$ and $\mathcal{P}_0$ matrix classes. There are no results specifically for sufficient matrices. The below listed properties are almost all negative results, namely state the lack of some kind of properties for the eigenvalues of sufficient matrices.

Recall that the definition of positive semidefinite matrices in this thesis is different than in linear algebra, because we do not require the symmetry of the matrix. Due to this, a positive semidefinite matrix may also have complex number as eigenvalues.

Proposition 2.11 ([39]) *Let $M \in \mathbb{R}^{n\times n}$ be a PSD-matrix. Then* $\mathrm{re}(\lambda) \geq 0$ *for all eigenvalues λ of M.*

The first two statements in the following proposition are a direct corollary of the definition. Furthermore, Illés and Wenzel considered the sequence of 3 dimensional $\mathcal{P}$-matrices

$M(a) := \left(\begin{smallmatrix} 1 & 0 & a \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{smallmatrix}\right)$, where $a > -1$. (The one and two dimensional principal minors of the matrix are one, and its determinant is $1 + a$, so this is a $\mathcal{P}$-matrix indeed.) The eigenvalues are $1 + a^{1/3}$ and $1 - 1/2\ a^{1/3} \pm i\ \sqrt{3}/2\ a^{1/3}$, therefore $\lim_{a\to\infty}\ \mathrm{re}(\lambda_{2,3}(\mathrm{a})) = -\infty$, namely the real part of one of the eigenvalues tends to infinity and the real part of the other two eigenvalues tend to minus infinity.

Proposition 2.12 ([39]) *The real part of all eigenvalues of $\mathcal{P}$-matrices of dimension 2 is positive. The real part of all eigenvalues of $\mathcal{P}_0$-matrices of dimension 2 is nonnegative.*
There is no lower bound on the real part of the eigenvalues of $\mathcal{P}$-matrices for dimension $n \geq 3$.

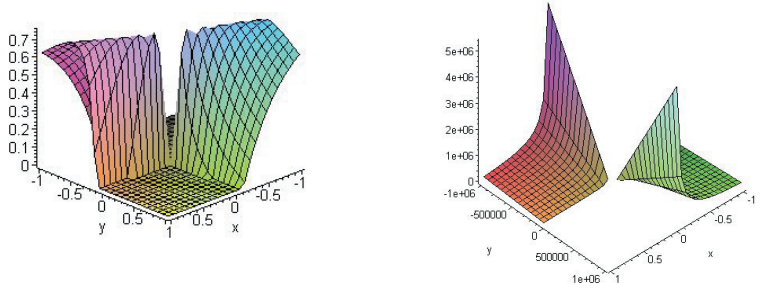
This result means that a sufficient matrix in 2 dimension has only eigenvalues with a non-negative real part. The result is tight, because skew-symmetric matrices have only pure imaginary eigenvalues.

The converse of the first statement is not true. There are matrices with positive real eigenvalues which are not even in $\mathcal{P}_0$ and matrices with imaginary eigenvalues and a positive real part which are not $\mathcal{P}_0$ [39].

Originally, the next result was due to [50], but see [33] for a simpler proof and [22] for an alternate proof.

Proposition 2.13 *A set of complex numbers $\{\lambda_1, \ldots, \lambda_n\}$ are the eigenvalues of an $n \times n$ $\mathcal{P}$-matrix ($\mathcal{P}_0$-matrix) if and only if the polynomial $\prod_1^n (t + \lambda_i) = \sum_0^n b_i t^i$ satisfies $b_i > 0$ ($b_i \geq 0$).*
A complex number $\lambda = re^{i\,\theta}$ is an eigenvalue of an $n \times n$ $\mathcal{P}$-matrix if and only if $|\theta - \pi| > \pi/n$.
A nonzero $\lambda = re^{i\,\theta}$ is an eigenvalue of an $n \times n$ $\mathcal{P}_0$-matrix if and only if $|\theta - \pi| \geq \pi/n$.

The following example shows that the eigenvalues do not determine the $\mathcal{P}_*$ property. Let us consider the matrix $M = \begin{pmatrix} 1 & 8 \\ -1 & 1 \end{pmatrix}$. Using the test of Väliaho [81], $\widehat{\kappa} = 0.75$. The matrix $U = \begin{pmatrix} 0.5 & \sqrt{0.75} \\ \sqrt{0.75} & -0.5 \end{pmatrix}$ is orthogonal, therefore the eigenvalues of M and UMU are the same. But the transformed matrix $UMU = \begin{pmatrix} 4.031 & -2.75 \\ 6.25 & -2.031 \end{pmatrix}$ is not $\mathcal{P}_*$, not even a $\mathcal{P}_0$-matrix, for example, because there is a negative value in the diagonal. Figure 2.3 illustrates the $\kappa(x, y)$ functions of the two matrices. Since the transformed matrix is not $\mathcal{P}_*$, there are points where this function is not defined.



$\kappa_M(\mathbf{x})$ $\kappa_{UMU}(\mathbf{x})$

Figure 2.3: The function $\kappa(x, y)$ for two matrices with same eigenvalues.

Based on these observations we can state that there is no connection between the eigenvalues of the matrix and the sufficiency.

Further results related to eigenvalues can be found in [22, 33, 34, 50].

2.4 Row and column scaling of $\mathcal{P}_*$-matrices

We now consider the column and row scaling of matrices belonging to different classes.

Definition 2.14 *The* $(\mathbf{p}, \mathbf{q})$-scaling *of a matrix* M *is the matrix* PMQ*, where* $P = diag(\mathbf{p})$ *and* $Q = diag(\mathbf{q})$ *are arbitrary diagonal matrices with* $p_i q_i > 0$ *for all* i*.*

It is a scaling of the rows with P and a scaling of the columns with Q.

Generally the $(\mathbf{p}, \mathbf{q})$-scaling cancels the positive semidefiniteness and the skew-symmetricity of a matrix, unless $\mathbf{p} = \mathbf{q}$. Kojima et al. and Väliaho independently showed that the matrix class $\mathcal{P}_*$ is closed under scaling, but the class of $\mathcal{P}_*(\kappa)$-matrices for a fix κ is not.

Lemma 2.15 ([51, 81]) *The set* $\mathcal{P}_*$ *is invariant under* $(\mathbf{p}, \mathbf{q})$*-scaling. If* $M \in \mathcal{P}_*(\kappa)$*, then* $PMQ \in \mathcal{P}_*(\kappa')$*, where* κ' *is such that* $(1 + 4\kappa') \min_i\{p_i/q_i\} = (1 + 4\kappa) \max_i\{p_i/q_i\}$*.*

Illés et al. [39] constructed via $(\mathbf{p}, \mathbf{q})$-scaling from a given matrix in $\mathcal{P}_*(\kappa)$ a matrix with arbitrary large κ': Let $M = \left(\begin{smallmatrix} 1 & -1 \\ 2 & 0 \end{smallmatrix}\right)$, which is sufficient by the test of Väliaho, but not PSD (take $\mathbf{x} = (1 \ \ -2)^T$, then $\mathbf{x}^T M \mathbf{x} = -1$) and not $\mathcal{P}$ (there is a zero in the diagonal). In this

situation $\widehat{\kappa}(M) = 0.25$. Scaling with $\mathbf{p} = (k, 1)$ for $k > 2$ and $\mathbf{q} = (1, 1)$ gives a matrix $PMQ \in \mathcal{P}_*(\kappa')$ with $\kappa' = (k-2)/8$, where κ' is the handicap of the matrix PMQ, namely it is minimal. Observe that the κ' given by Lemma 2.15 is not the handicap of the scaled matrix, only an upper bound on it (in the example above it is $2k$, however, the handicap is only $(k-2)/2$).

As a special case, it can be proved that scaling can lead out of the class of positive semidefinite matrices. However, the classes $\mathcal{P}$ and $\mathcal{P}_0$ are closed under scaling.

Lemma 2.16 ([39]) *The set $PSD = \mathcal{P}_*(0)$ is not invariant under $(\mathbf{p}, \mathbf{q})$-scaling. The classes $\mathcal{P}$ and $\mathcal{P}_0$ are invariant under $(\mathbf{p}, \mathbf{q})$-scaling.*

According to Lemma 2.12, $\mathcal{P}$-matrices may have arbitrary small eigenvalues. Thus, it is interesting to note that one may always find a $(\mathbf{p}, \mathbf{q})$-scaling for a $\mathcal{P}_0$-matrix (in this case the elements in the diagonal are nonnegative by definition), such that the real part of all eigenvalues is greater than some prescribed negative tolerance ν. The result is based on the Gerschgorins eigenvalue-inclusion theorem (see e.g. [76]).

Lemma 2.17 ([39]) *Let $\nu < 0$ be an arbitrary negative number and A be a real $n \times n$-matrix with nonnegative diagonal elements. Then there exists a positive diagonal matrix D such that $re(\lambda(DAD)) > \nu$ holds.*

We close this section with an observation. It is easy to see that by multiplying all elements of a sufficient matrix with a nonnegative number α^2 the handicap of the matrix will be the same. Although, the eigenvalues of the new matrix are α^2 times larger.

Lemma 2.18 ([39]) *Let $M \in \mathcal{P}_*(\kappa)$ with eigenvalues λ_i. Then applying symmetric scaling $(\alpha I)M(\alpha I)$ for arbitrary $\alpha \in \mathbb{R}$ leaves the handicap unchanged, but may scale the eigenvalues to arbitrary (except the sign) values $\alpha^2\lambda_i$.*

2.5 Structure of the sets $PSD, \mathcal{P}, \mathcal{P}_*$ and $\mathcal{P}_0$

It is known [51] that the following relations hold among matrix classes $SS \subsetneq PSD \subsetneq \mathcal{P}_* \subsetneq \mathcal{P}_0$, $SS \cap \mathcal{P} = \emptyset$, $PSD \cap \mathcal{P} \neq \emptyset$, $\mathcal{P} \subsetneq \mathcal{P}_*$, $\mathcal{P}_*(\kappa_1) \subsetneq \mathcal{P}_*(\kappa_2)$ for $\kappa_1 < \kappa_2$, $\mathcal{P}_*(0) \equiv PSD$.

According to definitions, it is easy to see that these sets are cones, moreover convex cones in case of semidefinite and skew-symmetric matrices. Illés and Wenzel gave two $\mathcal{P}$-matrices, $\left(\begin{smallmatrix} 1 & 4 \\ -1 & 1 \end{smallmatrix}\right)$ and its transpose $\left(\begin{smallmatrix} 1 & -1 \\ 4 & 1 \end{smallmatrix}\right)$, whose sum $\left(\begin{smallmatrix} 2 & 3 \\ 3 & 2 \end{smallmatrix}\right)$ is not even in $\mathcal{P}_0$ (its determinant is negative). Therefore the set of $\mathcal{P}$-matrices is not convex. Since $\mathcal{P} \subset \mathcal{P}_* \subset \mathcal{P}_0$ it means that the other two cones are also not convex. The closedness of the sets follows from the definitions and the continuity of the determinant function.

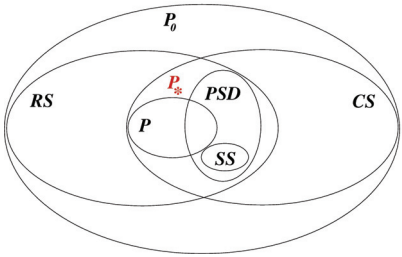


Figure 2.4: The matrix classes.

Proposition 2.19 ([39]) *SS and PSD are convex cones.*
Each of the sets $\mathcal{P}$, $\mathcal{P}_*$ and $\mathcal{P}_0$ form a cone, but none of them is convex.
The set $\mathcal{P}$ is open, and the sets SS, PSD and $\mathcal{P}_0$ are closed.

Illés and Wenzel proved the following statement with two counterexamples.

Proposition 2.20 ([39]) *The cone of $\mathcal{P}_*$-matrices is not closed, neither open.*

We close this section with a remark: all matrix classes mentioned above, namely SS, PSD, $\mathcal{P}$, $\mathcal{P}_*(\kappa)$, $\mathcal{P}_*$ and $\mathcal{P}_0$, enjoy the nice property that if a matrix belongs to one of these classes, then any principal submatrix of the matrix and any principal pivotal transformation of it does as well [11].

2.6 Complexity issues

We would like to investigate how difficult it is to decide whether a matrix is sufficient or not. With this aim in view we collect some related complexity results in this section. As we have already mentioned, the best known test for analyzing sufficiency of a matrix is given by Väliaho [80]. The procedure with the asymptotically best worst-case operation counted for general $n \times n$ matrices needs $\frac{1}{3}\,2^{n-3}n(n-1)(n+10)$ operations. The exponential complexity property is a general case because this algorithm is a recursive procedure. This fact, i.e., that only exponential time methods are known for decision problems related to sufficient matrices, is not surprising being aware of the result of Tseng [78] that the decision problem whether M is column sufficient, is co-$\mathbb{NP}$-complete.

As we have already seen a matrix is sufficient if and only if there is a finite κ with which the matrix is $\mathcal{P}_*(\kappa)$. We do not need the exact value of the handicap, it is enough to know

an upper bound on it. However, as in many problems, the decision on the feasibility of the problem (Is there a suitable κ?) is as difficult as to solve the problem (What is the handicap?) (see e.g. [83]). Väliaho also presented an exponential method to determine the handicap of a sufficient matrix [81]. This is the only kind of such an algorithm in the literature.

Chung showed that the 0-1 equality constrained knapsack problem can be reduced to an LCP with integer data.

Theorem 2.21 ([8]) *The problem of finding a rational solution* $(\mathbf{x}, \mathbf{s})$ *to an LCP with an integer square matrix* M *and an integer vector* $\mathbf{q}$ *is* $\mathbb{NP}$*-complete, even if* M *is restricted to be negative (semi)definite.*

Furthermore, Kojima et al. [51] show that the LCP for a $\mathcal{P}_0$-matrix is also $\mathbb{NP}$-complete.

Hereafter, we collect some complexity results about the decision problems in connection with matrix classes. Let us first consider the copositive matrices which are well known due to the Lemke algorithm. A matrix A is called *copositive*, if $\mathbf{x}^T A\mathbf{x} \geq 0$ for all $\mathbf{x} \geq \mathbf{0}$. Murty et al. gave a reduction of the subset sum problem to the decision problem "Is the given matrix copositive?".

Theorem 2.22 ([62]) *Let an integer square matrix* A *be given, the decision problem whether the matrix* A *is copositive, is co-*$\mathbb{NP}$*-complete.*

The equivalence of the problem whether $M \in \mathcal{P}$ and the regularity of rank-one matrix polytopes is shown by Coxson. The decision problem of the singularity of an interval matrix with a rank-one radius is known to be $\mathbb{NP}$-complete and is a special case of the latter problem [64].

Theorem 2.23 ([14]) *Let a real square matrix* M *be given, the decision problem "is* M *a* $\mathcal{P}$*-matrix?" is co-*$\mathbb{NP}$*-complete.*

Tseng gave a reduction of the problem of 1-norm maximization over a parallelotope to the decision problem for the $\mathcal{P}$ and the column sufficient matrix class. In addition, he showed the equivalence of the decision problem for the $\mathcal{P}_0$ matrix class and the decision problem for the $\mathcal{P}$ matrix class.

Theorem 2.24 ([78]) *Let an integer square matrix* M *be given, the decision problem for matrix class* $\mathcal{P}$*,* $\mathcal{P}_0$ *and for sufficient matrices is co-*$\mathbb{NP}$*-complete.*

Naturally, it means that the decision problem for the sufficient matrix class is also co-$\mathbb{NP}$-complete by the definition.

The last mentioned complexity result refers to the quadratic programming problem. On the one hand, when the constraints are quadratic and the objective function is linear, Sahni and later Garey and Johnson stated that the problem is $\mathbb{NP}$-hard. On the other hand, the quadratic programming problem with linear constraints and a symmetric matrix in the objective function is also $\mathbb{NP}$-hard, as claimed by Sahni and Garey and Johnson. Later Vavasis proved that this problem is in $\mathbb{NP}$, therefore it is $\mathbb{NP}$-complete.

Theorem 2.25 ([28, 72, 82]) *The quadratic programming problem* $\min\left\{\mathbf{x}^T Q\mathbf{x} : A\mathbf{x} \geq \mathbf{b}\right\}$ *with a symmetric matrix* Q *is* $\mathbb{NP}$*-complete. If the constraints are quadratic and the objective is linear,* $\mathbb{NP}$*-hardness is established.*

2.7 The solution set of LCPs for different matrix classes

This section discusses how the different properties of the solution set depend on which class the matrix of a problem belongs to. First, we present the result of Jansen.

Proposition 2.26 ([46]) *The solution set of an LCP is a finite union of polyhedral sets.*

Cottle et al. examined the cardinality of the solution set of the LCP.

Proposition 2.27 ([11]) *Let* M *be a real matrix. The following statements are equivalent.*

1. M *is nondegenerate (all principal minors are not* 0*).*
2. *The LCP has a finite number of solutions (possibly zero) for all* $\mathbf{q}$*.*
3. *Every solution of the LCP, if it exists, is locally unique.*

$\mathcal{P}$-matrices are not degenerate by definition. But if the coefficient matrix of the LCP belongs to $\mathcal{P}$, the solution set is even more special than in the case of nondegenerate matrices.

Proposition 2.28 ([11]) *Let* M *be a real matrix. Then the following two statements are equivalent.*

1. $M \in \mathcal{P}$*.*
2. *The LCP has a unique solution for all* $\mathbf{q} \in \mathbb{R}^n$*.*

In contrast with this nice result, if the matrix M is a $\mathcal{P}_0$-matrix (which is a generalization of a $\mathcal{P}$-matrix), there is an example, when the solution set $\mathcal{F}_*$ is empty, however there is a feasible solution. On the other hand the solution set can be unbounded and nonconnected [31].

Proposition 2.29 ([48]) *Let $M \in \mathcal{P}_0$. Then,*

1. *if $\mathcal{F}_*$ has a bounded connected component, then $\mathcal{F}_*$ is connected.*
2. *for any $\mathbf{q}$, an LCP has a unique solution if and only if it has a locally unique solution.*
3. *the number of solutions of an LCP is either zero, one, or infinite.*

Cao and Ferris [6] proved that in dimension 2 the solution set of an LCP with a $\mathcal{P}_0$-matrix is connected for all $\mathbf{q}$. On the contrary, Jones and Gowda [48] provide an example in dimension 3, where an LCP with a $\mathcal{P}_0$-matrix has a nonconnected and unbounded solution set.

Cottle et al. determined the set of matrices where the function $M\mathbf{x}$ is constant over the solution set of the LCP.

Proposition 2.30 ([11]) *Let M be a real matrix. The following two statements are equivalent.*

1. *$M \in \mathcal{P}_0$ and for each index set $\mathcal{J}$ with $\det M_{\mathcal{J}\mathcal{J}} = 0$, the columns in $\mathcal{J}$ of M are linearly dependent.*
2. *For all $\mathbf{q}$ for which the LCP has a solution, if $\mathbf{x}_1, \mathbf{x}_2$ are any two solutions of the LCP, then $M\mathbf{x}_1 = M\mathbf{x}_2$.*

The equivalence of the following first two assertions is a direct corollary of Jansen's result, Proposition 2.26.

Proposition 2.31 ([12]) *Let M be a real matrix. The following statements are equivalent.*

1. *The solution set of the LCP is polyhedral.*
2. *The solution set of the LCP is convex.*
3. *$\mathbf{x}_1^T(M\mathbf{x}_2 + \mathbf{q}) = \mathbf{x}_2^T(M\mathbf{x}_1 + \mathbf{q}) = 0$ for all solutions $\mathbf{x}_1, \mathbf{x}_2$.*

Let us take into consideration the following quadratic programming problem related to the LCP:

$$\left.\begin{aligned} \min\ & \mathbf{x}^T(M\mathbf{x} + \mathbf{q}) \\ & M\mathbf{x} + \mathbf{q} \geq \mathbf{0} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}\right\} \tag{2.2}$$

There are special connections between the Karush-Kuhn-Tucker (KKT) points of (2.2) and the solution set of the LCP if the matrix M is sufficient.

Proposition 2.32 ([12]) *Let $M \in \mathbb{R}^{n \times n}$ be a matrix.*

1. *M is column sufficient if and only if for all $\mathbf{q}$ the LCP has a (possibly empty) convex solution set.*

2. *M is row sufficient if and only if for all $\mathbf{q}$ if $(\mathbf{x}, \mathbf{u})$ is a Karush-Kuhn-Tucker pair for the quadratic program (2.2) then $\mathbf{x}$ is a solution to the LCP.*

3. *M is sufficient if and only if the set of KKT-points of the associated quadratic program is convex and equal to the solution set of the LCP for all $\mathbf{q}$.*

Chapter 3

Interior point methods

The aim of this section is to give an overview about the basic theory of interior point methods related to LCPs for $\mathcal{P}_*(\kappa)$-matrices. We collect the most important results and present a general schema of interior point algorithms.

We have already seen in Section 1.2.1 that the linear programming problem can be reformulated as a special LCP, where the coefficient matrix M is skew-symmetric. The primal-dual interior point methods for linear programming problems solve this LCP problem. As the first step of solving this LCP, the problem is relaxed to the *central path problem* as follows:

$$\left.\begin{aligned} -M\mathbf{x}+\mathbf{s} &= \mathbf{q} \\ \mathbf{x},\mathbf{s} &\geq \mathbf{0} \\ \mathbf{xs} &= \mu\mathbf{e} \end{aligned}\right\} \qquad (CPP_\mu)$$

where μ is a positive number. The solutions of the CPP problem constitute the *central path* which is denoted by

$$\mathcal{C} := \left\{(\mathbf{x},\mathbf{s}) \in \mathcal{F}^+ : \mathbf{xs} = \mu\mathbf{e}, \text{ for some } \mu > 0\right\}.$$

It was originally introduced in linear programming by Sonnevend [73, 74] and independently by Megiddo [54]. It is well known that the central path is a one dimensional, infinitely many times differentiable curve which tends to a solution of the LCP when μ tends to zero in case of linear programming. Furthermore, the limit point of the central path is a strictly positive solution, namely $\mathbf{x}+\mathbf{s} > 0$, thus for each coordinate i exactly one of x_i and s_i is zero (see e.g. [70]).

The question arises whether the central path remains well defined if the coefficient matrix is not skew-symmetric, only sufficient. Kojima et al. proved that if the interior point assumption holds, namely the set $\mathcal{F}^+$ is not empty, then the CPP problem has a unique solution for all positive μ, therefore the central path exists and is unique, furthermore, it remains a

one dimensional, infinitely many times differentiable curve, which tends to a solution of the LCP [51].

We show the main steps of the proof of Illés et al. following manuscript [38], as they did not use heavy mathematics like the implicit function theorem. The manuscript is not published, therefore the proofs are also presented. This approach can be found in book chapter [41] (in Hungarian) and in paper [77] for the linear programming problem, namely, for LCPs with skew-symmetric coefficient matrices.

3.1 The complementarity level set of the LCP

The *complementarity level set* of the LCP is defined for all $\mathbf{w} \in \mathbb{R}^n_{\oplus}$ by

$$\mathcal{L}_{\mathbf{w}} = \{\, (\mathbf{x}, \mathbf{s}) \in \mathcal{F} \,|\, \mathbf{x}\mathbf{s} \leq \mathbf{w} \,\}. \tag{3.1}$$

To show the compactness of the level sets $\mathcal{L}_{\mathbf{w}}$, we need to bound $\mathbf{x}^T M \mathbf{x}$ from below for all $\mathbf{x} \in \mathbb{R}^n$. If M is a positive semidefinite matrix, then $\mathbf{x}^T M \mathbf{x} \geq 0$ holds for all $\mathbf{x} \in \mathbb{R}^n$; however, a weaker condition is sufficient for our purpose. We assume that matrix M belongs to the $P_*(\kappa)$ matrix class.

Lemma 3.1 ([38]) *If $M \in \mathcal{P}_*(\kappa)$ for any $\kappa \geq 0$ and $\mathcal{F}^+ \neq \emptyset$ then for every $\mathbf{w} \in \mathbb{R}^n_{\oplus}$ the set $\mathcal{L}_{\mathbf{w}}$ is compact.*

Proof. An empty set is compact, so let us consider the case $\mathcal{L}_{\mathbf{w}} \neq \emptyset$. Since there exists an interior point $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{F}^+$ and by the definition of $\mathcal{P}_*(\kappa)$-matrices (see Definition 2.7), we have for every $(\mathbf{x}, \mathbf{s}) \in \mathcal{L}_{\mathbf{w}}$

$$\begin{aligned}
\left(\mathbf{x} - \mathbf{x}^0\right)^T \left(\mathbf{s} - \mathbf{s}^0\right) &= \left(\mathbf{x} - \mathbf{x}^0\right)^T M \left(\mathbf{x} - \mathbf{x}^0\right) \geq -4\,\kappa \sum_{i \in \mathcal{I}_+(\mathbf{x})} \left(x_i - x_i^0\right)\left(s_i - s_i^0\right) \\
&\geq -4\,\kappa \sum_{i \in \mathcal{I}_+(\mathbf{x})} \left(x_i s_i + x_i^0 s_i^0\right) \geq -4\,\kappa \left(\mathbf{x}^T \mathbf{s} + (\mathbf{x}^0)^T \mathbf{s}^0\right) \\
&\geq -4\,\kappa\, \mathbf{e}^T \left(\mathbf{w} + \mathbf{w}^0\right),
\end{aligned} \tag{3.2}$$

where $\mathbf{w}^0 = \mathbf{x}^0\,\mathbf{s}^0$. Furthermore,

$$\begin{aligned}
\mathbf{x}^T \mathbf{s}^0 + \left(\mathbf{x}^0\right)^T \mathbf{s} &= \mathbf{x}^T \mathbf{s} + \left(\mathbf{x}^0\right)^T \mathbf{s}^0 - \left(\mathbf{x} - \mathbf{x}^0\right)^T \left(\mathbf{s} - \mathbf{s}^0\right) \leq \mathbf{e}^T \mathbf{w} + \mathbf{e}^T \mathbf{w}^0 + 4\,\kappa\, \mathbf{e}^T \left(\mathbf{w} + \mathbf{w}^0\right) \\
&\leq (1 + 4\,\kappa)\, \mathbf{e}^T \left(\mathbf{w} + \mathbf{w}^0\right).
\end{aligned}$$

The points $(\mathbf{x}^0, \mathbf{s}^0)$ and $(\mathbf{x}, \mathbf{s})$ are feasible, therefore nonnegative, so

$$x_i s_i^0 \leq (1 + 4\kappa)\, \mathbf{e}^T (\mathbf{w} + \mathbf{w}^0) \qquad \forall\, i.$$

Using the positivity of the vector $\mathbf{s}^0$ we have

$$x_i \le (1+4\kappa)\,\frac{\mathbf{e}^T(\mathbf{w}+\mathbf{w}^0)}{s_i^0}$$

and similarly

$$s_i \le (1+4\kappa)\,\frac{\mathbf{e}^T(\mathbf{w}+\mathbf{w}^0)}{x_i^0}.$$

This shows that the set $\mathcal{L}_w$ is bounded. The closedness of the level set follows from the fact that $\mathcal{F}$ is closed and the mapping $(\mathbf{x},\mathbf{s}) \to \mathbf{x}\,\mathbf{s}$ is continuous. ■

The proof of boundedness in the previous lemma was based on the same idea as the proof of Lemma 4.5 in [51].

In the next lemma we present the main technical tool. It can be verified that if the interior of the rectangle $\mathcal{T}(\overline{\mathbf{w}},\hat{\mathbf{w}})$ is not empty for any positive $\overline{\mathbf{w}}$ and $\hat{\mathbf{w}}$, then we can take a step from an arbitrary interior point $(\overline{\mathbf{x}},\overline{\mathbf{s}})$, where $\overline{\mathbf{w}} = \overline{\mathbf{x}}\,\overline{\mathbf{s}}$, toward the level set $\mathcal{L}_{\hat{\mathbf{w}}}$.

Lemma 3.2 ([38]) *Let us suppose that $M \in \mathcal{P}_*(\kappa)$ for any $\kappa \ge 0$, $(\overline{\mathbf{x}},\overline{\mathbf{s}}) \in \mathcal{F}^+, \overline{\mathbf{w}} = \overline{\mathbf{x}}\,\overline{\mathbf{s}}$ and $\hat{\mathbf{w}} > \mathbf{0}$. Let us define the box*

$$\mathcal{T}(\overline{\mathbf{w}},\hat{\mathbf{w}}) := \Big\{\mathbf{w} \in \mathbb{R}^n : \min\{\overline{w}_i,\hat{w}_i\} \le w_i \le \max\{\overline{w}_i,\hat{w}_i\}\Big\}.$$

If $int\,\mathcal{T}(\overline{\mathbf{w}},\hat{\mathbf{w}}) \ne \emptyset$, then there exists such a direction $(\Delta\mathbf{x},\Delta\mathbf{s})$ and such a step length $\alpha \in (0,1]$ for which $\mathbf{x}^+ = \overline{\mathbf{x}}+\alpha\Delta\mathbf{x}$, $\mathbf{s}^+ = \overline{\mathbf{s}}+\alpha\Delta\mathbf{s}$ satisfies $(\mathbf{x}^+,\mathbf{s}^+) \in \mathcal{F}^+$ and $\mathbf{w}^+ \in int\,\mathcal{T}(\overline{\mathbf{w}},\hat{\mathbf{w}})$, where $\mathbf{w}^+ := \mathbf{x}^+\mathbf{s}^+$.

Proof. Let the direction $(\Delta\mathbf{x},\Delta\mathbf{s})$ be the solution of the system

$$\begin{aligned} -M\Delta\mathbf{x}+\Delta\mathbf{s} &= \mathbf{0} \\ \overline{S}\Delta\mathbf{x}+\overline{X}\Delta\mathbf{s} &= \hat{\mathbf{w}}-\overline{\mathbf{w}}, \end{aligned} \tag{3.3}$$

where $\overline{X} = diag(\overline{\mathbf{x}})$ and $\overline{S} = diag(\overline{\mathbf{s}})$. A $\mathcal{P}_*(\kappa)$-matrix is also a $\mathcal{P}_0$-matrix, therefore the system (3.3) has a unique solution by the last statement of Lemma 2.5. According to the assumption of the lemma, $int\,\mathcal{T} \ne \emptyset$, so there is no index i such that $\overline{w}_i = \hat{w}_i$. It means, that the direction $(\Delta\mathbf{x},\Delta\mathbf{s})$ can not be the zero vector.

We need to choose $\alpha > 0$ such that it satisfies inequalities

$$\mathbf{x}^+ = \overline{\mathbf{x}}+\alpha\Delta\mathbf{x} > \mathbf{0} \quad \text{and} \quad \mathbf{s}^+ = \overline{\mathbf{s}}+\alpha\Delta\mathbf{s} > \mathbf{0} \tag{3.4}$$

and

$$\min\{\overline{w}_i,\hat{w}_i\} < w_i^+ = x_i^+ s_i^+ < \max\{\overline{w}_i,\hat{w}_i\} \quad \forall i. \tag{3.5}$$

When Δx_i and Δs_i are nonnegative, then (3.4) holds for all positive α, otherwise we get upper bounds on α:

$$\alpha_1 := \min\left\{-\frac{\overline{x}_i}{\Delta x_i} \ : \ \Delta x_i < 0\right\}, \qquad \alpha_2 := \min\left\{-\frac{\overline{s}_i}{\Delta s_i} \ : \ \Delta s_i < 0\right\}.$$

Now we investigate inequality (3.5). By definition $\mathbf{x}^+\mathbf{s}^+ = (1-\alpha)\overline{\mathbf{w}} + \alpha\hat{\mathbf{w}} + \alpha^2\Delta\mathbf{x}\Delta\mathbf{s}$. Let us introduce the following index sets:

$$\mathcal{I}_{\overline{\mathbf{w}}} := \{i \ : \ \overline{w}_i < \hat{w}_i\} \quad \text{and} \quad \mathcal{I}_{\hat{\mathbf{w}}} := \{i \ : \ \hat{w}_i < \overline{w}_i\}.$$

If $i \in \mathcal{I}_{\overline{\mathbf{w}}}$, then inequality $\overline{w}_i < w_i^+$ implies the following upper bound on α:

$$\alpha_3 := \min_{i\in\mathcal{I}_{\overline{\mathbf{w}}}}\left\{-\frac{\hat{w}_i - \overline{w}_i}{\Delta x_i \Delta s_i} \ : \ \Delta x_i\Delta s_i < 0\right\},$$

and from $w_i^+ < \hat{w}_i$ we get the upper bound

$$\alpha_4 := \min_{i\in\mathcal{I}_{\overline{\mathbf{w}}}}\left\{-\frac{\hat{w}_i - \overline{w}_i - \sqrt{(\hat{w}_i - \overline{w}_i)^2 + 4\Delta x_i\Delta s_i(\hat{w}_i - \overline{w}_i)}}{2\Delta x_i\Delta s_i} \ : \ \Delta x_i\Delta s_i > 0\right\}.$$

Similarly, if $i \in \mathcal{I}_{\hat{w}}$, then we have the following bounds on α:

$$\alpha_5 := \min_{i\in\mathcal{I}_{\hat{w}}}\left\{\frac{\overline{w}_i - \hat{w}_i - \sqrt{(\overline{w}_i - \hat{w}_i)^2 - 4\Delta x_i\Delta s_i(\overline{w}_i - \hat{w}_i)}}{2\Delta x_i\Delta s_i} \ : \ \Delta x_i\Delta s_i < 0\right\},$$

and

$$\alpha_6 := \min_{i\in\mathcal{I}_{\hat{w}}}\left\{-\frac{\hat{w}_i - \overline{w}_i}{\Delta x_i\Delta s_i} \ : \ \Delta x_i\Delta s_i > 0\right\}.$$

Now, if $\alpha \in (0, \alpha^*)$, where

$$\alpha^* := \min\{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, 1\}, \tag{3.6}$$

then the inequalities (3.4) and (3.5) are satisfied. ■

Proposition 3.3 ([38]) *If $M \in \mathcal{P}_*(\kappa)$ for any $\kappa \geq 0$ and $\mathcal{F}^+ \neq \emptyset$, then for every $\mathbf{w} > 0$ the level set $\mathcal{L}_{\mathbf{w}}$ is nonempty and compact.*

Proof. The compactness of $\mathcal{L}_{\mathbf{w}}$ has already been proved in Lemma 3.1. What is left is to show the nonemptyness of the level set $\mathcal{L}_{\mathbf{w}}$. Let us assume on the contrary that there exists a $\hat{\mathbf{w}} > 0$ such that $\mathcal{L}_{\hat{\mathbf{w}}} = \emptyset$. From assumption $\mathcal{F}^+ \neq \emptyset$ follows that there is a $\mathbf{w}^0 > 0$ such that $\mathcal{L}_{\mathbf{w}^0} \neq \emptyset$. We may choose $\mathbf{w}' > 0$ such that $\mathbf{w}^0 < \mathbf{w}'$ and $\hat{\mathbf{w}} < \mathbf{w}'$, as well. Then $\mathcal{L}_{\mathbf{w}'}$ is nonempty, furthermore it is compact by Lemma 3.1.

Let us define the function $f : \mathcal{L}_{\mathbf{w}'} \to \mathbb{R}^n_+$ as

$$f_i(\mathbf{x},\mathbf{s}) = \begin{cases} 0, & \text{if } w_i \le \hat{w}_i, \\ w_i - \hat{w}_i, & \text{otherwise}, \end{cases} \tag{3.7}$$

where $\mathbf{w} = \mathbf{x}\,\mathbf{s}$. Then $f_i(\mathbf{x},\mathbf{s})$ is continuous for all index i, thus $f(\mathbf{x},\mathbf{s})$ is continuous, too. From the continuity of the infinity norm follows the continuity of the composed function $\|f(\mathbf{x},\mathbf{s})\|_\infty$, as well. Now, using the nonemptyness and compactness of the set $\mathcal{L}_{\mathbf{w}'}$ and the continuity of the function $\|f(\mathbf{x},\mathbf{s})\|_\infty$ we have

$$\gamma = \|f(\overline{\mathbf{x}},\overline{\mathbf{s}})\|_\infty = \min_{(\mathbf{x},\mathbf{s})\in\mathcal{L}_{\mathbf{w}'}} \|f(\mathbf{x},\mathbf{s})\|_\infty \le \left\|f(\mathbf{x}^0,\mathbf{s}^0)\right\|_\infty .$$

Our assumption $\mathcal{L}_{\hat{\mathbf{w}}} = \emptyset$ implies that the optimal value of the previous problem $\gamma = \|f(\overline{\mathbf{x}},\overline{\mathbf{s}})\|_\infty > 0$. If $\operatorname{int}\mathcal{T}(\overline{\mathbf{w}},\hat{\mathbf{w}}) \neq \emptyset$, then we may immediately apply Lemma 3.2 to $\overline{\mathbf{w}}$ and $\hat{\mathbf{w}}$ getting a point $\mathbf{w}^+$ such that

$$\left\|f(\mathbf{x}^+,\mathbf{s}^+)\right\|_\infty < \|f(\overline{\mathbf{x}},\overline{\mathbf{s}})\|_\infty \tag{3.8}$$

which is a contradiction.

Otherwise, take any point $\tilde{\mathbf{w}} \in \mathcal{B}_{\frac{\gamma}{2}}(\hat{\mathbf{w}}) \cap \operatorname{int}\mathcal{T}(\mathbf{0},\hat{\mathbf{w}})$, where $\mathcal{B}_{\frac{\gamma}{2}}(\hat{\mathbf{w}}) = \{\mathbf{z} \,:\, \|\mathbf{z}-\hat{\mathbf{w}}\| \le \gamma/2\}$. We may now apply Lemma 3.2 to $\overline{\mathbf{w}}$ and $\tilde{\mathbf{w}}$ getting again a point $\mathbf{w}^+$ such that (3.8) holds, which is a contradiction. ■

If $\mathbf{w} = \mathbf{0}$, then from the definition of $\mathcal{L}_{\mathbf{w}}$ and $\mathcal{F}$ it is easy to show that

$$\mathcal{L}_{\mathbf{0}} = \{(\mathbf{x},\mathbf{s}) \in \mathcal{F} :\ \mathbf{x}\mathbf{s} = \mathbf{0}\} = \mathcal{F}^*.$$

However, Proposition 3.3 claims the compactness and nonemptyness of the sets $\mathcal{L}_{\mathbf{w}}$ only for positive vectors $\mathbf{w}$. In the following corollary it will be verified for $\mathbf{w} = \mathbf{0}$ as well.

Corollary 3.4 ([38]) *If $M \in \mathcal{P}_*(\kappa)$ for any $\kappa \ge 0$ and $\mathcal{F}^+ \neq \emptyset$, then $\mathcal{F}^*$ is nonempty and compact.*

Proof. It is easily seen that $\mathcal{F}^* \subset \mathcal{L}_{\mathbf{w}}$ for any $\mathbf{w} > \mathbf{0}$, thus $\mathcal{F}^*$ is bounded.

Let $\mathbf{w}^i \in \mathbb{R}^n,\ \mathbf{w}^i > \mathbf{0}$ be a sequence of vectors such that $\mathbf{w}^i > \mathbf{w}^{i+1}$ and $\lim_{i\to\infty} \mathbf{w}^i = \mathbf{0}$. Then

$$\mathcal{L}_{\mathbf{w}^{i+1}} \subset \mathcal{L}_{\mathbf{w}^i}$$

holds and we have

$$\mathcal{F}^* = \mathcal{L}_{\mathbf{0}} = \bigcap_{i=1}^{\infty} \mathcal{L}_{\mathbf{w}^i}.$$

By a well-known result of elementary topology (Theorem 2.24 (b), [71]) the set $\mathcal{F}^*$ is closed. The nonemptyness of $\mathcal{F}^*$ follows from Cantor's theorem (Corollary of Theorem 2.36, [71]). ■

Theorem 3.5 ([38]) *If $M \in \mathcal{P}_*(\kappa)$ for any $\kappa \geq 0$ and $\mathcal{F}^+ \neq \emptyset$, then for every $\hat{\mathbf{w}} > 0$ there exists a unique $(\hat{\mathbf{x}}, \hat{\mathbf{s}}) \in \mathcal{F}$ such that $\hat{\mathbf{w}} = \hat{\mathbf{x}}\,\hat{\mathbf{s}}$.*

Proof. Let us suppose on the contrary that there is no $(\hat{\mathbf{x}}, \hat{\mathbf{s}}) \in \mathcal{F}$ such that $\hat{\mathbf{x}}\,\hat{\mathbf{s}} = \hat{\mathbf{w}}$ for a given $\hat{\mathbf{w}} > 0$. Thus $\mathbf{x}^T\mathbf{s} < \mathbf{e}^T\hat{\mathbf{w}}$ holds for every $(\mathbf{x}, \mathbf{s}) \in \mathcal{L}_{\hat{\mathbf{w}}}$. Based on Proposition 3.3 and the continuity of the scalar product there exists a point $(\overline{\mathbf{x}}, \overline{\mathbf{s}}) \in \mathcal{L}_{\hat{\mathbf{w}}}$ such that

$$\overline{\mathbf{x}}^T\overline{\mathbf{s}} = \max_{(\mathbf{x},\mathbf{s})\in\mathcal{L}_{\hat{\mathbf{w}}}} \mathbf{x}^T\mathbf{s}.$$

Obviously, $\overline{\mathbf{w}} = \overline{\mathbf{x}}\,\overline{\mathbf{s}} \leq \hat{\mathbf{w}}$, but $\overline{\mathbf{w}} \neq \hat{\mathbf{w}}$ according to our assumption, so $\beta := \mathbf{e}^T(\hat{\mathbf{w}} - \overline{\mathbf{w}})$ is positive. Let $\tilde{\mathbf{w}}$ be a positive vector such that $\hat{\mathbf{w}} > \tilde{\mathbf{w}} > \hat{\mathbf{w}} - \frac{\beta}{n}\,\mathbf{e}$ and $int\,\mathcal{T}(\tilde{\mathbf{w}}, \overline{\mathbf{w}}) \neq \emptyset$. Now we may apply Lemma 3.2, namely there is $(\mathbf{x}^+, \mathbf{s}^+) \in \mathcal{L}_{\tilde{\mathbf{w}}} \subset \mathcal{L}_{\hat{\mathbf{w}}}$ and $\mathbf{x}^+\mathbf{s}^+ = \mathbf{w}^+ \in int\,\mathcal{T}(\overline{\mathbf{w}}, \tilde{\mathbf{w}})$. It can be shown that the inequality $\mathbf{e}^T\mathbf{w}^+ > \mathbf{e}^T\overline{\mathbf{w}}$ also holds, with a suitable choice of the step length α in the construction of the point $(\mathbf{x}^+, \mathbf{s}^+)$ (see proof of Lemma 3.2). An equivalent formulation of the required inequality is the following

$$\begin{aligned} \mathbf{0} \leq \mathbf{e}^T(\mathbf{w}^+ - \overline{\mathbf{w}}) &= \mathbf{e}^T\left(\overline{\mathbf{x}}\,\overline{\mathbf{s}} + \alpha(\overline{\mathbf{s}}\,\Delta\mathbf{x} + \overline{\mathbf{x}}\,\Delta\mathbf{s}) + \alpha^2\Delta\mathbf{x}\,\Delta\mathbf{s} - \overline{\mathbf{x}}\,\overline{\mathbf{s}}\right) \\ &= \alpha\left(\mathbf{e}^T\left(\tilde{\mathbf{w}} - \overline{\mathbf{w}}\right) + \alpha^2\Delta\mathbf{x}^T\Delta\mathbf{s}\right), \end{aligned}$$

which holds for all nonnegative α if $\Delta\mathbf{x}^T\Delta\mathbf{s} \geq 0$. Otherwise, i.e, if $\Delta\mathbf{x}^T\Delta\mathbf{s} \leq 0$, we get an additional upper bound on α:

$$\alpha \leq -\frac{\mathbf{e}^T\left(\tilde{\mathbf{w}} - \overline{\mathbf{w}}\right)}{\Delta\mathbf{x}^T\Delta\mathbf{s}}.$$

Taking it also into consideration in (3.6), the given step length α determines such an $(\mathbf{x}^+, \mathbf{s}^+)$, for which $(\mathbf{x}^+)^T\,\mathbf{s}^+ = \mathbf{e}^T\mathbf{w}^+ \geq \mathbf{e}^T\overline{\mathbf{w}} = \overline{\mathbf{x}}^T\overline{\mathbf{s}}$, so we get a contradiction with the choice of the point $(\overline{\mathbf{x}}, \overline{\mathbf{s}})$.

The uniqueness follows similarly as in [51] (Lemma 4.2). Assume that there are two distinct points $(\mathbf{x}^1, \mathbf{s}^1), (\mathbf{x}^2, \mathbf{s}^2) \in \mathcal{F}^+$, then we have

$$M(\mathbf{x}^1 - \mathbf{x}^2) = \mathbf{s}^1 - \mathbf{s}^2 \quad \text{and} \quad \mathbf{x}^1\,\mathbf{s}^1 = \mathbf{x}^2\,\mathbf{s}^2 = \hat{\mathbf{w}}.$$

Since the matrix M is $\mathcal{P}_*(\kappa)$, it is $\mathcal{P}_0$, as well. Therefore, by the second statement of Lemma 2.5, there is an index j such that

$$x_j^1 \neq x_j^2 \quad \text{and} \quad 0 \leq (x_j^1 - x_j^2)[M(\mathbf{x}^1 - \mathbf{x}^2)]_j = (x_j^1 - x_j^2)(s_j^1 - s_j^2).$$

We may assume, without loss of generality, that $x_j^1 > x_j^2$. Hence the inequality above implies that $s_j^1 \geq s_j^2$. This contradicts the equality $x_j^1\,s_j^1 = x_j^2\,s_j^2$. ■

3.2 The central path

The results above can be summarized in the following theorem.

Theorem 3.6 ([38]) *Let an LCP with a $\mathcal{P}_*(\kappa)$-matrix M be given. Then the following statements are equivalent:*

1. $\mathcal{F}^+ \neq \emptyset$,
2. $\forall\, \mathbf{w} \in \mathbb{R}^n_+ \quad \exists!\, (\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+ : \mathbf{x}\,\mathbf{s} = \mathbf{w}$,
3. $\forall\, \mu > 0 \quad \exists!\, (\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+ : \mathbf{x}\,\mathbf{s} = \mu\,\mathbf{e}$, *i.e., the central path $\mathcal{C}$ exists and it is unique.*

Proof. The implication 1. $\Rightarrow$ 2. follows from Theorem 3.5. We get the third statement from the second with a special choice $\mathbf{w} = \mu\,\mathbf{e}$. Finally, it is obvious, that if the third statement holds, then the first is true. ■

The previous theorem means that the solution of the problem CPP_μ exists and is unique for all positive μ, if there is an interior point of the LCP, namely, a point satisfying the first constraint of the CPP_μ and being positive. Furthermore, instead of $\mu\,\mathbf{e}$, we can write an arbitrary positive vector on the right hand side of the third constraint and the system will still have a unique solution.

Corollary 3.7 ([38]) *Let $M \in \mathcal{P}_*(\kappa)$ for any $\kappa \geq 0$. If $\mathcal{F}^+ \neq \emptyset$ then for $\mu_0 > 0$ the following set*

$$\mathcal{C}_{\mu_0} := \{(\mathbf{x}(\mu), \mathbf{s}(\mu)) \in \mathcal{F} \mid \mathbf{x}(\mu)\mathbf{s}(\mu) = \mu\,\mathbf{e},\ 0 < \mu \leq \mu_0\}$$

is nonempty and compact. Furthermore, $\mathcal{C}_{\mu_0} \subset \mathcal{L}_{\mu_0\mathbf{e}}$ holds.

Let us recall the definition of the following index sets:

$$\begin{aligned} \mathcal{B} &:= \{i \in \mathcal{I} \,:\, x_i > 0 \text{ for some } (\mathbf{x}, \mathbf{s}) \in \mathcal{F}^*\}, \\ \mathcal{N} &:= \{i \in \mathcal{I} \,:\, s_i > 0 \text{ for some } (\mathbf{x}, \mathbf{s}) \in \mathcal{F}^*\}, \\ \mathcal{T} &:= \{i \in \mathcal{I} \,:\, x_i = s_i = 0 \text{ for all } (\mathbf{x}, \mathbf{s}) \in \mathcal{F}^*\}. \end{aligned}$$

Lemma 3.8 ([36]) *Let the matrix M be column sufficient. Then the index sets $\mathcal{B}$,$\mathcal{N}$ and $\mathcal{T}$ form a partition of the index set $\mathcal{I}$.*

An $(\mathbf{x}, \mathbf{s}) \in \mathcal{F}^*$ is called *maximally complementary*, if $\mathbf{x}_\mathcal{B} > \mathbf{0}$ and $\mathbf{s}_\mathcal{N} > \mathbf{0}$. If the matrix of the LCP is column sufficient, then the solution set $\mathcal{F}^*$ is convex and polyhedral by Proposition 2.32. Therefore, a maximally complementary solution exists.

The next result generalizes Theorem I.7. of [70], and shows that the central path converges to a maximally complementary solution of the LCP.

Theorem 3.9 ([38]) *If $M \in \mathcal{P}_*(\kappa)$ for any $\kappa \geq 0$ and $\mathcal{F}^+ \neq \emptyset$ then there exists $(\mathbf{x}^*, \mathbf{s}^*)$ such that*

(i) $(\mathbf{x}^*, \mathbf{s}^*) = \lim_{\mu \to 0} (\mathbf{x}(\mu), \mathbf{s}(\mu))$,

(ii) $(\mathbf{x}^*, \mathbf{s}^*) \in \mathcal{F}^*$,

(iii) $(\mathbf{x}^*, \mathbf{s}^*)$ *is a maximally complementary solution.*

Proof. Let $\{\mu_k\}_{k=0}^{\infty}$ be a positive sequence such that $\mu_k \to 0$ if $k \to \infty$. By Corollary 3.7, the set $\{(\mathbf{x}(\mu_k), \mathbf{s}(\mu_k))\} \subset \mathcal{C}_{\mu_0}$ is compact and hence contains a subsequence converging to a point $(\mathbf{x}^*, \mathbf{s}^*) \in \mathcal{L}_{\mu_0 \mathbf{e}}$. Since $\mathbf{x}(\mu_k)\, \mathbf{s}(\mu_k) = \mu_k \mathbf{e} \to \mathbf{0}$, we have $\mathbf{x}^*\, \mathbf{s}^* = \mathbf{0}$, so $(\mathbf{x}^*, \mathbf{s}^*) \in \mathcal{F}^*$. We claim that $(\mathbf{x}^*, \mathbf{s}^*)$ is a maximally complementary solution.

Let us denote by $(\overline{\mathbf{x}}, \overline{\mathbf{s}})$ a maximally complementary solution. From (3.2) we have that

$$(\mathbf{x}(\mu_k) - \overline{\mathbf{x}})^T (\mathbf{s}(\mu_k) - \overline{\mathbf{s}}) \geq -4\, \kappa\, n\, \mu_k.$$

This gives

$$(1 + 4\kappa)\, n\, \mu_k \geq \mathbf{x}(\mu_k)^T \overline{\mathbf{s}} + \mathbf{s}(\mu_k)^T \overline{\mathbf{x}}, \tag{3.9}$$

where we used that $\overline{\mathbf{x}}\, \overline{\mathbf{s}} = \mathbf{0}$. Noting that $\mathbf{x}(\mu_k)^T \mathbf{s}(\mu_k) = n\, \mu_k$, we deduce that

$$\sum_{j \in \mathcal{B}} \overline{x}_j s_j(\mu_k) + \sum_{j \in \mathcal{N}} x_j(\mu_k) \overline{s}_j \leq (1 + 4\kappa)\, n\, \mu_k.$$

Dividing both sides by μ_k and recalling that $x_j(\mu_k) s_j(\mu_k) = \mu_k$, we obtain

$$\sum_{j \in \mathcal{B}} \frac{\overline{x}_j}{x_j(\mu_k)} + \sum_{j \in \mathcal{N}} \frac{\overline{s}_j}{s_j(\mu_k)} \leq (1 + 4\kappa)\, n.$$

Letting $k \to \infty$, we see that

$$\sum_{j \in B} \frac{\overline{x}_j}{x_j^*} + \sum_{j \in N} \frac{\overline{s}_j}{s_j^*} \leq (1 + 4\kappa)\, n.$$

The previous inequality means that all x_j^*, $j \in \mathcal{B}$ and s_j^*, $j \in \mathcal{N}$ are nonzeros. We know that $(\mathbf{x}^*, \mathbf{s}^*) \geq \mathbf{0}$, therefore $\mathbf{x}_{\mathcal{B}}^* > \mathbf{0}$ and $\mathbf{s}_{\mathcal{N}}^* > \mathbf{0}$ follows. Then the solution $(\mathbf{x}^*, \mathbf{s}^*)$ is maximally complementary. ■

3.3 The Newton system

The CPP_μ is still a nonlinear programming problem, it has a similar difficulty as the original problem the LCP. Therefore, we only solve it approximately, using one step of the well known

Newton method. Suppose $(\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+$ and seek the unique solution $(\hat{\mathbf{x}}(\mu), \hat{\mathbf{s}}(\mu))$ of the CCP_μ with respect to μ in the form $\hat{\mathbf{x}}(\mu) = \mathbf{x}+\Delta\mathbf{x}$, $\hat{\mathbf{s}}(\mu) = \mathbf{s}+\Delta\mathbf{s}$. After substituting and neglecting the quadratic term $\Delta\mathbf{x}\,\Delta\mathbf{s}$, we obtain the following Newton system:

$$\left.\begin{array}{rcl} -M\Delta\mathbf{x} + \Delta\mathbf{s} &=& \mathbf{0} \\ S\Delta\mathbf{x} + X\Delta\mathbf{s} &=& \mu\,\mathbf{e} - \mathbf{x}\,\mathbf{s} \end{array}\right\} \tag{3.10}$$

The positivity condition of $(\hat{\mathbf{x}}(\mu), \hat{\mathbf{s}}(\mu))$ is ensured with the right choice of the Newton-step length.

The next statement is used to guarantee the existence and uniqueness of *Newton directions* that are the solutions of the system (3.11) for various values of vector $\mathbf{a} \in \mathbb{R}^n$, where $\mathbf{a}$ depends on the particular interior point algorithm. This statement is a direct corollary of the last statement in Lemma 2.5.

Corollary 3.10 *Let $M \in \mathbb{R}^{n\times n}$ be a $\mathcal{P}_0$-matrix, $\mathbf{x}, \mathbf{s} \in \mathbb{R}^n_+$. Then for all $\mathbf{a} \in \mathbb{R}^n$ the system*

$$\begin{array}{rcrcl} -M\Delta\mathbf{x} &+& \Delta\mathbf{s} &=& \mathbf{0} \\ \mathbf{s}\,\Delta\mathbf{x} &+& \mathbf{x}\,\Delta\mathbf{s} &=& \mathbf{a} \end{array} \tag{3.11}$$

has a unique solution $(\Delta\mathbf{x}, \Delta\mathbf{s})$.

Let the current point be $(\mathbf{x}, \mathbf{s})$ and $(\Delta\mathbf{x}, \Delta\mathbf{s})$ be the current Newton direction.[1] The new point with step length θ is given by $(\mathbf{x}(\theta), \mathbf{s}(\theta)) = (\mathbf{x} + \theta\Delta\mathbf{x}, \mathbf{s} + \theta\Delta\mathbf{s})$.

3.4 The proximity measures and the neighbourhoods of the central path

Since we can not take exact steps on the central path due to computational difficulties, it will only be used as a guide line, and some kind of deviation from it will be allowed. However, for proving convergence we need to control this variance. This is the reason why the proximity measure and the neighbourhood of the central path are introduced, two important components of interior point algorithms. There are several types of these in the literature of interior point methods (see for example book [70]).

The basic requirement on a proximity measure is the following: it has to take its minimum values in the points of the central path. In most cases the minimum is zero, but for example for the proximity measure

$$\delta_a(\mathbf{x}\,\mathbf{s}) = \frac{\max(\sqrt{\mathbf{x}\,\mathbf{s}}\,)}{\min(\sqrt{\mathbf{x}\,\mathbf{s}}\,)}$$

[1] Generally, the Newton direction is the unique solution of the system (3.11). We will discuss how to define the actual Newton directions for various algorithms in Chapter 5 and 6.

it is one. The idea behind this measure is that if the point $(\mathbf{x}, \mathbf{s})$ is on the central path, then the vector $\mathbf{x}\,\mathbf{s}$ has equal coordinates. We will use this measure in the affine scaling algorithm in Chapter 6. However, this type of proximity measure does not always meet our requirements for all purposes, as it takes only the maximal and the minimal coordinates of the product vector into consideration. The next measure was introduced by Kojima et al. [52]. It depends on every coordinate and it is zero in the points on the central path. It reflects our aims better since it measures the distance of the actual point and the central path.

$$\delta_0(\mathbf{x}\,\mathbf{s}, \mu) := \frac{1}{2} \left\| \mathbf{e} - \frac{\mathbf{x}\,\mathbf{s}}{\mu} \right\|.$$

One of the most frequently used proximity measures was introduced by Jansen [46], and it is defined as follows

$$\delta_c(\mathbf{x}\,\mathbf{s}, \mu) = \left\| \sqrt{\frac{\mathbf{x}\,\mathbf{s}}{\mu}} - \sqrt{\frac{\mu}{\mathbf{x}\,\mathbf{s}}} \right\|.$$

In addition, this measure penalizes the points approaching the boundary, because if some coordinates of $\mathbf{x}$ or $\mathbf{s}$ tend either to 0 or infinity, then the proximity measure approaches infinity. (Only the half of this property is satisfied by measure δ_0.)

In some sense, the generalization of δ_c is the family of proximity measures defined by self-regular functions introduced by Peng, Roos and Terlaky [63].

A $\psi(t) : \mathbb{R}_+ \to \mathbb{R}$ function is *self-regular* if it is a twice differentiable function and satisfies the following two conditions:

SR1 The function $\psi(t)$ is strictly convex respect to $t > 0$, its global minimum point is $t = 1$, where $\psi(1) = \psi'(1) = 0$. Furthermore, there are positive constants $\nu_2 \geq \nu_1 > 0$ and $p, q \geq 1$ such that for all $t > 0$

$$\nu_1(t^{p-1} + t^{-1-q}) \leq \psi''(t) \leq \nu_2(t^{p-1} + t^{-1-q}).$$

SR2 For any $t_1, t_2 > 0$ and $r \in [0, 1]$

$$\psi(t_1^r t_2^{1-r}) \leq r\psi(t_1) + (1 - r)\psi(t_2).$$

The parameter q is called the barrier degree and p the growth degree of the function $\psi(t)$, respectively. The first condition demands the function $\psi(t)$ to have an appropriate growing speed near zero and infinity. The second condition is like the exponential convexity property.

There are two special families of self-regular functions:

$$\Upsilon_{p,q}(t) = \frac{t^{p+1} - 1}{p(p+1)} + \frac{t^{1-q} - 1}{q(q-1)} + \frac{p-q}{pq}(t-1), \quad q > 1,$$

where $\nu_1 = \nu_2 = 1$ and

$$\Gamma_{p,q}(t) = \frac{t^{p+1}-1}{p+1} + \frac{t^{1-q}-1}{q-1}, \quad q > 1,$$

with $\nu_1 = \min(p,q)$ and $\nu_2 = \max(p,q)$.

Now we can define the following proximity measure:

$$\Psi(\mathbf{x}\,\mathbf{s}, \mu) := \sum_{i=1}^{n} \psi\left(\sqrt{\frac{x_i s_i}{\mu}}\right),$$

where ψ is a self-regular function.

The theoretical importance of the self-regular function lies in the fact that a better theoretical complexity of the long step interior point algorithms can be proved using this proximity measure [63].

A general form of the *neighbourhood*:

$$\mathcal{N}(\tau) := \left\{(\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+ : \ \delta(\mathbf{x}\,\mathbf{s}, \mu) \leq \tau\right\},$$

where $\delta(\mathbf{x}\,\mathbf{s})$ is a proximity measure and τ is the proximity parameter.
We mention here only a few of the well-known neighbourhoods:

- $\mathcal{N}_0(\tau) := \{(\mathbf{x},\, \mathbf{s}) \in \mathcal{F}^+ : \ \|\mathbf{x}\,\mathbf{s} - \mu\mathbf{e}\| \leq \tau\mu\}$,
- $\mathcal{N}_c(\tau) := \left\{(\mathbf{x},\, \mathbf{s}) \in \mathcal{F}^+ : \ \left\|\sqrt{\frac{\mathbf{x}\mathbf{s}}{\mu}} - \sqrt{\frac{\mu}{\mathbf{x}\mathbf{s}}}\right\| \leq \tau\right\}$,
- The negative infinity neighbourhood, which is defined for example by Potra in [66]. It is considered to be a 'wide neighbourhood'.

$$\mathcal{N}_\infty^-(\beta) := \left\{(\mathbf{x},\, \mathbf{s}) \in \mathcal{F}^+ : \ \mathbf{x}\,\mathbf{s} \geq \beta\frac{\mathbf{x}^T\mathbf{s}}{n}\right\},$$

3.5 Scaling and the sketch of the algorithm

The aim of scaling is to reformulate the Newton system to a simpler form to emphasize the structure of the problem. The following notations are introduced for *scaling*:

$$\mathbf{v} = \sqrt{\frac{\mathbf{x}\mathbf{s}}{\mu}}, \quad \mathbf{d} = \sqrt{\frac{\mathbf{x}}{\mathbf{s}}}, \quad \mathbf{d}_x = \frac{\mathbf{v}\Delta\mathbf{x}}{\mathbf{x}}, \quad \mathbf{d}_s = \frac{\mathbf{v}\Delta\mathbf{s}}{\mathbf{s}}, \quad \mathbf{r} = \mathbf{d}_x\mathbf{d}_s, \quad \mathbf{p} = \mathbf{d}_x + \mathbf{d}_s, \qquad (3.12)$$

where we have $\mu \equiv 1$ in the affine scaling algorithm. Accordingly, the proximity measure on the scaled space is $\delta(\mathbf{v}) = \delta(\mathbf{x}\,\mathbf{s}, \mu)$, so for example $\delta_c(\mathbf{v}) = \|\mathbf{v} - \mathbf{v}^{-1}\|$.

After rescaling the Newton system we have

$$\left.\begin{array}{rcl} -\bar{M}\mathbf{d}_x + \mathbf{d}_s &=& \mathbf{0} \\ \mathbf{d}_x + \mathbf{d}_s &=& \mathbf{v}^{-1} - \mathbf{v} \end{array}\right\}$$

where $\bar{M} = DMD$ and $D = diag(\mathbf{d})$. It has the following solution if the matrix $I + \bar{M}$ is regular[2]

$$\mathbf{d}_x = -(I + \bar{M})^{-1}\mathbf{v} \quad \text{and} \quad \mathbf{d}_s = -\bar{M}(I + \bar{M})^{-1}\mathbf{v}.$$

To summarize the above mentioned ideas, the general sketch of the interior point methods is the following. While the *duality gap*, namely $\mathbf{x}^T\mathbf{s}$ is too large, we take steps – i.e., we solve the Newton system and determine the step length θ such that the new point will be in the appropriate neighbourhood of the central path $\mathcal{N}$. Let us note that the positivity of the point is built into the definition of the neighbourhood.

Input:

an accuracy parameter $\varepsilon > 0$;
an initial point $(\mathbf{x}^0, \mathbf{s}^0, \mu^0) \in \mathcal{N}$,

begin
$\mathbf{x} := \mathbf{x}^0$, $\mathbf{s} := \mathbf{s}^0$, $\mu := \mu^0$
while $\mathbf{x}^T\mathbf{s} \geq \varepsilon$ **do**
update the centrality parameter μ;
calculate the Newton directions by solving (3.10);
determine the step length θ such that $(\mathbf{x}(\theta), \mathbf{s}(\theta)) \in \mathcal{N}$;
$\mathbf{x} = \mathbf{x} + \theta\Delta\mathbf{x}$, $\mathbf{s} = \mathbf{s} + \theta\Delta\mathbf{s}$;
end
end.

The family of interior point algorithms is very diverse. There are three main factors for freedom in the algorithms:

1. *The updating strategy of the centrality parameter μ.* In this aspect, there are three main types. The long step, the short step and the adaptive algorithms. In both cases the new value of μ is $\mu = (1-\gamma)\mu$, where $\gamma \in (0, 1]$. But γ is independent of the dimension

[2]If the matrix M is a $\mathcal{P}_0$-matrix, then $\bar{M}$ also belongs to $\mathcal{P}_0$, furthermore the matrix $I + \bar{M}$ is a $\mathcal{P}$-matrix by the seventh statement of Lemma 2.5, and a $\mathcal{P}$-matrix is regular by definition (its determinant is positive, so not zero). Therefore, if $M \in \mathcal{P}_0$, then matrix $I + \bar{M}$ is certainly regular.

of the problem n at the long step algorithm (in practice, in a linear programming case a usual choice is $\gamma = 0.99$), on the contrary, in the short step case it depends on n, for example $\gamma = 1/(2\sqrt{n})$, and in the last case γ is not constant, but depends on the actual point.

2. *The right hand side vector* $\mathbf{a}$ *of the Newton system (3.11).* We can eagerly choose $\mu = 0$ in (3.10), namely $\mathbf{a} = -\mathbf{x}\,\mathbf{s}$ in (3.11). This is called the affine scaling or the predictor step. If $\mathbf{a} = \mu\,\mathbf{e} - \mathbf{x}\,\mathbf{s}$ we talk about a centering or a corrector step. Predictor-corrector algorithms take predictor and corrector steps alternatively.

3. *The choice of the central path neighbourhood* $\mathcal{N}$. (See a few variants above.)

3.6 Estimates of the Newton directions

When the Newton directions $\Delta\mathbf{x}$ and $\Delta\mathbf{s}$ are determined, we make a reduction to a linear system, the quadratic term $\Delta\mathbf{x}\,\Delta\mathbf{s}$ is neglected. However, this term has to be considered at the iteration complexity analysis of interior point methods, so we need to estimate it.

The next lemma is crucial for the analysis of the interior point algorithm. It shows, that if M is a $\mathcal{P}_*(\kappa)$-matrix, then the scalar product of the Newton directions, $\Delta\mathbf{x}^T\Delta\mathbf{s}$ can be bounded from below.

Lemma 3.11 ([51]) *Let $\kappa \geq 0$. The following three statements are equivalent:*

1. $M \in \mathcal{P}_*(\kappa)$.

2. *For every positive diagonal matrix D and every $\xi, \nu, h \in \mathbb{R}^n$, the relations*
$$D^{-1}\xi + D\eta = h,$$
$$-M\xi + \eta = 0$$
always imply
$$\xi^T\eta \geqq -\kappa\|h\|_2^2$$

3. *For every $\xi \in \mathbb{R}^n$ it is*
$$\xi^T M\xi \geqq -\kappa \inf_D \left\|D^{-1}\xi + DM\xi\right\|_2^2,$$
where the infimum is taken over all positive diagonal matrices D.

The following estimations for the Newton directions are used in the complexity analysis of interior point methods. The next lemma is proved by Potra in [65].

Lemma 3.12 *Let M be an arbitrary $n \times n$ real matrix and $(\Delta\mathbf{x}, \Delta\mathbf{s})$ be a solution of the system (3.11). Then*

$$\sum_{i \in \mathcal{I}_+} \Delta x_i \Delta s_i \leq \frac{1}{4} \left\| \frac{\mathbf{a}}{\sqrt{\mathbf{xs}}} \right\|^2 .$$

The proof of the first statement in the next lemma is similar to the proof of Lemma 5.1 by Illés, Roos and Terlaky [37]. The second estimation follows from the previous lemma by using some properties of $\mathcal{P}_*(\kappa)$-matrices, and the last estimation is a corollary of the first and second statements using some properties of norms.

Lemma 3.13 *Let the matrix M be a $\mathcal{P}_*(\kappa)$-matrix and $\mathbf{x}, \mathbf{s} \in \mathbb{R}^n_+$, $\mathbf{a} \in \mathbb{R}^n$. Let $(\Delta\mathbf{x}, \Delta\mathbf{s})$ be the solution of the system (3.11). Then*

$$\|\Delta\mathbf{x}\Delta\mathbf{s}\|_\infty \leq \left(\frac{1}{4} + \kappa\right) \left\| \frac{\mathbf{a}}{\sqrt{\mathbf{xs}}} \right\|^2, \qquad \|\Delta\mathbf{x}\Delta\mathbf{s}\|_1 \leq \left(\frac{1}{2} + \kappa\right) \left\| \frac{\mathbf{a}}{\sqrt{\mathbf{xs}}} \right\|^2,$$

$$\|\Delta\mathbf{x}\Delta\mathbf{s}\|_2 \leq \sqrt{\left(\frac{1}{4} + \kappa\right)\left(\frac{1}{2} + \kappa\right)} \left\| \frac{\mathbf{a}}{\sqrt{\mathbf{xs}}} \right\|^2 .$$

3.7 An embedding model for LCPs

In this section we deal with a technique that allows us to handle the initialization problem of interior point methods for LCPs, i.e., how to ensure that $\mathcal{F}^+$ is nonempty and how to get a well centered initial interior point. The embedding model discussed in this section was introduced by Kojima et al. [51]. The following lemma plays a crucial role in this model.

Lemma 3.14 (Lemma 5.3 in [51]) *Let M be a real matrix. The matrix $M' = \begin{pmatrix} M & I \\ -I & O \end{pmatrix}$ belongs to the class $\mathcal{P}_0$, column sufficient, $\mathcal{P}_*$, $\mathcal{P}_*(\kappa)$, positive semidefinite or skew-symmetric, if and only if M belongs to the same matrix class, respectively.*

Let us consider the LCP as given by (1.1). We assume that all the entries of matrix M and vector $\mathbf{q}$ are integral. The input length L of the LCP is defined as

$$L = \sum_{i=1}^{n} \sum_{j=1}^{n} \log_2(|m_{ij} + 1|) + \sum_{i=1}^{n} \log_2(|q_i| + 1) + 2 \log_2 n,$$

and let

$$\tilde{\mathbf{q}} = \frac{2^{L+1}}{n^2}\, \mathbf{e}.$$

The *embedding problem* of Kojima et al. [51] is as follows:

$$\left.\begin{array}{rcl} -M'\mathbf{x}' + \mathbf{s}' & = & \mathbf{q}' \\ \mathbf{x}'\mathbf{s}' & = & \mathbf{0} \\ \mathbf{x}', \mathbf{s}' & \geq & \mathbf{0} \end{array}\right\} \qquad \textit{(LCP')}$$

where

$$\mathbf{x}' = \begin{pmatrix} \mathbf{x} \\ \tilde{\mathbf{x}} \end{pmatrix}, \ \mathbf{s}' = \begin{pmatrix} \mathbf{s} \\ \tilde{\mathbf{s}} \end{pmatrix}, \ \mathbf{q}' = \begin{pmatrix} \mathbf{q} \\ \tilde{\mathbf{q}} \end{pmatrix}, \ M' = \begin{pmatrix} M & I \\ -I & O \end{pmatrix}.$$

An initial interior point for the embedding model (LCP') is readily available:

$$\mathbf{x} = \frac{2^L}{n^2}\mathbf{e}, \quad \tilde{\mathbf{x}} = \frac{2^{2L}}{n^3}\mathbf{e}, \quad \mathbf{s} = \frac{2^L}{n^2}M\mathbf{e} + \frac{2^{2L}}{n^3}\mathbf{e} + \mathbf{q}, \quad \tilde{\mathbf{s}} = \frac{2^L}{n^2}\mathbf{e}.$$

The following lemma indicates the connection between the solutions of the embedding problem and the solutions of the original LCP.

Lemma 3.15 (Lemma 5.4 in [51]) *Let* $(\mathbf{x}', \mathbf{s}') = \left(\begin{pmatrix} \mathbf{x} \\ \tilde{\mathbf{x}} \end{pmatrix}, \begin{pmatrix} \mathbf{s} \\ \tilde{\mathbf{s}} \end{pmatrix}\right)$ *be a solution of the problem (LCP').*

(i) If $\tilde{\mathbf{x}} = \mathbf{0}$*, then* $(\mathbf{x}, \mathbf{s})$ *is a solution of the original LCP.*

(ii) If M *is column sufficient and* $\tilde{\mathbf{x}} \neq \mathbf{0}$*, then the original LCP has no solution.*

We will return to Lemma 3.15 in Chapter 6, where interior point algorithms for LCPs with arbitrary matrices will be discussed.

Chapter 4

The dual linear complementarity problem

Let us now consider the *dual linear complementarity problem* (DLCP) [16]: find vectors $\mathbf{u}, \mathbf{z} \in \mathbb{R}^n$ which satisfy the constraints

$$\mathbf{u} + M^T\mathbf{z} = \mathbf{0}, \quad \mathbf{q}^T\mathbf{z} = -1, \quad \mathbf{u}\,\mathbf{z} = \mathbf{0}, \quad \mathbf{u}, \mathbf{z} \geq \mathbf{0}. \tag{4.1}$$

The set of feasible solutions of the DLCP is

$$\mathcal{F}_D = \{(\mathbf{u}, \mathbf{z}) \geq \mathbf{0} \ : \ \mathbf{u} + M^T\mathbf{z} = \mathbf{0}, \ \mathbf{q}^T\mathbf{z} = -1\}.$$

We show that the dual LCP can be solved in polynomial time if the matrix is row sufficient, as in this case all feasible solutions are complementary (see Lemma 4.3). This result yields an improvement compared to earlier known polynomial time complexity results, according to which an LCP is solvable in polynomial time for $\mathcal{P}_*(\kappa)$-matrices with known $\kappa \geq 0$. Due to the special structure of the DLCP, the polynomial time complexity of interior point methods depends on the row sufficiency property of the coefficient matrix M. Furthermore, we present an EP theorem for the dual LCP with arbitrary matrix M. The results of this chapter have been published in paper [43].

The concept of EP (existentially polynomial-time) theorems was introduced by Cameron and Edmonds [5]. It is a theorem of the form:

$$[\forall x : F_1(x), F_2(x), \ldots, F_k(x)],$$

where $F_i(x)$ is a predicate formula which has the form

$$F_i(x) = [\exists y_i \text{ such that } \|y_i\| \leq \|x\|^{n_i} \text{ and } f_i(x, y_i)].$$

Here $n_i \in \mathbb{Z}^+$, $\|z\|$ denotes the encoding length of z and $f_i(x, y_i)$ is a predicate for which there is a polynomial-size certificate.

The LCP duality theorem in an EP form was given by Fukuda, Namiki and Tamura [25]:

Theorem 4.1 *Let a matrix $M \in \mathbb{Q}^{n\times n}$ and a vector $q \in \mathbb{Q}^n$ be given. Then at least one of the following statements holds:*

(i) *the LCP has a complementary feasible solution (x, s), whose encoding size is polynomially bounded.*

(ii) *the DLCP has a complementary feasible solution (u, z), whose encoding size is polynomially bounded.*

(iii) *the matrix M is not sufficient and there is a certificate whose encoding size is polynomially bounded.*

Fukuda and Terlaky [26] proved a fundamental theorem for quadratic programming in oriented matroids. As they stated in their paper, the LCP duality theorem follows from Theorem 4.5 of [26] for sufficient matrix LCPs.

Theorem 4.2 *Let a sufficient matrix $M \in \mathbb{Q}^{n\times n}$ and a vector $\mathbf{q} \in \mathbb{Q}^n$ be given. Then exactly one of the following statements holds:*

(i) *the LCP has a solution $(\mathbf{x}, \mathbf{s})$, whose encoding size is polynomially bounded.*

(ii) *the DLCP has a solution $(\mathbf{u}, \mathbf{v})$, whose encoding size is polynomially bounded.*

A direct and constructive proof of the LCP duality theorem can be found in [16].

Now we show that if the matrix is row sufficient, then all feasible solutions of the DLCP are not only nonnegative, but complementary as well. Based on this result we get an EP theorem for the DLCP.

Lemma 4.3 *Let the matrix M be row sufficient. If $(\mathbf{u}, \mathbf{z}) \in \mathcal{F}_D$, then $(\mathbf{u}, \mathbf{z})$ is a solution to the DLCP.*

Proof. The vector $(\mathbf{u}, \mathbf{z})$ is a feasible solution to the DLCP, therefore $\mathbf{u}, \mathbf{z} \geq \mathbf{0}$ and $\mathbf{u} = -M^T\mathbf{z}$, so the complementarity gap is nonnegative

$$\mathbf{0} \leq \mathbf{u}\,\mathbf{z} = -\mathbf{z}\,M^T\mathbf{z} = -ZM^T\mathbf{z}.$$

From here, and by Definition 2.9, if the matrix M is a row sufficient matrix, then $ZM^T\mathbf{z} = \mathbf{0}$, thus $\mathbf{u}\,\mathbf{z} = \mathbf{0}$. ■

Corollary 4.4 *Let the matrix M be row sufficient. Then the DLCP can be solved in polynomial time.*

Proof. By Lemma 4.3, if M is row sufficient, one only needs to solve the feasibility problem of the DLCP, that is one needs to solve only a linear feasibility problem what can be solved in polynomial time, e.g. by interior point methods [70]. ■

We have already mentioned that there is no known polynomial time algorithm to check whether a matrix is row sufficient or not. The following theorem presents the properties which can be proved about an LCP problem with an arbitrary matrix using a polynomial time algorithm.

Theorem 4.5 *Let the matrix $M \in \mathbb{Q}^{n\times n}$ and the vector $\mathbf{q} \in \mathbb{Q}^n$ be given. Then it can be shown in polynomial time that at least one of the following statements holds:*

(i) *the DLCP has a feasible complementary solution $(\mathbf{u}, \mathbf{z})$, whose encoding size is polynomially bounded.*

(ii) *the LCP has a feasible solution, whose encoding size is polynomially bounded.*

(iii) *the matrix M is not row sufficient and there is a certificate whose encoding size is polynomially bounded.*

Proof. Apply a polynomial time algorithm to solve the feasibility problem of the DLCP, i.e., to find a point in the set $\mathcal{F}_D$. This is a linear feasibility problem, thus it can be solved in polynomial time with e.g. interior point methods using the self-dual embedding technique (see [70]). If $\mathcal{F}_D = \emptyset$, then by the Farkas Lemma $\mathcal{F}_P \neq \emptyset$, and a primal feasible point can be read out from the solution to the embedding problem, thus we get the second case in polynomial time as well. Otherwise, we get a point $(\mathbf{u}, \mathbf{z}) \in \mathcal{F}_D$. If the complementarity condition $\mathbf{uz} = \mathbf{0}$ holds, too, then the point $(\mathbf{u}, \mathbf{z})$ is a solution to the DLCP, so we get the first case. Finally, if the feasible solution $(\mathbf{u}, \mathbf{z})$ is not complementary, then according to Lemma 4.3 vector $\mathbf{z}$ provides a certificate that the matrix M is not a row sufficient matrix. As the encoding size of the solution to the self-dual embedding problem – after a proper rounding procedure – is polynomially bounded, the third option holds. ■

Observe that Theorem 4.5 is in an EP form. Both Theorems 4.1 and 4.5 deal with the LCP, but Theorem 4.1 approaches the problem from the primal, while Theorem 4.5 from the dual side. The advantage of Theorem 4.5 is that we determine certificates in polynomial time. The proof of Theorem 4.1 is constructive, too, it is based on the criss-cross algorithm (for details see [16, 25]). In the first two cases the LCP duality theorem gives not only a

feasible, but also complementary solutions. We deal with the second case of Theorem 4.5 in Chapter 6, where some modified interior point methods are presented which either solve the LCP with a given arbitrary matrix, or provide a polynomial size certificate in polynomial time, that the matrix of the problem is not $\mathcal{P}_*(\kappa)$ with an a priori given, but an arbitrary large κ.

Chapter 5

Linear complementarity problems for sufficient matrices

We analyze a version of the Mizuno–Todd–Ye predictor-corrector interior point algorithm for LCPs for $\mathcal{P}_*(\kappa)$-matrices in this chapter. We assume that there is a strictly positive feasible solution, namely that the interior point assumption holds.

One of the first versions of the predictor-corrector interior point algorithms for linear programming problems was introduced by Sonnevend, Stoer and Zhao [75]. This algorithm needs more corrector steps after each predictor step in order to return to the appropriate neighbourhood of the central path. Mizuno, Todd and Ye [56] published such a predictor-corrector interior point method for the linear programming problem in which each predictor step is followed by a single corrector step and whose iteration complexity is the best known in the linear programming literature. Anstreicher and Ye [87] extended this result to LCPs with a positive semidefinite matrix with the same iteration complexity.

In one of the best papers on interior point algorithms, Kojima, Mizuno and Yoshise [52] offered a polynomial primal-dual interior point method for positive semidefinite matrix LCPs. The properties of a more general matrix class can be formulated in a natural way from the iteration complexity analysis of their algorithm. As we have already mentioned in Chapter 2, this class is called $\mathcal{P}_*(\kappa)$ matrix class by Kojima et al. [51]. The primal-dual interior point algorithm of Kojima, Mizuno and Yoshise is generalized to $\mathcal{P}_*(\kappa)$-matrices. This algorithm is also polynomial if such a $\hat{\kappa} \geq 0$ is known, that the matrix of the problem is $\mathcal{P}_*(\hat{\kappa})$-matrix. The iteration complexity is a polynomial of $\hat{\kappa}$, the dimension n and the bit length L of the problem.

The natural outcome of the book on interior point methods for LCPs by Kojima et al. [51] was the emergence of different interior point algorithms for $\mathcal{P}_*(\kappa)$-matrix LCPs in the mid' 90s.

Several variants of the Mizuno–Todd–Ye type predictor-corrector interior point algorithm are known in the literature. First Miao [55], later Potra and Sheng [67] gave a generalization of Mizuno–Todd–Ye predictor-corrector algorithm for LCPs with a $\mathcal{P}_*(\kappa)$-matrix assuming the existence of a strictly positive solution. Miao updated the central path parameter μ in such a way that the equality $\mathbf{x}^T\mathbf{s}/n = \mu$ holds throughout. Therefore, the updating of μ is more complicated than in the skew-symmetric case, where $\mu' = (1-\alpha)\mu$ and θ is the length of the Newton step in the predictor phase. Further generalization has been established: Ji, Potra and Sheng [47] extended the algorithm to the infeasible LCP, Potra and Sheng [68, 69] to the infeasible and degenerate problem. In these methods the parameter μ is updated by $\mu' = (1-\theta)\mu$, thus generally $\mathbf{x}^T\mathbf{s}/n \neq \mu$ in their cases.

The Mizuno–Todd–Ye type predictor-corrector algorithm for the skew-symmetric or positive semidefinite LCP (horizontal linear complementarity problem, HLCP) of [65] is the basis of this chapter. As it is explained in [65] the Mizuno–Todd–Ye predictor-corrector method is based on a very simple and elegant idea that is used in various other fields of computational mathematics – such as the numerical methods of differential equations and continuation methods. We have already mentioned that there are $\mathbb{NP}$-complete LCPs. The $\mathbb{NP}$-completeness of LCPs is related to the properties of the matrix M. Therefore, our aim is to generalize the Mizuno–Todd–Ye algorithm – which is one of the most remarkable interior point methods for linear programming and quadratic programming – for the widest possible matrix class where the method is polynomial. This is the $\mathcal{P}_*$ matrix class defined by Kojima et al. [51].

When choosing the proximity measure we followed Potra ($\|\mathbf{v}^{-1}-\mathbf{v}\|$) in contrast with the previous works ($\|\mathbf{v}-\mathbf{e}\|$, where $\mathbf{v} = \sqrt{\mathbf{xs}/\mu}$). The reason for this was that in practice interior point algorithms make longer steps in a wide neighbourhood, therefore their practical performance might be even better than the theoretical one. Furthermore, the Mizuno–Todd–Ye algorithm for LCPs with a $\mathcal{P}_*(\kappa)$-matrix was not generalized earlier for large neighbourhoods.

Summarizing the discussion above, in this chapter we present a new variant of the Mizuno–Todd–Ye predictor-corrector algorithm for $\mathcal{P}_*(\kappa)$ LCPs that uses self-regular proximity measure $\|\mathbf{v}^{-1} - \mathbf{v}\|$, and therefore the iterates lies in a wider neighbourhood of the central path than in the earlier published Mizuno–Todd–Ye type algorithms for this class of problems. The iteration complexity of our algorithm is $O((1+\kappa)^{\frac{3}{2}}\sqrt{n}L)$.

The rest of the chapter is organized as follows. Section 5.1 describes the predictor-corrector algorithm and the following sections analyze the method. Section 5.1.1 deals with the predictor step and determines the length of the Newton-step. The next part examines the corrector step and the relationship between the proximity parameters τ and $\hat{\tau}$. The last section provides the iteration complexity of the algorithm. The results of this chapter have

been published in paper [40].

Throughout the chapter we make the following assumptions:

1. the initial point $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{F}^+$, i.e., $\mathcal{F}^+ \neq \emptyset$ and an interior point is known;
2. the coefficient matrix M is a $\mathcal{P}_*(\kappa)$-matrix with an a priori known $\kappa \geq 0$.

5.1 The predictor-corrector algorithm

We use the proximity measure δ_c and the central path neighbourhood $\mathcal{N}_c$ through the construction and analysis of the algorithm (see Chapter 3 for the definitions). Let $0 < \tau < \hat{\tau}$ be the suitable proximity parameters defining the neighbourhood of the central path (their relationship will be determined later).

In the predictor[1] step we solve the Newton system (3.10) with a greedy $\mu = 0$ choice from a given initial point $(\mathbf{x}^0, \mathbf{s}^0, \mu_0)$, which is in the smaller neighbourhood of the central path, namely $\delta_c(\mathbf{x}^0\,\mathbf{s}^0, \mu_0) < \tau$. We denote the solution by $(\Delta_p\mathbf{x}, \Delta_p\mathbf{s})$. Let our new point be

$$\mathbf{x}^p = \mathbf{x}^0 + \theta_p\,\Delta_p\mathbf{x}, \quad \mathbf{s}^p = \mathbf{s}^0 + \theta_p\,\Delta_p\mathbf{s}, \quad \mu_p = (1-\theta_p)\mu_0,$$

where $\theta \in (0,1]$ is the maximal real number for which $\delta_c(\mathbf{x}^p\,\mathbf{s}^p, \mu_p) \leq \hat{\tau}$ is satisfied, i.e., we allow certain amount of deviation from the smaller neighbourhood after the predictor step.

The aim of the corrector step is to return to the τ-neighbourhood of the central path. By solving the Newton system (3.10) with $\mathbf{x} = \mathbf{x}^p$, $\mathbf{s} = \mathbf{s}^p$, $\mu = \mu_p$, we get $(\Delta_c\mathbf{x}, \Delta_c\mathbf{s})$. Let the new point be

$$\mathbf{x}^c = \mathbf{x}^p + \Delta_c\mathbf{x}, \quad \mathbf{s}^c = \mathbf{s}^p + \Delta_c\mathbf{s}, \quad \mu_c = \mu_p,$$

so in contrast with the predictor step we do a full Newton step here. The point $(\mathbf{x}^c, \mathbf{s}^c, \mu_c)$ will be in the τ neighbourhood again, so we can continue the iteration with this point as an initial point $(\mathbf{x}^0, \mathbf{s}^0, \mu_0)$, until the duality gap is not small enough.

Accordingly, we will use the following notations in this chapter: in each iteration the current initial point is $(\mathbf{x}, \mathbf{s})$. The predictor Newton direction is $(\Delta_p\mathbf{x}, \Delta_p\mathbf{s})$ and $\mathbf{x}(\theta) = \mathbf{x}+\theta\Delta_p\mathbf{x}$, $\mathbf{s}(\theta) = \mathbf{s}+\theta\Delta_p\mathbf{s}$, $\mu(\theta) = (1-\theta)\mu$. We get the predictor point with the step length θ_p, namely $\mathbf{x}^p = \mathbf{x}(\theta_p)$, $\mathbf{s}^p = \mathbf{s}(\theta_p)$, $\mu_p = \mu(\theta_p)$. The corrector Newton direction is $(\Delta_c\mathbf{x}, \Delta_c\mathbf{s})$ and the corrector point is $\mathbf{x}^c = \mathbf{x}^p + \Delta_c\mathbf{x}$, $\mathbf{s}^c = \mathbf{s}^p + \Delta_c\mathbf{s}$. Similarly we use the superscripts for the notations of scaling, i.e., without a superscript it is for the initial point $(\mathbf{x}, \mathbf{s}, \mu)$, with superscript p it is for the predictor point $(\mathbf{x}^p, \mathbf{s}^p, \mu_p)$ and with superscript c it is for the corrector point $(\mathbf{x}^c, \mathbf{s}^c, \mu_c)$, for example $\mathbf{v} = \sqrt{\mathbf{x}\,\mathbf{s}/\mu}$, $\mathbf{v}^p = \sqrt{\mathbf{x}^p\,\mathbf{s}^p/\mu_p}$ and $\mathbf{v}^c = \sqrt{\mathbf{x}^c\,\mathbf{s}^c/\mu_c}$. Furthermore, $\mathbf{v}(\theta) = \sqrt{\mathbf{x}(\theta)\mathbf{s}(\theta)/\mu(\theta)}$.

[1] Sometimes this kind of predictor step is called the *affine scaling step* in the literature [70].

Now let us outline the algorithm as it follows:

Mizuno–Todd–Ye predictor-corrector algorithm

Input:

an accuracy parameter $\varepsilon > 0$;
proximity parameters $\tau,\ \hat{\tau}, \qquad 0 < \tau < \hat{\tau}$;
an initial point $(\mathbf{x}^0, \mathbf{s}^0, \mu_0), \qquad \delta_c(\mathbf{x}^0\,\mathbf{s}^0, \mu_0) \le \tau$;

begin
$\mathbf{x} := \mathbf{x}^0$, $\mathbf{s} := \mathbf{s}^0$, $\mu := \mu_0$
while $\mathbf{x}^T\mathbf{s} \ge \varepsilon$ **do**
Predictor step
calculate the predictor Newton direction $(\Delta_p\mathbf{x}, \Delta_p\mathbf{s})$ by solving (3.10) with $\mu = 0$;
determine $\theta_p = \max\{\theta > 0 : \big(\mathbf{x}(\theta), \mathbf{s}(\theta)\big) \in \mathcal{F}^+ \text{ and } \delta_c\big(\mathbf{x}(\theta)\,\mathbf{s}(\theta), \mu(\theta)\big) \le \hat{\tau}\}$;
$\mathbf{x}^p = \mathbf{x} + \theta_p\,\Delta_p\mathbf{x},\ \mathbf{s}^p = \mathbf{s} + \theta_p\,\Delta_p\mathbf{s},\ \mu_p = (1-\theta_p)\mu$;
Corrector step
calculate the corrector Newton direction $(\Delta_c\mathbf{x}, \Delta_c\mathbf{s})$
by solving (3.10) with $\mathbf{x} = \mathbf{x}^p,\ \mathbf{s} = \mathbf{s}^p,\ \mu = \mu_p$;
$\mathbf{x}^c = \mathbf{x}^p + \Delta_c\mathbf{x},\ \mathbf{s}^c = \mathbf{s}^p + \Delta_c\mathbf{s},\ \mu_c = \mu_p$;
$\mathbf{x} = \mathbf{x}^c,\ \mathbf{s} = \mathbf{s}^c,\ \mu = \mu_c$;
end
end.

In the rest of the chapter we deal with the analysis of the previous algorithm. In the analysis of the predictor step we will give conditions to the step length θ and at the corrector step to the proximity parameters τ and $\hat{\tau}$.

5.1.1 The predictor step

Let $(\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+$ and $\delta_c(\mathbf{x}\,\mathbf{s}, \mu) < \tau$. We solve the Newton system with parameter $\mu = 0$ that means we get the following scaled predictor system[2]

$$\left.\begin{aligned} -\bar{M}_p\,\mathbf{d}_x^p + \mathbf{d}_s^p &= \mathbf{0} \\ \mathbf{d}_x^p + \mathbf{d}_s^p &= -\mathbf{v} \end{aligned}\right\}$$

[2]See the notations of scaling in Section 3.5.

where

$$\mathbf{v} = \sqrt{\frac{\mathbf{x}\,\mathbf{s}}{\mu}}, \quad \mathbf{d}_x^p = \frac{\mathbf{v}\Delta_p\mathbf{x}}{\mathbf{x}}, \quad \mathbf{d}_s^p = \frac{\mathbf{v}\Delta_p\mathbf{s}}{\mathbf{s}}, \quad \mathbf{d} = \sqrt{\frac{\mathbf{x}}{\mathbf{s}}} \quad \text{and} \quad \bar{M}_p = DMD.$$

We choose the step length θ so that the following two conditions are satisfied:

1. $(\mathbf{x}^p, \mathbf{s}^p) \in \mathcal{F}^+$,
2. $\delta_c(\mathbf{x}^p\,\mathbf{s}^p, \mu_p) < \hat{\tau}$,

namely, the positivity condition is satisfied at the new point and we do not deviate from the central path too much.

We can examine the first condition as it is usual in the interior point literature, thus

$$\mathbf{x}^p\mathbf{s}^p = \mu_p\,\mathbf{v}^2 + \theta_p^2\,\mu\,\mathbf{d}_x^p\mathbf{d}_s^p\,, \quad \text{and so} \quad (\mathbf{v}^p)^2 = \frac{\mathbf{x}^p\mathbf{s}^p}{(1-\theta_p)\mu} = \mathbf{v}^2 + \frac{\theta_p^2}{1-\theta_p}\,\mathbf{d}_x^p\mathbf{d}_s^p > 0$$

is necessary for $(\mathbf{x}^p, \mathbf{s}^p) \in \mathcal{F}^+$. In details, it means if $(d_x^p d_s^p)_i \geq 0$, then obviously $v_i^p > 0$ is true, but if $(d_x^p d_s^p)_i < 0$, then the condition $\theta_p^2/(1-\theta_p) < -v_i^2/(d_x^p d_s^p)_i$ has to hold for the step length.

Let us denote $\hat{\varphi} = \min\left\{-\frac{v_i^2}{(d_x^p d_s^p)_i} : (d_x^p d_s^p)_i < 0\right\}$. Let $\varphi = \frac{\theta^2}{1-\theta}$, so there is a one-to-one correspondence between a positive φ and $0 < \theta < 1$. Then $\varphi \in [0, \hat{\varphi})$ is the necessary condition for $(\mathbf{x}^p, \mathbf{s}^p) \in \mathcal{F}^+$ to be satisfied.

Let us now examine the second condition for choosing step length θ. For this purpose, compute the proximity measure of vector $\mathbf{v}(\theta)$

$$\delta_c(\mathbf{v}(\theta))^2 = \|\mathbf{v}(\theta)^{-1} - \mathbf{v}(\theta)\|^2 = \delta_c(\mathbf{v})^2 + \mathbf{e}^T\left(\frac{\mathbf{e}}{\mathbf{v}^2 + \varphi\,\mathbf{d}_x^p\mathbf{d}_s^p} - \frac{\mathbf{e}}{\mathbf{v}^2} + \varphi\,\mathbf{d}_x^p\mathbf{d}_s^p\right).$$

Analogously to [65], let $f : [0, \hat{\varphi}) \to \mathbb{R}$ be the function of the φ defined as the difference of the square of new and old proximity measures, therefore

$$f(\varphi) = \delta_c(\mathbf{v}(\theta))^2 - \delta_c(\mathbf{v})^2 = \mathbf{e}^T\left(\frac{\mathbf{e}}{\mathbf{v}^2 + \varphi\,\mathbf{d}_x^p\mathbf{d}_s^p} - \frac{\mathbf{e}}{\mathbf{v}^2} + \varphi\,\mathbf{d}_x^p\mathbf{d}_s^p\right).$$

It is easy to show that the function f is strictly convex and nonnegative on the interval $[0, \hat{\varphi})$, $f(0) = 0$ and $\lim_{\varphi \to \hat{\varphi}} f(\varphi) = \infty$. From the properties of function f, it follows that the equation

$$f(\varphi) = \hat{\tau}^2 - \delta_c(\mathbf{v})^2$$

has a unique solution on the interval $[0, \hat{\varphi})$. Let it be $\varphi^* = \frac{(\theta^*)^2}{1-\theta^*}$, and then the step length can be computed as $\theta^* = \frac{-\varphi^* + \sqrt{(\varphi^*)^2 + 4\varphi^*}}{2}$.

Lemma 5.1 *The θ^* is the maximal feasible step length such that $\delta_c(\mathbf{v}(\theta)) \leq \hat{\tau}$ and the new point $(\mathbf{x}(\theta), \mathbf{s}(\theta)) \in \mathcal{F}^+$ for all $0 < \theta \leq \theta^*$.*

Proof. If $\theta > \theta^*$, then $\delta_c(\mathbf{v}(\theta)) > \hat{\tau}$ holds by definition of θ^*. For the feasibility of the step length θ^* the following properties have to be satisfied.

$$\begin{aligned} \mathbf{q} + M\mathbf{x}(\theta) &= \mathbf{s}(\theta), \\ \mathbf{x}(\theta), \mathbf{s}(\theta) &> 0, \qquad \forall\, \theta \in [0, \theta^*]. \end{aligned}$$

Suppose on the contrary that $\exists\, \theta' \in [0, \theta^*] : \mathbf{x}(\theta')\mathbf{s}(\theta') = 0$. Then $\lim_{\theta \to \theta'} \delta_c(\mathbf{v}(\theta)) = \infty$. This contradicts to $\delta_c(\mathbf{v}(\theta))^2 = \delta_c(\mathbf{v})^2 + f(\varphi) < \infty$. ■

5.1.2 The corrector step

In the predictor step the point $(\mathbf{x}^p, \mathbf{s}^p, \mu_p)$ is obtained. Our aim now is to return from the $\hat{\tau}$-neighbourhood to the τ-neighbourhood of the central path. For this, we solve the following Newton system

$$\left.\begin{aligned} -M\Delta_c\mathbf{x} + \Delta_c\mathbf{s} &= 0 \\ \mathbf{x}^p\Delta_c\mathbf{s} + \mathbf{s}^p\Delta_c\mathbf{x} &= \mu_p\mathbf{e} - \mathbf{x}^p\mathbf{s}^p \end{aligned}\right\}$$

After scaling, one has

$$\left.\begin{aligned} -\bar{M}_c\,\mathbf{d}_x^c + \mathbf{d}_s^c &= 0 \\ \mathbf{d}_x^c + \mathbf{d}_s^c &= (\mathbf{v}^p)^{-1} - \mathbf{v}^p \end{aligned}\right\}$$

where

$$\mathbf{v}^p = \sqrt{\frac{\mathbf{x}^p\,\mathbf{s}^p}{\mu_p}},\ \mathbf{d}_x^c = \frac{\mathbf{v}^p\Delta_c\mathbf{x}}{\mathbf{x}^p},\ \mathbf{d}_s^c = \frac{\mathbf{v}^p\Delta_c\mathbf{s}}{\mathbf{s}^p},\ \mathbf{d}^p = \sqrt{\frac{\mathbf{x}^p}{\mathbf{s}^p}}, \quad D_p = diag(\mathbf{d}^p),\ \bar{M}_c = D_pMD_p.$$

Since $\mathbf{x}^c = \mathbf{x}^p + \Delta_c\mathbf{x}$, $\mathbf{s}^c = \mathbf{s}^p + \Delta_c\mathbf{s}$, $\mu_c = \mu_p = (1-\theta_p)\mu$, and $\mathbf{v}^c = \sqrt{\mathbf{x}^c\,\mathbf{s}^c/\mu_c}$, we have $(\mathbf{v}^c)^2 = \mathbf{e} + \mathbf{d}_x^c\mathbf{d}_s^c$. We will see that $(\mathbf{v}^c)^2 = \mathbf{e} + \mathbf{d}_x^c\mathbf{d}_s^c > 0$, namely, the corrector point is feasible if τ parameter is in the given interval (Proposition 5.3), because in that case $\|\mathbf{d}_x^c\mathbf{d}_s^c\|_\infty < 1$.

By similar computation we get $\left((\mathbf{v}^c)^{-1}\right)^2 = \dfrac{\mathbf{e}}{\mathbf{e} + \mathbf{d}_x^c\mathbf{d}_s^c}$. Then the proximity measure is

$$\delta_c(\mathbf{v}^c)^2 = \|\mathbf{v}^c - (\mathbf{v}^c)^{-1}\|^2 = \mathbf{e}^T\frac{(\mathbf{d}_x^c\mathbf{d}_s^c)^2}{\mathbf{e} + \mathbf{d}_x^c\mathbf{d}_s^c}.$$

Introducing $\mathbf{r}^c = \Delta_c\mathbf{x}\Delta_c\mathbf{s}/\mu_p$ we get the following expression:

$$\delta_c(\mathbf{v}^c)^2 = \mathbf{e}^T\frac{(\mathbf{r}^c)^2}{\mathbf{e} + \mathbf{r}^c} = \sum_{i=1}^{n}\frac{(r_i^c)^2}{1 + r_i^c} = \sum_{i\in\mathcal{I}_+}\frac{(r_i^c)^2}{1 + r_i^c} + \sum_{i\in\mathcal{I}_-}\frac{(r_i^c)^2}{1 - |r_i^c|},$$

where $\mathcal{I}_+ = \{1 \le i \le n : r_i^c > 0\}$ and $\mathcal{I}_- = \{1 \le i \le n : r_i^c < 0\}$.

As the matrix M is a $\mathcal{P}_*(\kappa)$-matrix, the further analysis of the corrector step is different from the analysis of [65], since in our case $\kappa > 0$ is possible, too, but Potra dealt only with the case $\kappa = 0$. Because M is a $\mathcal{P}_*(\kappa)$-matrix by definition the following estimation holds:

$$\sum_{i\in\mathcal{I}_-} |r_i^c| \le (1+4\kappa) \sum_{i\in\mathcal{I}_+} r_i^c. \tag{5.1}$$

From inequality (5.1) follows that $\max_j r_j^c > 0$ (so $\max_j (\Delta_c x)_j (\Delta_c s)_j > 0$). This consequence will be used for example in Proposition 5.5.

We can estimate the distance of the new point and the central path by applying the property of the infinity norm and (5.1) as it follows:

$$\begin{aligned}
\delta_c(\mathbf{x}^c\,\mathbf{s}^c,\mu_c)^2 &= \sum_{i\in\mathcal{I}_+}^{n} \frac{(r^c)_i^2}{1+r_i^c} + \sum_{i\in\mathcal{I}_-}^{n} \frac{(r^c)_i^2}{1-|r_i^c|} \le \sum_{i\in\mathcal{I}_+}^{n} \frac{r_i^c\,\|\mathbf{r}^c\|_\infty}{1+\|\mathbf{r}^c\|_\infty} + \sum_{i\in\mathcal{I}_-}^{n} \frac{|r_i^c|\,\|\mathbf{r}^c\|_\infty}{1-\|\mathbf{r}^c\|_\infty} \\
&\le \frac{\|\mathbf{r}^c\|_\infty}{1+\|\mathbf{r}^c\|_\infty} \sum_{i\in\mathcal{I}_+} r_i^c + \frac{\|\mathbf{r}^c\|_\infty}{1-\|\mathbf{r}^c\|_\infty} (1+4\kappa) \sum_{i\in\mathcal{I}_+} r_i^c \\
&= \frac{2\|\mathbf{r}^c\|_\infty}{1-\|\mathbf{r}^c\|_\infty^2} \left(1+2\kappa+2\kappa\|\mathbf{r}^c\|_\infty\right) \sum_{i\in\mathcal{I}_+} r_i^c.
\end{aligned}$$

Applying the estimation of Newton directions in Lemma 3.13 with $\mathbf{a} = \mu_p \mathbf{e} - \mathbf{x}^p\,\mathbf{s}^p$ we have

$$\|\mathbf{r}^c\|_\infty \le \frac{1+4\kappa}{4}\,\delta_c(\mathbf{x}^p\,\mathbf{s}^p,\mu_p)^2 \quad \text{and} \quad \sum_{i\in\mathcal{I}_+} r_i^c \le \frac{1}{4}\,\delta_c(\mathbf{x}^p\,\mathbf{s}^p,\mu_p)^2. \tag{5.2}$$

Now we can continue the estimation of $\delta_c(\mathbf{x}^c\,\mathbf{s}^c,\mu_c)$ as it follows:

$$\begin{aligned}
\delta_c(\mathbf{x}^c\,\mathbf{s}^c,\mu_c)^2 &\le \frac{2\|\mathbf{r}^c\|_\infty}{1-\|\mathbf{r}^c\|_\infty^2} \left(1+2\kappa+2\kappa\|\mathbf{r}^c\|_\infty\right) \sum_{i\in\mathcal{I}_+} r_i^c \\
&\le \frac{(1+4\kappa)\delta_c^4}{16-(1+4\kappa)^2\delta_c^4} \left(2+4\kappa+\kappa\,(1+4\kappa)\,\delta_c^2\right),
\end{aligned}$$

where $\delta_c := \delta_c(\mathbf{x}^p\,\mathbf{s}^p,\mu_p)$.

We would like to satisfy $\delta_c(\mathbf{x}^c\,\mathbf{s}^c,\mu_c)^2 < \tau^2$. To achieve this, we require a stronger condition, i.e.,

$$\frac{(1+4\kappa)\delta_c^4}{16-(1+4\kappa)^2\delta_c^4} \left(2+4\kappa+\kappa(1+4\kappa)\delta_c^2\right) < \tau^2.$$

After simple computation we get

$$\kappa(1+4\kappa)^2\delta_c^6 + \left[(2+4\kappa)(1+4\kappa) + (1+4\kappa)^2\tau^2\right] \delta_c^4 < 16\tau^2. \tag{5.3}$$

We know that $\delta_c = \delta_c(\mathbf{x}^p \mathbf{s}^p, \mu_p) < \hat{\tau}$, therefore with a suitable choice of τ and $\hat{\tau}$, inequality (5.3) has a solution in terms of δ_c. Combining these results we can express the relationship between τ and $\hat{\tau}$ as follows

$$\kappa(1+4\kappa)^2\hat{\tau}^6 + \left[(2+4\kappa)(1+4\kappa) + (1+4\kappa)^2\tau^2\right]\hat{\tau}^4 < 16\tau^2. \tag{5.4}$$

Lemma 5.2 *If*

$$\hat{\tau} < \frac{2\sqrt{\tau}}{\sqrt{1+4\kappa}\sqrt[4]{2+\tau^2}}, \tag{5.5}$$

then the inequality (5.4) is true for all $\hat{\tau} > \tau > 0$ and for all $\kappa \geq 0$.

Proof. By inequality (5.5) one has

$$\hat{\tau}^6 < \frac{64\tau^3}{(1+4\kappa)^3(2+\tau^2)\sqrt{2+\tau^2}} \quad \text{and} \quad \hat{\tau}^4 < \frac{16\tau^2}{(1+4\kappa)^2(2+\tau^2)}.$$

Then applying inequality (5.4) we obtain

$$\begin{aligned} \kappa(1+4\kappa)^2\hat{\tau}^6 &+ \left[(2+4\kappa)(1+4\kappa) + (1+4\kappa)^2\tau^2\right]\hat{\tau}^4 \\ &< 16\tau^2 \frac{1}{1+4\kappa}\left[\frac{4\tau\kappa}{(2+\tau^2)\sqrt{2+\tau^2}} + \frac{2+4\kappa+(1+4\kappa)\tau^2}{2+\tau^2}\right] \\ &< 16\tau^2 \frac{1}{1+4\kappa}\frac{2(1+4\kappa)+(1+4\kappa)\tau^2}{2+\tau^2} = 16\tau^2. \end{aligned}$$

This completes the proof. ■

Considering the relationship between corrector and predictor proximity parameters τ and $\hat{\tau}$ given in Proposition 5.2, we can determine the suitable value of τ depending on the parameter κ.

Lemma 5.3 *If $\tau < \sqrt{-1+\sqrt{1+\frac{16}{(1+4\kappa)^2}}}$, then there exists a positive $\hat{\tau}$ that satisfies (5.5) and the full Newton step is feasible in the corrector step.*

Proof. By Proposition 5.2 we have

$$\tau < \hat{\tau} < \frac{2\sqrt{\tau}}{\sqrt{1+4\kappa}\sqrt[4]{2+\tau^2}},$$

which implies

$$\tau^4 < \frac{16\tau^2}{(1+4\kappa)^2(2+\tau^2)}.$$

After rearrangement we get

$$(1+4\kappa)^2\tau^4 + 2(1+4\kappa)^2\tau^2 - 16 < 0.$$

The zero points of the quadratic expression on the left side of the previous inequality are

$$(\tau^2)_{1,2} = \frac{-2(1+4\kappa)^2 \pm \sqrt{4(1+4\kappa)^4 + 4\cdot 16(1+4\kappa)^2}}{2(1+4\kappa)^2} = -1 \pm \sqrt{1+\frac{16}{(1+4\kappa)^2}}.$$

Considering the signs, we get the first statement.
The condition $(\mathbf{x}^c, \mathbf{s}^c) \in \mathcal{F}^+$, namely that $\mathbf{x}^c\,\mathbf{s}^c > 0$ is equivalent with $(\mathbf{v}^c)^2 = \mathbf{e} + \mathbf{d}_x^c\mathbf{d}_s^c > 0$. From inequality (5.2) follows

$$\|\mathbf{d}_x^c\mathbf{d}_s^c\|_\infty = \|\mathbf{r}^c\|_\infty \le \frac{1+4\kappa}{4}\,\delta_c(\mathbf{x}^p\,\mathbf{s}^p, \mu_p)^2 \le \frac{1+4\kappa}{4}\,\tau^2.$$

Using the given upper bound on τ

$$\|\mathbf{d}_x^c\mathbf{d}_s^c\|_\infty < \frac{1+4\kappa}{4}\left(-1+\sqrt{1+\frac{16}{(1+4\kappa)^2}}\right) < \frac{1+4\kappa}{4}\sqrt{\frac{16}{(1+4\kappa)^2}} = 1.$$

■

Obviously the value of κ effects the size of the neighbourhood parameters τ and $\hat{\tau}$. The larger the value of κ is, the smaller the τ and $\hat{\tau}$ neighbourhoods are which ensures that the Mizuno–Todd–Ye predictor-corrector algorithm takes one predictor and one corrector step alternately.

5.2 Iteration complexity analysis

Thereafter we deal with the iteration complexity of the algorithm following the analysis of [65]. For this purpose, we need a lower bound on the maximal feasible step length of the predictor Newton step θ^*. We determine it in three steps. At first we give an upper bound on the range of function f and using this we determine a lower bound on φ^*. Finally, a bound on θ is computed.

We will use the following lemma in our estimations.

Lemma 5.4 *([65]) Let $\mathbf{x}, \mathbf{s} \in \mathbb{R}^n_+$ and $\mu > 0$ number and let $\delta_c^2(\mathbf{x}\,\mathbf{s}, \mu) \le 2\eta$. Then*

$$\frac{1}{1+\eta+\sqrt{2\eta+\eta^2}} \le \frac{x_i s_i}{\mu} \le 1+\eta+\sqrt{2\eta+\eta^2},\ \textit{for all } i = 1,\dots,n.$$

Now the application of Lemma 5.4 with $\eta = \frac{\tau^2}{2}$ gives

$$\frac{1}{m(\tau)} \le \min_i v_i^2 \le \left\|\mathbf{v}^2\right\|_\infty \le m(\tau), \quad \text{where} \quad m(\tau) = 1 + \frac{\tau^2}{2} + \sqrt{\tau^2 + \frac{\tau^4}{4}}. \tag{5.6}$$

First, we divide the feasible set of φ into two parts:

$$\left[0\,,\ \frac{3}{8}\,\frac{\min_i x_i s_i}{\max_j (\Delta_p x)_j (\Delta_p s)_j}\right] \qquad \text{and} \qquad \left(\frac{3}{8}\,\frac{\min_i x_i s_i}{\max_j (\Delta_p x)_j (\Delta_p s)_j}\,,\ \hat{\varphi}\right].$$

We give a lower bound on φ^* on the two intervals separately. In the second interval $\frac{3}{8}\, m(\tau)/\|\mathbf{r}^p\|_\infty$ is a trivial lower bound (where $\mathbf{r}^p = \Delta_p \mathbf{x}\, \Delta_p \mathbf{s}/\mu$). In the first interval the function f is finite because $\frac{3}{8}\frac{\min_i x_i s_i}{\max_j (\Delta x)_j (\Delta s)_j} \le \hat{\varphi}$ (the constant 3/8 will be analyzed in the Section 5.3). Therefore, we can get an upper bound on the function value of f on the first interval.

Lemma 5.5 *For any* $\varphi \in \left[0, \frac{3}{8}\,\frac{\min_i x_i s_i}{\max_j (\Delta_p x)_j (\Delta_p s)_j}\right]$ *one has*

$$f(\varphi) \le (1 + 4\kappa)\, \varphi\, h(\tau) \sum_{i \in \mathcal{I}_+} r_i^p, \tag{5.7}$$

where $h(\tau) = 8\left(\frac{m^2(\tau)}{5} - \frac{1}{11 m^2(\tau)}\right)$ *and* $\mathbf{r}^p = \frac{\Delta_p \mathbf{x}\, \Delta_p \mathbf{s}}{\mu}$.

Proof. Using the properties of the infinity norm and $\mathcal{P}_*(\kappa)$-matrices one has

$$\begin{aligned}
f(\varphi) &= \sum_{i=1}^n \left(\frac{1}{v_i^2 + \varphi\, r_i^p} - \frac{1}{v_i^2} + \varphi\, r_i^p\right) = \sum_{i=1}^n \varphi\, r_i^p \left(1 - \frac{1}{v_i^2 (v_i^2 + \varphi\, r_i^p)}\right) \\
&= \sum_{i \in \mathcal{I}_+} \varphi\, r_i^p \left(1 - \frac{1}{v_i^2 (v_i^2 + \varphi\, r_i^p)}\right) - \sum_{i \in \mathcal{I}_-} \varphi\, |r_i^p| \left(1 - \frac{1}{v_i^2 (v_i^2 - \varphi\, |r_i^p|)}\right) \\
&\le \sum_{i \in \mathcal{I}_+} \varphi\, r_i^p \left(1 - \frac{1}{\|\mathbf{v}\|_\infty^2 (\|\mathbf{v}\|_\infty^2 + \varphi \|\mathbf{r}^p\|_\infty)}\right) \\
&\quad + \sum_{i \in \mathcal{I}_-} \varphi\, |r_i^p| \left(\frac{1}{\min_i v_i^2 (\min_i v_i^2 - \varphi\, \|\mathbf{r}^p\|_\infty)} - 1\right) \\
&\le \sum_{i \in \mathcal{I}_+} \varphi\, r_i^p \left(1 - \frac{1}{\|\mathbf{v}\|_\infty^2 (\|\mathbf{v}\|_\infty^2 + \varphi \|\mathbf{r}^p\|_\infty)}\right) \\
&\quad + \sum_{i \in \mathcal{I}_+} \varphi\, r_i^p (1 + 4\kappa) \left(\frac{1}{\min_i v_i^2 (\min_i v_i^2 - \varphi\, \|\mathbf{r}^p\|_\infty)} - 1\right).
\end{aligned}$$

Considering the domain of function f and applying estimation (5.6)

$$\varphi \|\mathbf{r}^p\|_\infty \le \frac{3}{8} \min_i v_i^2 \le \frac{3}{8} \|\mathbf{v}\|_\infty^2 \le \frac{3}{8}\, m(\tau) \quad \text{and} \quad \min_i v_i^2 \ge \frac{1}{m(\tau)}$$

holds and we can continue the estimation using previous inequalities:

$$\begin{aligned} f(\varphi) &\leq \varphi \sum_{i\in\mathcal{I}_+} r_i^p \left[1 - \frac{1}{\frac{11}{8}m^2(\tau)} + (1+4\kappa)\left(\frac{1}{\frac{5}{8}\frac{1}{m^2(\tau)}} - 1\right)\right] \\ &= \varphi \sum_{i\in\mathcal{I}_+} r_i^p \left[8\frac{m^2(\tau)}{5} - \frac{8}{11m^2(\tau)} + 4\kappa\left(\frac{8m^2(\tau)}{5} - 1\right)\right], \end{aligned}$$

where $m^2(\tau) \geq 1$, therefore $\frac{8m^2(\tau)}{5} - 1 \leq \frac{8m^2(\tau)}{5} - \frac{8}{11m^2(\tau)}$. ■

Corollary 5.6 *If* $\varphi \in \left[0, \frac{3}{8}\frac{\min_i x_i s_i}{\max_j (\Delta_p x)_j (\Delta_p s)_j}\right]$, *then*

$$\varphi \geq \frac{f(\varphi)}{(1+4\kappa)\, h(\tau) \sum\limits_{i\in\mathcal{I}_+} r_i^p}. \tag{5.8}$$

Therefore, for the uniquely determined φ^* *one has*

$$\varphi^* \geq \tilde{\varphi} = \min_i \left\{ \frac{\hat{\tau}^2 - \delta_c^2(\mathbf{x}\,\mathbf{s}, \mu)}{(1+4\kappa)h(\tau)\sum_{i\in\mathcal{I}_+} r_i^p}\ ,\ \frac{3}{8m(\tau)\|\mathbf{r}^p\|_\infty} \right\}.$$

The lower bound on φ^* is obtained by using the first statement of Corollary 5.6 with substituting the maximal value of $f(\varphi)$, considering the restricted domain of function f.

The next lemma gives estimations for further reordering of inequality (5.8).

Lemma 5.7 *Let* $\mathbf{r}^p = \frac{\Delta_p \mathbf{x}\, \Delta_p \mathbf{s}}{\mu}$. *Then*

$$\sum_{i\in\mathcal{I}_+} r_i^p \leq \frac{n}{4}\, m\!\left(\frac{\tau}{\sqrt{n}}\right) \quad \textit{and} \quad \|\mathbf{r}^p\|_\infty \leq \frac{1+4\kappa}{4}\, n\, m\!\left(\frac{\tau}{\sqrt{n}}\right).$$

Proof. If $\delta_c(\mathbf{x}\,\mathbf{s}, \mu)^2 \leq 2\eta$, then based on Lemma 5.4 and [65]

$$\frac{1}{m\!\left(\sqrt{\frac{2\eta}{n}}\right)} = \frac{1}{1 + \frac{\eta}{n} + \sqrt{\frac{2\eta}{n} + \frac{\eta^2}{n^2}}} \leq \frac{\mathbf{x}^T\mathbf{s}}{n\mu} \leq 1 + \frac{\eta}{n} + \sqrt{\frac{2\eta}{n} + \frac{\eta^2}{n^2}} = m\!\left(\sqrt{\frac{2\eta}{n}}\right).$$

By applying Lemma 3.13 on the vector $\mathbf{a} = -\mathbf{x}\mathbf{s}$ one has

$$\sum_{i\in\mathcal{I}_+} \Delta x_i \Delta s_i \leq \frac{1}{4}\left\|\sqrt{\mathbf{x}\mathbf{s}}\right\|^2 = \frac{\mathbf{x}^T\mathbf{s}}{4} \quad \text{and} \quad \sum_{i\in\mathcal{I}_+} r_i^p \leq \frac{\mathbf{x}^T\mathbf{s}}{4\mu} \leq \frac{n}{4}\, m\!\left(\frac{\tau}{\sqrt{n}}\right).$$

The statement on $\|\mathbf{r}^p\|_\infty$ can be deduced similarly. ■

Now by Lemma 5.7 one has

$$\begin{aligned}\varphi^* \geq \tilde{\varphi} &\geq \frac{4}{n\,m\left(\frac{\tau}{\sqrt{n}}\right)} \min\left\{\frac{\hat{\tau}^2-\tau^2}{(1+4\kappa)\,h(\tau)}, \frac{3}{8\,m(\tau)\,(1+4\kappa)}\right\} \\ &= \frac{4}{n\,m\left(\frac{\tau}{\sqrt{n}}\right)(1+4\kappa)} \min\left\{\frac{\hat{\tau}^2-\tau^2}{h(\tau)}, \frac{3}{8\,m(\tau)}\right\} = \frac{4\gamma}{n\,m\left(\frac{\tau}{\sqrt{n}}\right)(1+4\kappa)},\end{aligned}$$

where $\gamma = \min\left\{\frac{\hat{\tau}^2-\tau^2}{h(\tau)}, \frac{3}{8\,m(\tau)}\right\}$.

By rearranging the inequality on φ^* to θ^* we get a lower bound on the maximal feasible step length θ^*.

Theorem 5.8 *In the predictor step the maximal feasible step length θ^* satisfies $\theta^* \geq \frac{\chi_n}{\sqrt{n}}$, where*

$$\chi_n = 2\sqrt{\frac{\gamma}{m\left(\frac{\tau}{\sqrt{n}}\right)(1+4\kappa)}}\left(\sqrt{\frac{\gamma}{n\,m\left(\frac{\tau}{\sqrt{n}}\right)(1+4\kappa)}+1} - \sqrt{\frac{\gamma}{n\,m\left(\frac{\tau}{\sqrt{n}}\right)(1+4\kappa)}}\right) \tag{5.9}$$

is a bounded quantity.

Proof.

$$\begin{aligned}\theta &= \frac{-\varphi+\sqrt{\varphi^2+4\varphi}}{2} \\ &\geq \frac{1}{2}\left(-\frac{4\gamma}{n\,m\left(\frac{\tau}{\sqrt{n}}\right)(1+4\kappa)} + \sqrt{\frac{16\gamma^2}{n^2m^2\left(\frac{\tau}{\sqrt{n}}\right)(1+4\kappa)^2} + \frac{16\gamma}{n\,m\left(\frac{\tau}{\sqrt{n}}\right)(1+4\kappa)}}\right) \\ &= 2\sqrt{\frac{\gamma}{n\,m\left(\frac{\tau}{\sqrt{n}}\right)(1+4\kappa)}}\left(-\sqrt{\frac{\gamma}{n\,m\left(\frac{\tau}{\sqrt{n}}\right)(1+4\kappa)}} + \sqrt{\frac{\gamma}{n\,m\left(\frac{\tau}{\sqrt{n}}\right)(1+4\kappa)}+1}\right) \\ &= \frac{\chi_n}{\sqrt{n}}.\end{aligned}$$

The χ_n expression (5.9) is bounded because $\lim\limits_{n\to\infty} m\left(\frac{\tau}{\sqrt{n}}\right) = 1$, namely $\lim\limits_{n\to\infty}\chi_n = 2\sqrt{\frac{\gamma}{1+4\kappa}}$ is finite. ■

It is easy to show that the expression χ_n monotonically increases in size n of the problem. The χ_n expression (5.9) is introduced in [65] for positive semidefinite matrices, corresponding to the case $\kappa = 0$.

We derive the iteration complexity of the algorithm by applying the previous bound of θ^*. At first a lower and an upper bound of the duality gap are determined.

Theorem 5.9 *Denote the length of the predictor Newton step in the i^{th} iteration by θ_i. Then the duality gap of the point obtained in step k satisfies*

$$n\left(\prod_{i=1}^{k}(1-\theta_i)\right)\mu_0\left[1-\frac{\kappa\hat{\tau}^2}{n}\right] \leq \mathbf{x}^{k^T}\mathbf{s}^k \leq n\left(\prod_{i=1}^{k}(1-\theta_i)\right)\mu_0\left[1+\frac{\hat{\tau}^2}{4n}\right].$$

Proof. Denote the point where we arrive by the predictor step from point $(\mathbf{x}^{k-1}, \mathbf{s}^{k-1})$ applying the Newton step with step length θ_k and direction $(\Delta_p\mathbf{x}, \Delta_p\mathbf{s})$ by $(\mathbf{x}^p, \mathbf{s}^p)$. Let $(\Delta_c\mathbf{x}, \Delta_c\mathbf{s})$ be the direction of the corrector Newton step. Then

$$\mathbf{x}^k\mathbf{s}^k = (\mathbf{x}^p + \Delta_c\mathbf{x})(\mathbf{s}^p + \Delta_c\mathbf{s}) = \mu_k\mathbf{e} + \Delta_c\mathbf{x}\Delta_c\mathbf{s}.$$

Thus $\mathbf{x}^{k^T}\mathbf{s}^k = \mu_k n + \Delta_c\mathbf{x}^T\Delta_c\mathbf{s}$. Estimate the scalar product $\Delta_c\mathbf{x}^T\Delta_c\mathbf{s}$ using Lemma 3.13 $(\mathbf{a} = \mu_k\mathbf{e} - \mathbf{x}^p\mathbf{s}^p)$

$$\begin{aligned}\Delta_c\mathbf{x}^T\Delta_c\mathbf{s} &= \sum_{i=1}^{n}(\Delta_c x)_i(\Delta_c s)_i \leq \sum_{i\in\mathcal{I}_+}(\Delta_c x)_i(\Delta_c s)_i \leq \frac{1}{4}\left\|\frac{\mu_k\mathbf{e} - \mathbf{x}^p\mathbf{s}^p}{\sqrt{\mathbf{x}^p\mathbf{s}^p}}\right\|^2 \\ &= \frac{1}{4}\mu_k\left\|\sqrt{\frac{\mu_k}{\mathbf{x}^p\mathbf{s}^p}} - \sqrt{\frac{\mathbf{x}^p\mathbf{s}^p}{\mu_k}}\right\|^2 \leq \frac{1}{4}\mu_k\hat{\tau}^2,\end{aligned}$$

where $\mathcal{I}_+ = \{1 \leq i \leq n : (\Delta_c x)_i(\Delta_c s)_i > 0\}$. Because of the properties of $\mathcal{P}_*(\kappa)$-matrices

$$\begin{aligned}\Delta_c\mathbf{x}^T\Delta_c\mathbf{s} &= \sum_{i\in\mathcal{I}_+}^{n}(\Delta_c x)_i(\Delta_c s)_i + \sum_{i\in\mathcal{I}_-}^{n}(\Delta_c x)_i(\Delta_c s)_i \geq -4\kappa\sum_{i\in\mathcal{I}_+}^{n}(\Delta_c x)_i(\Delta_c s)_i \\ &\geq -\kappa\left\|\frac{\mu_k\mathbf{e} - \mathbf{x}^p\mathbf{s}^p}{\sqrt{\mathbf{x}^p\mathbf{s}^p}}\right\|^2 \geq -\kappa\mu_k\hat{\tau}^2.\end{aligned}$$

Substituting the bounds and considering $\mu_k = \mu_0\prod_{i=1}^{k}(1-\theta_i)$ completes the proof. ■

Using the estimation of the duality gap we derive the iteration complexity of the algorithm.

Theorem 5.10 *Let the LCP for any $\mathcal{P}_*(\kappa)$-matrix M be given, where $\kappa \geq 0$ and let $\mu_0 = 1$. Then the Mizuno–Todd–Ye algorithm generates an $(\mathbf{x}, \mathbf{s}, \mu)$ point satisfying $\mathbf{x}^T\mathbf{s} < \varepsilon$ in at most*

$$\left\lceil \frac{\sqrt{n}}{\chi_1}\log\frac{4n+\hat{\tau}^2}{4\varepsilon} \right\rceil$$

iterations.

Proof. The Mizuno–Todd–Ye algorithm takes only one step both in the predictor and the corrector steps, so it is sufficient to count the updates of parameter μ. Since $\mu_0 = 1$, therefore

after iteration k the duality gap is certainly smaller than ε if

$$n\left(\prod_{i=1}^{k}(1-\theta_i)\right)\left[1+\frac{\hat{\tau}^2}{4n}\right] \;<\; \varepsilon.$$

Consider the given lower bound of θ^* for each iteration

$$\left(1-\frac{\chi_n}{\sqrt{n}}\right)^k \;<\; \frac{\varepsilon}{n\left[1+\frac{\hat{\tau}^2}{4n}\right]},$$

with a simple computing we get

$$k > \frac{\log\dfrac{\varepsilon}{n\left[1+\frac{\hat{\tau}^2}{4n}\right]}}{\log\left(1-\frac{\chi_n}{\sqrt{n}}\right)} = \frac{\log\dfrac{n\left[1+\frac{\hat{\tau}^2}{4n}\right]}{\varepsilon}}{-\log\left(1-\frac{\chi_n}{\sqrt{n}}\right)}.$$

Using properties of the logarithm $-\log(1-\theta)\geq\theta,\ \theta\in(0,1)$ and monotonicity of χ_n

$$\frac{\log\dfrac{n\left[1+\frac{\hat{\tau}^2}{4n}\right]}{\varepsilon}}{-\log\left(1-\frac{\chi_n}{\sqrt{n}}\right)} \leq \frac{\sqrt{n}}{\chi_n}\log\frac{n\left[1+\frac{\hat{\tau}^2}{4n}\right]}{\varepsilon} \leq \frac{\sqrt{n}}{\chi_1}\log\frac{n\left[1+\frac{\hat{\tau}^2}{4n}\right]}{\varepsilon}.$$

■

In Theorem 5.10 the determined number of steps are independent from κ only at first sight, because both χ_1 and $\hat{\tau}$ depend on κ.

Corollary 5.11 *Let an LCP with a $\mathcal{P}_*(\kappa)$-matrix M be given, where $\kappa\geq 0$ and let $\mu_0=1$, $\tau=\frac{1}{1+4\kappa}$ and $\hat{\tau}=\frac{\sqrt{2}}{1+4\kappa}$. Then the Mizuno–Todd–Ye algorithm generates an $(\mathbf{x},\mathbf{s},\mu)$ point satisfying $\mathbf{x}^T\mathbf{s}<\varepsilon$ in at most $O((1+\kappa)^{\frac{3}{2}}\sqrt{n}\log\frac{n}{\varepsilon})$ iterations.*

Proof. It is easy to verify that the given value of parameter τ and $\hat{\tau}$ are feasible, namely, they satisfy inequality (5.5). After that we estimate the result of Theorem 5.10. For a lower estimation of χ_1 examine the value of $m(\tau)$, $h(\tau)$ and γ:

$$1\leq m(\tau)=1+\frac{\tau^2}{2}+\sqrt{\tau^2+\frac{\tau^4}{4}}<(1+\tau)^2,$$

$$h(\tau)=\frac{8}{5}\,m^2(\tau)-\frac{8}{11\,m^2(\tau)}<\frac{8}{5}\,m^2(\tau)<\frac{8}{5}(1+\tau)^4,$$

$$\gamma=\min\left\{\frac{\hat{\tau}^2-\tau^2}{h(\tau)},\frac{3}{8m(\tau)}\right\}>\min\left\{\frac{\tau^2}{8/5(1+\tau)^4},\frac{3}{8(1+\tau)^2}\right\}>\frac{\tau^2}{8(1+\tau)^4},$$

and $\gamma < \frac{3}{8\,m(\tau)} < \frac{3}{8}$. Thus

$$\begin{aligned}
\chi_1 &= 2\sqrt{\frac{\gamma}{m(\tau)(1+4\kappa)}}\left(\sqrt{\frac{\gamma}{m(\tau)(1+4\kappa)}+1}-\sqrt{\frac{\gamma}{m(\tau)(1+4\kappa)}}\right) \\
&= 2\sqrt{\frac{\gamma}{m(\tau)(1+4\kappa)}}\frac{1}{\sqrt{\frac{\gamma}{m(\tau)(1+4\kappa)}+1}+\sqrt{\frac{\gamma}{m(\tau)(1+4\kappa)}}} \\
&> \sqrt{\frac{\gamma}{m(\tau)(1+4\kappa)}}\frac{1}{\sqrt{\frac{\gamma}{m(\tau)(1+4\kappa)}+1}} > \sqrt{\frac{\tau^2}{[3+8(1+4\kappa)](1+\tau)^6}} \\
&> \sqrt{\frac{\tau^2}{11(1+4\kappa)(1+\tau)^6}} = \sqrt{\frac{(1+4\kappa)^6}{11(1+4\kappa)^3(2+4\kappa)^6}} > \frac{1}{8\sqrt{11}}(1+4\kappa)^{-\frac{3}{2}}.
\end{aligned}$$

Substitute the lower bound of χ_1

$$\frac{\sqrt{n}}{\chi_1}\log\frac{n\left[1+\frac{\hat{\tau}^2}{4n}\right]}{\varepsilon} < 8\sqrt{11}(1+4\kappa)^{\frac{3}{2}}\sqrt{n}\log\frac{2n}{\varepsilon}.$$

So the iteration complexity of the algorithm is $O((1+\kappa)^{\frac{3}{2}}\sqrt{n}\log\frac{n}{\varepsilon})$. ■

5.3 Conclusion

We showed that a Mizuno–Todd–Ye type predictor-corrector algorithm can solve LCPs with a $\mathcal{P}_*(\kappa)$-matrix, too. The τ and $\hat{\tau}$ neighbourhood parameters can be chosen in such a way that after each predictor step only one corrector step is needed to return into the pre-established neighbourhood of the central path.

We generalized the analysis of [65] for LCP problems with $\mathcal{P}_*(\kappa)$-matrices. Since we are working with $\mathcal{P}_*(\kappa)$-matrices our analysis is more complicated, because the acceptable values of parameters τ and $\hat{\tau}$ depend on κ. For larger values of κ the values of τ and $\hat{\tau}$ decrease quickly, therefore the constant in the iteration complexity increase. For example if $\kappa = 0.3274$, then the feasible values of parameters τ and $\hat{\tau}$ are less than one, but if $\kappa = 0$ the maximal value of τ is $\sqrt{-1+\sqrt{17}}$.

In the analysis we used 3/8 as the constant (Proposition 5.5, Corollary 5.6) which can be replaced with an arbitrary number $0 < c < 1$. Looking for the best choice of c would make the analysis of the algorithm more complicated without any notable improvement. The $c = 3/8$ choice (first used by Potra) can be considered as quite close to the upper bound on the optimal value of c in case $\kappa = 0$. (A slightly better choice is $c = 0.27975$ but that would make the computation in several proofs much more complicated.)

The update strategy of the central path parameter μ is the same $\mu' = (1-\theta)\mu$ as in [65], but it is different from Miao's, because he updated the parameter μ in such a way that the

equation $\mathbf{x}^T\mathbf{s}/n = \mu$ must hold. Suppose we take a step from point $(\mathbf{x}^0, \mathbf{s}^0)$ to $(\mathbf{x}, \mathbf{s})$ with vector $(\Delta\mathbf{x}, \Delta\mathbf{s})$. Let $\mu := \frac{(\mathbf{x}^0)^T\mathbf{s}^0}{n}$ and denote μ_M Miao's and μ_P Potra's updated value of parameter μ. Then $\mu_M = \frac{\mathbf{x}^T\mathbf{s}}{n}$ and $\mu_P = (1-\theta)\mu$, so

$$\mu_M = \frac{\mathbf{x}^T\mathbf{s}}{n} = (1-\theta)\mu + \theta^2\frac{\Delta\mathbf{x}^T\Delta\mathbf{s}}{n} = \mu_P + \theta^2\frac{\Delta\mathbf{x}^T\Delta\mathbf{s}}{n}.$$

Using the result of Lemma 3.13 we obtain the following connection

$$\mu_M - \frac{1}{4}\theta^2\mu \leq \mu_P \leq \mu_M + \kappa\theta^2\mu.$$

Therefore, Potra's way of updating is always better in the case $\kappa = 0$ (the matrix is a positive semidefinite) and sometimes it is better and sometimes it is worse than Miao's in the case $\kappa > 0$, but the algorithm and its analysis is easier.

Let us note that we can hope for further improvements by using some other self-regular proximity measures [63] in the algorithm.

Chapter 6

Linear complementarity problems for arbitrary matrices

In the previous chapter we made two assumptions: 1. an initial interior point is known; 2. the coefficient matrix is a $\mathcal{P}_*(\kappa)$-matrix with an a priori known $\kappa \geq 0$. The first assumption is mainly technical, as we can use either an infeasible IPM, or the problem of an initial interior point can be handled by an embedding technique (see Section 3.7). We will deal with the latter one. However, the second assumption, the a priori knowledge of the parameter κ is too strong. In contrast with the positive semidefiniteness of a matrix – which can be checked in strongly polynomial time [60] – the best known test to check whether a matrix is $\mathcal{P}_*(\kappa)$ or not, introduced by Väliaho [81], is not polynomial (see Chapter 2). Potra et al. relaxed this assumption [66], they modified their IPM in such a way, that we only need to know the sufficiency of the matrix. However, this is still a condition, that can not be verified in polynomial time. Indeed, Tseng [78], proved that the decision problem 'whether a square matrix with rational entries is a column sufficient matrix' is co-$\mathbb{NP}$-complete, suggesting that it can not be decided in polynomial time whether there is a finite nonnegative κ with which matrix M is $\mathcal{P}_*(\kappa)$ (see Theorem 2.24). Furthermore, while a lot of applications can be formed as an LCP, in most cases the coefficient matrix is not sufficient as we have pointed out in Section 1.2.

The LCP belongs to the class of $\mathbb{NP}$-complete problems (see Chapter 2, Theorem 2.21). Therefore, we can not expect an efficient (polynomial time) solution method for LCPs without requiring some special property of the matrix M. Consequently, there is a need to design such an algorithm, that can handle any LCP with an arbitrary matrix. Therefore, in this chapter we give a general schema how to modify interior point algorithms for $\mathcal{P}_*(\kappa)$-matrix LCPs, with the goal to *process* general LCPs in polynomial time.

Another motivation of the results of this chapter was the paper of Csizmadia and Illés

[16]. The criss-cross algorithm for sufficient matrix LCPs was introduced by Hertog, Roos and Terlaky [18]. The first criss-cross type pivot algorithm in EP form, which does not use a priori knowledge of sufficiency of the matrix M, was given by Fukuda, Namiki and Tamura [25]. They utilized the LCP duality theorem of Fukuda and Terlaky [26] (see Chapter 4). Csizmadia and Illés [16] extended this method to several flexible finite pivot rules. These variants of the criss-cross method solve LCPs with an arbitrary matrix. They either solve the primal LCP or give a dual solution, or detect that the algorithm may begin cycling (due to lack of sufficiency) and in this case, they give a polynomial size certificate for the lack of sufficiency. However, there is not yet any EP theorem for LCPs based on an interior point algorithm.

To sum up, our aim is to construct such interior point algorithms, that according to the duality theorem of LCPs in EP form either give a solution for the original LCP or for the dual LCP, or detect the lack of property $\mathcal{P}_*(\tilde{\kappa})$ (where $\tilde{\kappa}$ is an arbitrary large, but a priori given parameter), and give a polynomial sized certificate in all cases in polynomial time. Throughout the first part of this chapter, we assume that a feasible interior point of the LCP is known. According to Corollary 3.4, it means that the LCP has a solution. Thus in this situation the EP theorem has only two cases, either we find a solution of the LCP or detect the lack of $\mathcal{P}_*(\tilde{\kappa})$-property. After that we deal with the problem of the initial interior point and present an EP theorem for the general case.

We have to note that we speak about solutions, but these are only ε-optimal solutions, where ε is an arbitrary small positive number and it is an upper bound of the duality gap of an ε-optimal solution. We touch upon the question of the rounding procedure in the chapter Open problems.

The rest of the chapter is organized as follows. In Section 6.1 we describe the general idea of modified IPMs and then present the modification of three well-known IPMs: the long-step path-following, the affine scaling and the predictor-corrector algorithms. Section 6.2 addresses the question how the interior point assumption can be eliminated by using the embedding technique. For the ease of understanding and to be self contained, in each subsection of the algorithms we introduce the necessary results of the original papers [63], [37] and [66] first and then present the modified algorithms and their analysis.

The notations are the same as in the previous chapter, namely the current point is $(\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+$ and $(\Delta\mathbf{x}, \Delta\mathbf{s})$ is the corresponding Newton direction, the solution of system (3.11). The new point with step length θ is given by $(\mathbf{x}(\theta), \mathbf{s}(\theta)) = (\mathbf{x} + \theta\Delta\mathbf{x}, \mathbf{s} + \theta\Delta\mathbf{s})$. We use the notations for scaling defined in Section 3 by (3.12), where in the affine scaling algorithm for the purpose of scaling we have $\mu \equiv 1$, otherwise $\mu > 0$.

6.1 Interior point algorithms in EP form

Our goal is to modify interior point algorithms in such a way, that they either solve LCPs with any arbitrary matrix, or give a certificate, that the matrix is not $\mathcal{P}_*(\tilde{\kappa})$, where $\tilde{\kappa}$ is a given (arbitrary big) number. Potra et al. [66] gave the first interior point algorithm, where we do not need to know a priori the value of κ – it is enough to know that the matrix is $\mathcal{P}_*$. Their algorithm initially assumes that the matrix is $\mathcal{P}_*(1)$. At each iteration they check whether the new point is in the appropriate neighbourhood of the central path or not. In the latter case they double the value of κ. We use this idea in a modified way. The larger κ is, the worse the iteration complexity estimate becomes, thus we only take the necessary enlargement of κ (until it reaches $\tilde{\kappa}$). The inequality in the definition of $\mathcal{P}_*(\kappa)$-matrices gives the following lower bound on κ for any vector $\mathbf{x} \in \mathbb{R}^n$:

$$\kappa \geq \kappa(\mathbf{x}) = -\frac{1}{4}\frac{\mathbf{x}^T M \mathbf{x}}{\sum_{i\in\mathcal{I}_+} x_i (Mx)_i}.$$

In IPMs the $\mathcal{P}_*(\kappa)$ property needs to hold only for the actual Newton direction $\Delta\mathbf{x}$ in various ways, for example, this property ensures that with a certain step size the new iterate is in an appropriate neighbourhood of the central path and/or the complementarity gap is sufficiently reduced. Consequently, if the desired results do not hold with the current κ value, we update κ to the lower bound determined by the Newton direction $\Delta\mathbf{x}$, i.e.,

$$\kappa(\Delta\mathbf{x}) = -\frac{1}{4}\frac{\Delta\mathbf{x}^T\Delta\mathbf{s}}{\sum_{i\in\mathcal{I}_+}\Delta x_i \Delta s_i} \qquad (\Delta\mathbf{s} = M\Delta\mathbf{x}). \tag{6.1}$$

The following two lemmas are immediate consequences of the definition of $\mathcal{P}_*(\kappa)$ and $\mathcal{P}_*$-matrices.

Lemma 6.1 *Let M be a real $n\times n$ matrix. If there exists a vector $\mathbf{x} \in \mathbb{R}^n$ such that $\kappa(\mathbf{x}) > \tilde{\kappa}$, then the matrix M is not $\mathcal{P}_*(\tilde{\kappa})$ and $\mathbf{x}$ is a certificate for this fact.*

Lemma 6.2 *Let M be a real $n \times n$ matrix. If there exists a vector $\mathbf{x} \in \mathbb{R}^n$ such that $\mathcal{I}_+(\mathbf{x}) = \{i \in \mathcal{I} : x_i(Mx)_i > 0\} = \emptyset$, then the matrix M is not $\mathcal{P}_*$ and $\mathbf{x}$ is a certificate for this fact.*

Therefore, if there exists such a vector $\Delta\mathbf{x}$ for which $\mathcal{I}_+(\Delta\mathbf{x}) = \emptyset$, and thus $\kappa(\Delta\mathbf{x})$ is not defined, then the matrix M of the LCP is not a $\mathcal{P}_*$-matrix. In this case we stop the algorithm, and the output will be $\Delta\mathbf{x}$ as a certificate to prove that M is not a $\mathcal{P}_*$-matrix.

There is another point where IPMs may fail if the matrix of the LCP is not $\mathcal{P}_*$. If the matrix is not $\mathcal{P}_0$, then the Newton system may not have a solution, or the solution may not

be unique (see Corollary 3.10). If this is the case, then the actual point $(\mathbf{x}, \mathbf{s})$ is a certificate which proves that the matrix is not $\mathcal{P}_0$, so it is not $\mathcal{P}_*$ either.

To summarize, we make three tests in our algorithms. In each test the property of the LCP matrix M is examined indirectly. When we inquire about the existence and uniqueness of the solution of the Newton system, we check whether the matrix is $\mathcal{P}_0$ or not. When we test some properties of the new point, for example whether it is in the appropriate neighbourhood of the central path or not, we examine the $\mathcal{P}_*(\kappa)$ property for the current value of κ. Finally, if the $\kappa(\Delta\mathbf{x})$ value given by (6.1) is not defined, then the matrix is not $\mathcal{P}_*$. We note that at each step all the properties are checked locally, only for one vector of $\mathbb{R}^n$. Consequently, it is possible, that the matrix is neither a $\mathcal{P}_0$ nor a $\mathcal{P}_*$-matrix, but the algorithm does not discover it and solves the LCP in polynomial time, because those properties are true for the vectors $\mathbf{x}$ and $\Delta\mathbf{x}$ that were generated by the algorithm. It may also occur that the matrix is not $\mathcal{P}_*$, but the algorithm does not detect it. It only increases the value of κ if $\kappa < \kappa(\Delta\mathbf{x})$ and then it proceeds to the next iterate. This is the reason why we need the threshold $\tilde{\kappa}$ parameter that enables us to get a finite algorithm.

Hereafter we modify the following three popular IPMs:

- Long-step path-following algorithm [63]:
  The Newton direction is the solution of the system (3.11) with $\mathbf{a} = \mu\mathbf{e} - \mathbf{xs}$, where $\mu = \dfrac{\mathbf{x}^T\mathbf{s}}{n}$.
  The choice of $\mathbf{a}$ implies $\left\| \dfrac{\mathbf{a}}{\sqrt{\mathbf{xs}}} \right\|^2 = \mu \left\| \sqrt{\dfrac{\mathbf{xs}}{\mu}} - \sqrt{\dfrac{\mu}{\mathbf{xs}}} \right\|^2$.

- A family of affine scaling algorithms [37]:
  The Newton direction is the solution of the system (3.11) with $\mu = 1$ and $\mathbf{a} = -\dfrac{\mathbf{v}^{2\rho+2}}{\|\mathbf{v}^{2\rho}\|}$, where $\rho \geq 0$ is the degree of the algorithm.
  The choice of $\mathbf{a}$ implies $\left\| \dfrac{\mathbf{a}}{\sqrt{\mathbf{xs}}} \right\|^2 = \dfrac{\|\mathbf{v}^{2\rho+1}\|^2}{\|\mathbf{v}^{2\rho}\|^2}$.

- Predictor-corrector algorithm [66]:
  The predictor Newton direction is the solution of the system (3.11) with $\mathbf{a} = -\mathbf{xs}$ (the affine scaling direction with $\rho = 0$).
  The choice of $\mathbf{a}$ implies $\left\| \dfrac{\mathbf{a}}{\sqrt{\mathbf{xs}}} \right\|^2 = \mathbf{x}^T\mathbf{s}$;
  The corrector (centering) Newton direction is the solution of the system (3.11) with $\mathbf{a} = \mu\mathbf{e} - \mathbf{xs}$, where $\mu = \dfrac{\mathbf{x}^T\mathbf{s}}{n}$.
  The choice of $\mathbf{a}$ implies $\left\| \dfrac{\mathbf{a}}{\sqrt{\mathbf{xs}}} \right\|^2 = \mu \left\| \sqrt{\dfrac{\mathbf{xs}}{\mu}} - \sqrt{\dfrac{\mu}{\mathbf{xs}}} \right\|^2$.

The following lemma is our main tool to verify when the $\mathcal{P}_*(\kappa)$ property does not hold. Furthermore, the concerned vector $\Delta\mathbf{x}$ is a certificate, whose encoding size is polynomial when it is computed as the solution of the Newton system (3.11). We use this lemma during the analysis. The first statement is simply the negation of the definition. We point out in Lemma 6.7 that if Theorem 10.5 of [63] does not hold, then the second or the third statement is realized. Furthermore, we claim in Lemma 6.13 that if Lemma 4.3 of [45] does not hold, then the second statement is realized. Finally, we show in Lemma 6.22 and Lemma 6.23 that if Theorem 3.3 of [66] does not hold, then the second, the third or the fourth statement is realized.

Lemma 6.3 *Let M be a real $n \times n$ matrix, $\kappa \geq 0$ be a given parameter, and $(\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+$. If one of the following statements holds, then the matrix M is not a $\mathcal{P}_*(\kappa)$-matrix.*

1. *There exists a vector $\mathbf{y} \in \mathbb{R}^n$ such that*

$$(1+4\kappa) \sum_{i \in \mathcal{I}_+(\mathbf{y})} y_i w_i + \sum_{i \in \mathcal{I}_-(\mathbf{y})} y_i w_i < 0,$$

   where $\mathbf{w} = M\mathbf{y}$ and $\mathcal{I}_+(\mathbf{y}) = \{i \in I : y_i w_i > 0\}$, $\mathcal{I}_-(\mathbf{y}) = \{i \in I : y_i w_i < 0\}$.

2. *There exists a solution $(\Delta\mathbf{x}, \Delta\mathbf{s})$ of the system (3.11) such that*

$$\|\Delta\mathbf{x}\Delta\mathbf{s}\|_\infty > \frac{1+4\kappa}{4} \left\| \frac{\mathbf{a}}{\sqrt{\mathbf{xs}}} \right\|^2.$$

3. *There exists a solution $(\Delta\mathbf{x}, \Delta\mathbf{s})$ of the system (3.11) such that*

$$\max\left(\sum_{i \in \mathcal{I}_+} \Delta x_i \Delta s_i, \; -\sum_{i \in \mathcal{I}_-} \Delta x_i \Delta s_i \right) > \frac{1+4\kappa}{4} \left\| \frac{\mathbf{a}}{\sqrt{\mathbf{xs}}} \right\|^2.$$

4. *There exists a solution $(\Delta\mathbf{x}, \Delta\mathbf{s})$ of the system (3.11) such that*

$$\Delta\mathbf{x}^T \Delta\mathbf{s} < -\kappa \left\| \frac{\mathbf{a}}{\sqrt{\mathbf{xs}}} \right\|^2.$$

Proof. The first statement is the negation of the definition of $\mathcal{P}_*(\kappa)$ matrices. Now we prove that the first statement follows from the others. By Lemma 3.12, one has

$$\sum_{i \in \mathcal{I}_+} \Delta x_i \Delta s_i \leq \frac{1}{4} \left\| \frac{\mathbf{a}}{\sqrt{\mathbf{xs}}} \right\|^2, \tag{6.2}$$

so $\Delta x_i \Delta s_i \leq \| \mathbf{a}/\sqrt{\mathbf{xs}} \|^2/4$ for all $i \in \mathcal{I}_+$. Accordingly, if the inequality of the second statement holds, let $j \in \mathcal{I}$ such that $\|\Delta \mathbf{x}\, \Delta \mathbf{s}\|_\infty = |\Delta x_j\, \Delta s_j|$, then $j \in \mathcal{I}_-$, i.e., $\Delta x_j \Delta s_j < 0$. Therefore,

$$\begin{aligned}(1+4\kappa)\sum_{i\in\mathcal{I}_+}\Delta x_i\Delta s_i \;\; + \;\; \sum_{i\in\mathcal{I}_-}\Delta x_i\Delta s_i &\leq (1+4\kappa)\sum_{i\in\mathcal{I}_+}\Delta x_i\Delta s_i + \Delta x_j\Delta s_j \\ &< \;\; (1+4\kappa)\frac{1}{4}\left\|\frac{\mathbf{a}}{\sqrt{\mathbf{xs}}}\right\|^2 - \frac{1+4\kappa}{4}\left\|\frac{\mathbf{a}}{\sqrt{\mathbf{xs}}}\right\|^2 = 0. \qquad (6.3)\end{aligned}$$

This is the same as the first statement with $\mathbf{y} = \Delta \mathbf{x},\ \mathbf{w} = \Delta \mathbf{s}$.

From the assumption of statement 3 and inequality (6.2), the second term is greater in the maximum, hence one has

$$\sum_{i\in\mathcal{I}_-}\Delta x_i\Delta s_i < -\frac{1+4\kappa}{4}\left\|\frac{\mathbf{a}}{\sqrt{\mathbf{xs}}}\right\|^2.$$

Therefore, $(\Delta \mathbf{x}, \Delta \mathbf{s})$ satisfies inequality (6.3), so $\mathbf{y} = \Delta \mathbf{x},\ \mathbf{w} = \Delta \mathbf{s}$ proves that the first statement holds.

The proof of the last statement, by using inequality (6.2) comes from the following inequality

$$\begin{aligned}(1+4\kappa)\sum_{i\in\mathcal{I}_+}\Delta x_i\Delta s_i + \sum_{i\in\mathcal{I}_-}\Delta x_i\Delta s_i \;\; &= \;\; 4\kappa\sum_{i\in\mathcal{I}_+}\Delta x_i\Delta s_i + \sum_{i\in\mathcal{I}}\Delta x_i\Delta s_i \\ &< \;\; \kappa\left\|\frac{\mathbf{a}}{\sqrt{\mathbf{xs}}}\right\|^2 - \kappa\left\|\frac{\mathbf{a}}{\sqrt{\mathbf{xs}}}\right\|^2 = 0,\end{aligned}$$

where we can use $\mathbf{y} = \Delta \mathbf{x},\ \mathbf{w} = \Delta \mathbf{s}$ again to get the first statement. ■

6.1.1 The long-step path-following interior point algorithm

In this section we deal with the algorithm proposed in [63]. The long-step algorithm has two loops. In the inner loop one takes steps towards the central path and in the outer loop the parameter μ is updated. The original algorithm for LCPs with $\mathcal{P}_*(\kappa)$-matrix is as follows:

Long-step path-following IPM for $\mathcal{P}_*(\kappa)$-matrix LCPs

---

Input:

a proximity parameter $\tau \geq 2$;
an accuracy parameter $\varepsilon > 0$;

a damping parameter $\theta > 0$;
a fixed barrier update parameter $\gamma \in (0,1)$;
an initial point $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{F}^+$, and $\mu^0 > 0$ such that $\delta_c(\mathbf{x}^0\mathbf{s}^0, \mu^0) \leq \tau$.

begin
$\mathbf{x} := \mathbf{x}^0$, $\mathbf{s} := \mathbf{s}^0$, $\mu := \mu^0$;
while $\mathbf{x}^T\mathbf{s} \geq \varepsilon$ **do**
$\mu = (1-\gamma)\mu$;
while $\delta_c(\mathbf{xs}, \mu) \geq \tau$ **do**
calculate the Newton direction $(\Delta\mathbf{x}, \Delta\mathbf{s})$ by solving (3.11) with $\mathbf{a} = \mu\mathbf{e} - \mathbf{xs}$;
$\mathbf{x} = \mathbf{x} + \theta\Delta\mathbf{x}$, $\mathbf{s} = \mathbf{s} + \theta\Delta\mathbf{s}$;
end
end
end.

We use the notations of [63]:

$$\sigma_+ = \frac{1}{\mu}\sum_{i\in\mathcal{I}_+} \Delta x_i \Delta s_i, \qquad \sigma_- = -\frac{1}{\mu}\sum_{i\in\mathcal{I}_-} \Delta x_i \Delta s_i, \qquad \sigma = \max(\sigma_+, \sigma_-).$$

Furthermore, let

$$\theta^*_\ell = \theta^*_\ell(\kappa) := \frac{2}{(1+4\kappa)\delta_c^2(\mathbf{xs}, \mu)}.$$

To simplify the notation we often write δ and δ^* instead of $\delta_c(\mathbf{xs}, \mu)$, $\delta_c(\mathbf{x}(\theta^*_\ell)\mathbf{x}(\theta^*_\ell), \mu)$, respectively.

For self-containedness, here we present those results from [63] that are needed for our developments. All theorems are converted to our notation.

The following three theorems are the main steps of the complexity analysis of the above algorithm. Due to the definition of σ, it is easy to see that the new point is certainly strictly feasible, if the step length θ is less than $1/\sigma$. First, an upper bound on the proximity measure of the new point is given.

Theorem 6.4 (Theorem 10.2 in [63]) *Let M be an arbitrary matrix and $(\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+$. Furthermore, let $(\Delta\mathbf{x}, \Delta\mathbf{s})$ be the Newton direction of the long-step path-following algorithm and $\delta^+ := \delta_c(\mathbf{x}(\theta)\mathbf{s}(\theta), \mu)$. Then for all $0 \leq \theta \leq 1/\sigma$, one has*

$$\delta^+ \leq (1-\theta)\delta^2 + \frac{2\,\theta^3\sigma^2}{1-\theta^2\sigma^2}.$$

When the matrix of the LCP is a $\mathcal{P}_*(\kappa)$-matrix with a known κ, then we can give a feasible step length, namely the step length θ_ℓ^* defines a feasible point because it is less than $1/\sigma$. Furthermore, in this case the decrease of the proximity measure has the following lower bound.

Theorem 6.5 (Theorem 10.5 in [63]) *Let $M \in \mathcal{P}_*(\kappa)$, $(\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+$ and $(\Delta\mathbf{x}, \Delta\mathbf{s})$ be the Newton direction of the long-step path-following algorithm. Then*

$$(\delta^*)^2 - \delta^2 \leq -\frac{5}{3(1+4\kappa)}. \tag{6.4}$$

Based on the above estimation, an upper bound on the number of inner iterations can be determined between two updates of the parameter μ. The number of outer iterations is well known from the theory of the long-step IPMs, it is $\lceil 1/\gamma \log\left((\mathbf{x}^0)^T \mathbf{s}^0/\varepsilon \right) \rceil$. Multiplication of these bounds provides the following complexity result.

Theorem 6.6 (From Theorem 10.10 and the subsequent remarks in [63]) *Let the matrix $M \in \mathcal{P}_*(\kappa)$, $\tau = 2$, $\gamma = 1/2$ and $(\mathbf{x}^0, \mathbf{s}^0)$ be a feasible interior point such that $\delta_c(\mathbf{x}^0\mathbf{s}^0, \mu^0) \leq \tau$. Then the long-step path-following algorithm produces a point $(\hat{\mathbf{x}}, \hat{\mathbf{s}})$ such that $\delta_c(\hat{\mathbf{x}}\hat{\mathbf{s}}, \hat{\mu}) \leq \tau$ and $\hat{\mathbf{x}}^T\hat{\mathbf{s}} \leq \varepsilon$ in at most*

$$\mathcal{O}\left((1+4\kappa) n \log \frac{(\mathbf{x}^0)^T\mathbf{s}^0}{\varepsilon} \right) \quad \textit{iterations.}$$

Now, after introducing the original algorithm, let us discuss the modified algorithm. In this algorithm we check the decrease of the centrality measure after one inner step and if it is too small, then κ is updated by (6.1), or a certificate is obtained showing that M is not a $\mathcal{P}_*(\tilde{\kappa})$ matrix. As stated in the previous subsection, if $\kappa(\Delta\mathbf{x})$ is not defined, then the matrix is not $\mathcal{P}_*$ and $\Delta\mathbf{x}$ is a certificate for it. Furthermore, if $\kappa(\Delta\mathbf{x}) > \tilde{\kappa}$, then matrix M is not $\mathcal{P}_*(\tilde{\kappa})$ and the Newton direction $\Delta\mathbf{x}$ is a certificate for this fact. The modified algorithm is as follows:

Modified long-step path-following IPM

Input:

an upper bound $\tilde{\kappa} > 0$ on the value of κ;
a proximity parameter $\tau \geq 2$;
an accuracy parameter $\varepsilon > 0$;

a fixed barrier update parameter $\gamma \in (0,1)$;
an initial point $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{F}^+$, and $\mu^0 > 0$ such that $\delta_c(\mathbf{x}^0\mathbf{s}^0, \mu^0) \leq \tau$.

begin
$\mathbf{x} := \mathbf{x}^0$, $\mathbf{s} := \mathbf{s}^0$, $\mu := \mu^0$, $\kappa := 0$;
while $\mathbf{x}^T\mathbf{s} \geq \varepsilon$ **do**
$\mu = (1-\gamma)\mu$;
while $\delta_c(\mathbf{xs}, \mu) \geq \tau$ **do**
calculate the Newton direction $(\Delta\mathbf{x}, \Delta\mathbf{s})$ by solving (3.11) with $\mathbf{a} = \mu\mathbf{e} - \mathbf{xs}$;
if (the Newton direction does not exist or it is not unique) **then**
return the matrix is not $\mathcal{P}_0$; *% see Corollary 3.10*
$\bar{\theta} = \operatorname{argmin}\{\delta_c(\mathbf{x}(\theta)\mathbf{s}(\theta), \mu) : \ (\mathbf{x}(\theta), \mathbf{s}(\theta)) > \mathbf{0}\}$;
if $\left(\delta_c^2(\mathbf{xs}, \mu) - \delta_c^2(\mathbf{x}(\bar{\theta})\mathbf{s}(\bar{\theta}), \mu) < \frac{5}{3(1+4\kappa)}\right)$ **then**
calculate $\kappa(\Delta\mathbf{x})$; *% see (6.1)*
if ($\kappa(\Delta\mathbf{x})$ is not defined) **then**
return the matrix is not $\mathcal{P}_*$; *% see Lemma 6.2*
if ($\kappa(\Delta\mathbf{x}) > \tilde{\kappa}$) **then**
return the matrix is not $\mathcal{P}_*(\tilde{\kappa})$; *% see Lemma 6.1*
$\kappa = \kappa(\Delta\mathbf{x})$;
$\mathbf{x} = \mathbf{x}(\bar{\theta}), \ \mathbf{s} = \mathbf{s}(\bar{\theta})$;
end
end
end.

Peng et al. [63] proved that for $\mathcal{P}_*(\kappa)$ LCPs the step length θ_ℓ^* is feasible, and taking this step, the decrease of the proximity measure is sufficient to ensure the polynomiality of the algorithm (see the above theorems). The following lemma shows that if the specified sufficiently large decrease does not take place, then the matrix of the problem is not $\mathcal{P}_*(\kappa)$.

Lemma 6.7 *If after an inner iteration the decrease of the proximity is not sufficient, i.e.,* $\delta^2(\mathbf{xs}, \mu) - \delta^2(\mathbf{x}(\bar{\theta})\mathbf{s}(\bar{\theta}), \mu) < \frac{5}{3(1+4\kappa)}$, *then the matrix of the LCP is not* $\mathcal{P}_*(\kappa)$ *with the actual* κ *value, and the Newton direction* $\Delta\mathbf{x}$ *is a certificate for this fact.*

Proof. By Theorem 6.5, if the matrix is $\mathcal{P}_*(\kappa)$ we achieve the sufficient decrease of the centrality measure with step length θ_ℓ^*. Therefore, if the maximum decrease is smaller, then

either $(\mathbf{x}^*, \mathbf{s}^*)$ is not feasible or the decrease of the proximity with step size θ_ℓ^* is not sufficient, i.e., $\delta^2(\mathbf{xs}, \mu) - \delta^2(\mathbf{x}^*\mathbf{s}^*, \mu) < \frac{5}{3(1+4\kappa)}$. We prove in both cases that the matrix of the problem is not $\mathcal{P}_*(\kappa)$ and $\Delta\mathbf{x}$ is a certificate for it.

If the point $(\mathbf{x}^*, \mathbf{s}^*)$ is not feasible, then $\mu\mathbf{e} + \theta_\ell^*\Delta\mathbf{x}\Delta\mathbf{s} \not> 0$, so there exists such an index k, that $\mu + \theta_\ell^*\Delta x_k \Delta s_k \leq 0$. It means, that $\Delta x_k \Delta s_k \leq -\mu/\theta_\ell^* = -\frac{\mu}{2}(1+4\kappa)\delta^2 < 0$. Since $(\Delta\mathbf{x}, \Delta\mathbf{s})$ is a solution of the system (3.11) with $\mathbf{a} = \mu\mathbf{e} - \mathbf{xs}$, therefore

$$\|\Delta\mathbf{x}\Delta\mathbf{s}\|_\infty > \frac{1+4\kappa}{4}\,\mu\,\delta^2 = \frac{1+4\kappa}{4}\left\|\frac{\mathbf{a}}{\sqrt{\mathbf{xs}}}\right\|^2,$$

but this contradicts the $\mathcal{P}_*(\kappa)$ property by the second statement of Lemma 6.3.

Now let us analyze the case when the decrease of the proximity measure is not sufficient with step length θ_ℓ^*. According to the condition of Theorem 6.4, let us consider the cases $\theta_\ell^* < 1/\sigma$ and $\theta_\ell^* \geq 1/\sigma$ separately. If $\theta_\ell^* < 1/\sigma$, by the definition of θ_ℓ^* and Theorem 6.4, one has

$$(\delta^*)^2 - \delta^2 \leq -\frac{2}{1+4\kappa} + \frac{2(\theta_\ell^*)^3\sigma^2}{1-(\theta_\ell^*)^2\sigma^2}. \tag{6.5}$$

Since $\delta \geq 2$, we can write

$$-\frac{2}{1+4\kappa} + \frac{4}{3(1+4\kappa)\delta^2} \leq -\frac{5}{3(1+4\kappa)}. \tag{6.6}$$

Therefore, by inequalities (6.5) and (6.6) and by the assumption of the lemma, the following inequalities hold

$$-\frac{2}{1+4\kappa} + \frac{4}{3(1+4\kappa)\delta^2} \leq -\frac{5}{3(1+4\kappa)} < (\delta^*)^2 - \delta^2 \leq -\frac{2}{1+4\kappa} + \frac{2(\theta_\ell^*)^3\sigma^2}{1-(\theta_\ell^*)^2\sigma^2}.$$

Using the definition of θ_ℓ^* we get

$$\frac{4}{3(1+4\kappa)\delta^2} < \frac{2(\theta_\ell^*)^2\sigma^2}{1-(\theta_\ell^*)^2\sigma^2}\;\frac{2}{(1+4\kappa)\delta^2}.$$

After reordering one has $\frac{1}{2} < \theta_\ell^*\sigma$. Substituting the definition of θ_ℓ^* we get the following lower bound on σ

$$\max(\sigma_+, \sigma_-) = \sigma > \frac{1+4\kappa}{4}\,\delta^2. \tag{6.7}$$

By the definitions of σ and the proximity measure, one has

$$\max\left(\sum_{i\in\mathcal{I}_+}\Delta x_i\Delta s_i, -\sum_{i\in\mathcal{I}_-}\Delta x_i\Delta s_i\right) > \frac{1+4\kappa}{4}\,\mu\,\delta^2 = \frac{1+4\kappa}{4}\left\|\frac{\mu\mathbf{e}-\mathbf{xs}}{\sqrt{\mathbf{xs}}}\right\|^2.$$

By the third statement of Lemma 6.3 this implies that matrix M is not $\mathcal{P}_*(\kappa)$ and the vector $\Delta\mathbf{x}$ is a certificate for it.

If $\theta_\ell^* \geq 1/\sigma$, then by the definition of θ_ℓ^* one has $\sigma \geq (1+4\kappa)\delta^2/2$, therefore inequality (6.7) holds, so the lemma is true in this case, too. ■

The following lemma proves, that the long-step path-following IPM is well defined.

Lemma 6.8 *At each iteration when the value of κ is updated, the new value of κ satisfies the inequality $\delta^2(\mathbf{xs},\mu) - \delta^2(\mathbf{x}(\bar{\theta})\mathbf{s}(\bar{\theta}),\mu) \geq \frac{5}{3(1+4\kappa)}$.*

Proof. In the proof of Theorem 6.5 we use the $\mathcal{P}_*(\kappa)$ property only for the vector $\Delta\mathbf{x}$. When parameter κ is updated, we choose the new value such that the inequality in the definition of $\mathcal{P}_*(\kappa)$-matrices (2.1) holds for the vector $\Delta\mathbf{x}$. Therefore, the new point defined by the updated value of step size θ_ℓ^* is strictly feasible and $\delta^2(\mathbf{xs},\mu) - \delta^2(\mathbf{x}^*\mathbf{s}^*,\mu)) \geq \frac{5}{3(1+4\kappa)}$. Thus the new value of θ_ℓ^* is considered in the definition of $\bar{\theta}$ as $\delta^2(\mathbf{xs},\mu) - \delta^2(\mathbf{x}(\bar{\theta})\mathbf{s}(\bar{\theta}),\mu) \geq \delta^2(\mathbf{xs},\mu) - \delta^2(\mathbf{x}^*\mathbf{s}^*,\mu) \geq \frac{5}{3(1+4\kappa)}$. ■

Now we are ready to state the complexity result for the modified long-step path-following interior point algorithm for a general LCP in case an initial interior point is given.

Theorem 6.9 *Let $\tau = 2$, $\gamma = 1/2$ and $(\mathbf{x}^0,\mathbf{s}^0)$ be a feasible interior point such that $\delta_c(\mathbf{x}^0\mathbf{s}^0,\mu^0) \leq \tau$. Then after at most $\mathcal{O}\left((1+4\hat{\kappa})n\log\frac{(\mathbf{x}^0)^T\mathbf{s}^0}{\varepsilon}\right)$ steps, where $\hat{\kappa} \leq \tilde{\kappa}$ is the largest value of parameter κ throughout the algorithm, the long-step path-following interior point algorithm either produces a point $(\hat{\mathbf{x}},\hat{\mathbf{s}})$ such that $\hat{\mathbf{x}}^T\hat{\mathbf{s}} \leq \varepsilon$ and $\delta_c(\hat{\mathbf{x}}\hat{\mathbf{s}},\hat{\mu}) \leq \tau$ or it gives a certificate that the matrix of the LCP is not $\mathcal{P}_*(\tilde{\kappa})$.*

Proof. The algorithm at each iteration either takes a step, or detects that the matrix is not $\mathcal{P}_*(\tilde{\kappa})$ and stops. If we take a Newton step, then by the definition of the algorithm and by Lemma 6.8 the decrease of the squared proximity measure is at least $5/[3(1+4\kappa)]$. We can see that a larger κ means a smaller lower bound on the decrease of the proximity measure. Therefore, if the algorithm stops with an ε-optimal solution, then after each Newton step the decrease of the squared proximity measure is at least $5/[3(1+4\hat{\kappa})]$. Thus at each outer iteration we take at most as many inner iterations as in the original long-step algorithm with a $\mathcal{P}_*(\hat{\kappa})$-matrix is done, or the algorithm stops earlier with a certificate that M is not a $\mathcal{P}_*(\tilde{\kappa})$-matrix. By the complexity theorem of the original algorithm (see Theorem 6.6) we proved our statement. ■

6.1.2 The affine scaling interior point algorithm

Now we deal with affine scaling IPMs. We modify the family of the algorithms proposed in [37], where the particular algorithms correspond to the degree $\rho \geq 0$ of the algorithm, where

$\rho = 0$ gives the classical primal-dual affine scaling algorithm, while $\rho = 1$ gives the primal-dual Dikin affine scaling algorithm [19]. Furthermore, there is a step length parameter ν, that depends on the degree ρ (defined among the inputs of the algorithm), and $\mu \equiv 1$ in scaling (3.12). The original affine scaling algorithm for LCPs with $\mathcal{P}_*(\kappa)$-matrix is as follows:

Affine scaling IPM for $\mathcal{P}_*(\kappa)$-matrix LCPs

Input:

an accuracy parameter $\varepsilon > 0$;
a centrality parameter τ;
the degree of scaling $\rho > 0$;
a step size $\theta > 0$;
a strictly feasible initial point $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{F}^+$ such that $\delta_a(\mathbf{x}^0\,\mathbf{s}^0) \leq \tau$;

$$\nu := \begin{cases} 2/\sqrt{n}, & \text{if } 0 < \rho \leq 1 \\ 2\tau^{2-2\rho}/\sqrt{n}, & \text{if } 1 \leq \rho. \end{cases}$$

begin
$\mathbf{x} := \mathbf{x}^0$, $\mathbf{s} := \mathbf{s}^0$;
while $\mathbf{x}^T\mathbf{s} \geq \varepsilon$ **do**
calculate the Newton direction $(\Delta\mathbf{x}, \Delta\mathbf{s})$ by solving (3.11) with $\mathbf{a} = -\mathbf{v}^{2\rho+2}/\|\mathbf{v}^{2\rho}\|$;
$\mathbf{x} = \mathbf{x} + \theta\Delta\mathbf{x}$, $\mathbf{s} = \mathbf{s} + \theta\Delta\mathbf{s}$;
end
end.

We will use the following results of [45, 37] in the analysis of the modified affine scaling IPM. Jansen et al. gave an upper bound on the complementarity gap of the new point. This upper bound is true for all matrices in general, because only an upper bound is used on $\Delta\mathbf{x}^T\Delta\mathbf{s}$, which is independent of κ, see Lemma 3.12.

Lemma 6.10 (Lemma 4.3 in [45]) *Let M be an arbitrary real matrix, $\delta_a(\mathbf{xs}) < \tau$ and $(\Delta\mathbf{x}, \Delta\mathbf{s})$ is the affine scaling direction.*
(i) If $0 \leq \rho \leq 1$ and $\theta \leq 2/\sqrt{n}$, then

$$\mathbf{x}(\theta)^T\mathbf{s}(\theta) \leq \left(1 - \frac{\theta}{2\sqrt{n}}\right)\|\mathbf{v}\|^2.$$

(ii) If $1 \leq \rho$ and $\theta \leq 2\tau^{2-2\rho}/\sqrt{n}$, then

$$\mathbf{x}(\theta)^T\mathbf{s}(\theta) \leq \left(1 - \frac{\theta\tau^{2-2\rho}}{2\sqrt{n}}\right)\|\mathbf{v}\|^2.$$

Illés et al. set a feasible interval of step lengths, namely step lengths which determine strictly feasible points in the τ neighbourhood of the central path according to δ_a proximity measure. The first term in the minimum is due to the feasibility of the new point, while the third is due to the neighbourhood restriction. The second term is a kind of technical bound.

Theorem 6.11 (Theorem 6.1 in [37]) *Let M be a $\mathcal{P}_*(\kappa)$-matrix, $\rho > 0$, $\tau > 1$ and let $(\Delta\mathbf{x}, \Delta\mathbf{s})$ be the affine scaling direction. If $(\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+$, $\delta_a(\mathbf{xs}) \leq \tau$ and*

$$0 \leq \theta \leq \min\left\{ \frac{2}{(1+4\kappa)\tau}\left(\sqrt{1+4\kappa+\frac{1}{\tau^2 n}} - \frac{1}{\tau\sqrt{n}}\right), \frac{\sqrt{n}}{(\rho+1)\tau^{2\rho}}, \frac{4(\tau^{2\rho}-1)}{(1+4\kappa)(1+\tau^2)\tau^{2\rho}\sqrt{n}} \right\},$$

then $(\mathbf{x}(\theta), \mathbf{s}(\theta))$ is strictly feasible and $\delta_a(\mathbf{x}(\theta)\mathbf{s}(\theta)) \leq \tau$.

The last referred result in connection with affine scaling IPMs for LCPs with $\mathcal{P}_*(\kappa)$-matrices is the complexity of the algorithm. This is a direct consequence of the two previous statements, Lemma 6.10 and Theorem 6.11.

Theorem 6.12 (Corollary 6.1 in [37]) *Let $M \in \mathcal{P}_*(\kappa)$ and $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{F}^+$ such that $\delta_a(\mathbf{x}^0\mathbf{s}^0) \leq \tau = \sqrt{2}$.*

- *If $0 < \rho \leq 1$ and $n \geq 4$, then we may choose $\theta = \frac{4(1-2^{-\rho})}{3(1+4\kappa)\sqrt{n}}$, hence the complexity of the affine scaling algorithm is $\mathcal{O}\left(\frac{n(1+4\kappa)}{1-2^{-\rho}} \log \frac{(\mathbf{x}^0)^T\mathbf{s}^0}{\varepsilon}\right)$.*
- *If $\rho = 1$ and $n \geq 4$, then we may choose $\theta = \frac{1}{2(1+4\kappa)\sqrt{n}}$, hence the complexity of the affine scaling algorithm is $\mathcal{O}\left(n(1+4\kappa) \log \frac{(\mathbf{x}^0)^T\mathbf{s}^0}{\varepsilon}\right)$.*
- *If $\rho > 1$ and n is sufficiently large, then we may choose $\theta = \frac{1}{2^\rho(1+4\kappa)\sqrt{n}}$, hence the complexity of the affine scaling algorithm is $\mathcal{O}\left(2^{2\rho-2} n(1+4\kappa) \log \frac{(\mathbf{x}^0)^T\mathbf{s}^0}{\varepsilon}\right)$.*

Similarly to the long-step path following IPMs in the previous subsection, we check the solvability and uniqueness of the Newton system in each iteration. Furthermore, here we verify the sufficient decrease not of the proximity measure, but of the complementarity gap after each step. For the actual value of κ we determine $\theta_a^*(\kappa)$, which is a theoretical lower bound for the maximal feasible step length in the specified neighbourhood if the matrix M satisfies the $\mathcal{P}_*(\kappa)$ property. Therefore, if after a step the decrease of the complementarity gap is not large enough, it means that the matrix M is not $\mathcal{P}_*(\kappa)$ with the actual value of κ, so either we update κ or exit the algorithm with a corresponding certificate. If the new value of κ can not be defined by (6.1), then the matrix M is not $\mathcal{P}_*$, so we stop and the Newton direction $\Delta\mathbf{x}$ is a certificate. If the new value of κ is larger than $\tilde{\kappa}$, then the matrix is not $\mathcal{P}_*(\tilde{\kappa})$, therefore the algorithm stops as well and $\Delta\mathbf{x}$ is a certificate. In the rest of this

subsection we consider the case $\rho > 0$. The modified algorithm is as follows.

Modified affine scaling IPM

Input:

an upper bound $\tilde{\kappa} > 0$ on the value of κ;
an accuracy parameter $\varepsilon > 0$;
a centrality parameter τ;
the degree of scaling $\rho > 0$;
a strictly feasible initial point $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{F}^+$ such that $\delta_a(\mathbf{x}^0\,\mathbf{s}^0) \le \tau$;

$$\nu := \begin{cases} 2/\sqrt{n}, & \text{if } 0 < \rho \le 1 \\ 2\tau^{2-2\rho}/\sqrt{n}, & \text{if } 1 \le \rho; \end{cases}$$

$$\theta_a^*(\kappa) := \min\left\{ \frac{2}{(1+4\kappa)\tau}\left(\sqrt{1+4\kappa+\frac{1}{\tau^2\,n}} - \frac{1}{\tau\sqrt{n}}\right), \frac{\sqrt{n}}{(r+1)\tau^{2\rho}}, \frac{4(\tau^{2\rho}-1)}{(1+4\kappa)(1+\tau^2)\tau^{2\rho}\sqrt{n}}, \nu \right\}.$$

begin
$\mathbf{x} := \mathbf{x}^0$, $\mathbf{s} := \mathbf{s}^0$, $\kappa := 0$;
while $\mathbf{x}^T\mathbf{s} \ge \varepsilon$ **do**
calculate the Newton direction $(\Delta\mathbf{x}, \Delta\mathbf{s})$ by solving (3.11) with $\mathbf{a} = -\mathbf{v}^{2\rho+2}/\|\mathbf{v}^{2\rho}\|$;
if (the Newton direction does not exist or it is not unique) **then**
return the matrix is not $\mathcal{P}_0$; *% see Corollary 3.10*
$\bar{\theta} = \operatorname{argmin}\left\{\mathbf{x}(\theta)^{\mathrm{T}}\mathbf{s}(\theta) : \ \delta_{\mathrm{a}}\left(\mathbf{x}(\theta), \mathbf{s}(\theta)\right) \le \tau, \ (\mathbf{x}(\theta), \mathbf{s}(\theta)) \ge \mathbf{0}\right\}$;
if $\left(\mathbf{x}(\bar{\theta})^T\mathbf{s}(\bar{\theta}) > (1 - 0.25\,\nu\,\theta_a^*(\kappa))\,\mathbf{x}^T\mathbf{s}\right)$ **then**
calculate $\kappa(\Delta\mathbf{x})$; *% see (6.1)*
if ($\kappa(\Delta\mathbf{x})$ is not defined) **then**
return the matrix is not $\mathcal{P}_*$; *% see Lemma 6.2*
if ($\kappa(\Delta\mathbf{x}) > \tilde{\kappa}$) **then**
return the matrix is not $\mathcal{P}_*(\tilde{\kappa})$; *% see Lemma 6.1*
$\kappa = \kappa(\Delta\mathbf{x})$;
update $\theta_a^*(\kappa)$; *% it depends on κ*
$\mathbf{x} = \mathbf{x}(\bar{\theta})$, $\mathbf{s} = \mathbf{s}(\bar{\theta})$;
end
end.

Illés et al. proved [37] that if the matrix M is a $\mathcal{P}_*(\kappa)$-matrix, then the step length $\theta_a^*(\kappa)$ is feasible; with this step size the new iterate stays within the specified neighbourhood and

it provides the required decrease of the complementarity gap. The following lemma shows that if the decrease of the complementarity gap is not sufficient, then the matrix M does not belong to the class of $\mathcal{P}_*(\kappa)$-matrices.

Lemma 6.13 *If* $\mathbf{x}(\bar{\theta})^T\mathbf{s}(\bar{\theta}) > (1-0.25\,\nu\,\theta_a^*(\kappa))\,\mathbf{x}^T\mathbf{s}$, *that is, the decrease of the complementarity gap within the* $\delta_a \leq \tau$ *neighbourhood is not sufficient, then the matrix* M *of the LCP is not* $\mathcal{P}_*(\kappa)$ *with the actual value of* κ. *The Newton direction* $\Delta\mathbf{x}$ *serves as a certificate.*

Proof. Based on Lemma 6.10 the complementarity gap at $\theta_a^*(\kappa)$ is smaller than $(1-0.25\,\nu\,\theta_a^*(\kappa))\,\mathbf{x}^T\mathbf{s}$, furthermore by Theorem 6.11, if M is a $\mathcal{P}_*(\kappa)$-matrix, then the point $(\mathbf{x}^*,\,\mathbf{s}^*) = (\mathbf{x}(\theta_a^*(\kappa)), \mathbf{s}(\theta_a^*(\kappa)))$ is feasible. Therefore, if $\mathbf{x}(\bar{\theta})^T\mathbf{s}(\bar{\theta}) > (1-0.25\,\nu\,\theta_a^*(\kappa))\,\mathbf{x}^T\mathbf{s}$, then because the step length $\theta_a^*(\kappa)$ is not considered in the definition of $\bar{\theta}$ (see the affine scaling algorithm), either $(\mathbf{x}^*,\mathbf{s}^*)$ is not feasible, or this point is not in the τ neighbourhood of the central path, namely $\delta_a(\mathbf{x}^*\mathbf{s}^*) > \tau$. We show that both cases imply that the matrix M is not $\mathcal{P}_*(\kappa)$ with the actual κ value.

Let us denote the first three terms in the definition of $\theta_a^*(\kappa)$ by θ_1, θ_2, and θ_3, respectively. We follow the proof of Theorem 6.1 in [37] (see Theorem 6.11). We need to reconsider only the expressions depending on κ. Therefore, the function $\varphi(t) = t - \theta\frac{t^{\rho+1}}{\|\mathbf{v}^{2\rho}\|}$ remains monotonically increasing for $\theta \leq \theta_2$, and there exist positive constants α and β such that $\frac{\beta}{\alpha} = \tau^2$ and $\alpha\mathbf{e} \leq \mathbf{v}^2 \leq \beta\mathbf{e}$. Additionally, inequalities (17) in [37] hold for $\theta \leq \theta_2$, thus for $\theta_a^*(\kappa)$, too:

$$\min(\mathbf{v}^{*2}) \geq \alpha - \theta_a^*(\kappa)\frac{\alpha^{\rho+1}}{\|\mathbf{v}^{2\rho}\|} - (\theta_a^*(\kappa))^2\|\mathbf{r}\|_\infty, \tag{6.8}$$

$$\max(\mathbf{v}^{*2}) \leq \beta - \theta_a^*(\kappa)\frac{\beta^{\rho+1}}{\|\mathbf{v}^{2\rho}\|} + (\theta_a^*(\kappa))^2\|\mathbf{r}\|_\infty, \tag{6.9}$$

where $\mathbf{r}$ is defined by (3.12).

Let us first consider the case $\delta_a(\mathbf{x}^*\mathbf{s}^*) > \tau$, i.e., $\max(\mathbf{x}^*\mathbf{s}^*) > \tau^2\min(\mathbf{x}^*\mathbf{s}^*)$. From the inequalities (6.8) and (6.9) one has

$$\tau^2\left(\alpha - \theta_a^*(\kappa)\frac{\alpha^{\rho+1}}{\|\mathbf{v}^{2\rho}\|} - (\theta_a^*(\kappa))^2\|\mathbf{r}\|_\infty\right) < \beta - \theta_a^*(\kappa)\frac{\beta^{\rho+1}}{\|\mathbf{v}^{2\rho}\|} + (\theta_a^*(\kappa))^2\|\mathbf{r}\|_\infty,$$

$$\frac{\beta^\rho - \alpha^\rho}{\|\mathbf{v}^{2\rho}\|} < \theta_a^*(\kappa)\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)\|\mathbf{r}\|_\infty. \tag{6.10}$$

If $\theta_a^*(\kappa)$ is substituted by $\theta_3 = \frac{4(\tau^{2\rho}-1)}{(1+4\kappa)(1+\tau^2)\tau^{2\rho}\sqrt{n}}$, the right hand side of inequality (6.10) increases, so the inequality is still true. After substitution one has

$$\frac{1+4\kappa}{4}\,\beta < \|\mathbf{r}\|_\infty. \tag{6.11}$$

Since $\mathbf{v}^2 \leq \beta \mathbf{e}$ and $\mathbf{r} = \Delta\mathbf{x}\Delta\mathbf{s}$, inequality (6.11) gives

$$\|\mathbf{v}\|_\infty^2 \leq \beta < \frac{4}{1+4\kappa}\,\|\mathbf{r}\|_\infty = \frac{4}{1+4\kappa}\,\|\Delta\mathbf{x}\Delta\mathbf{s}\|_\infty. \tag{6.12}$$

One can check that

$$\|\mathbf{v}^{2\rho+1}\|^2 \leq \|\mathbf{v}\|_\infty^2 \|\mathbf{v}^{2\rho}\|^2. \tag{6.13}$$

Since $(\Delta\mathbf{x}, \Delta\mathbf{s})$ is the solution of the system (3.11) with $\mathbf{a} = -\mathbf{v}^{2\rho+2}/\|\mathbf{v}^{2\rho}\|$, by inequalities (6.12) and (6.13) we have

$$\left\| \frac{\mathbf{a}}{\sqrt{\mathbf{xs}}} \right\|^2 = \left\| \frac{\mathbf{v}^{2\rho+1}}{\|\mathbf{v}^{2\rho}\|} \right\|^2 \leq \|\mathbf{v}\|_\infty^2.$$

Therefore, by the second statement of Lemma 6.3, we get that inequality (6.12) contradicts the $\mathcal{P}_*(\kappa)$ property and vector $\Delta\mathbf{x}$ is a certificate for this fact.

Now we consider the case when $(\mathbf{x}^*, \mathbf{s}^*)$ is not feasible, so there exists such an index i, that either $x_i^* < 0$ or $s_i^* < 0$. Let us consider the maximum feasible step size $\hat{\theta} < \theta_a^*(\kappa)$, for which $(\mathbf{x}(\hat{\theta}), \mathbf{s}(\hat{\theta})) = (\hat{\mathbf{x}}, \hat{\mathbf{s}}) \geq \mathbf{0}$ holds and at least one of its coordinates is 0. For this point $\hat{\mathbf{x}}\,\hat{\mathbf{s}} \neq \mathbf{0}$, else $\bar{\theta} = \hat{\theta}$ by the definition of $\bar{\theta}$, and the new point would be an exact solution, so the decrease of the complementarity gap would be $\mathbf{x}^T\mathbf{s}$ contradicting the assumption of the lemma. Therefore, $0 \neq \max(\hat{\mathbf{x}}\,\hat{\mathbf{s}}) > \tau^2 \min(\hat{\mathbf{x}}\,\hat{\mathbf{s}}) = 0$, so inequality (6.10) holds with $\hat{\theta}$. Because of $\theta_3 \geq \theta_a^*(\kappa) > \hat{\theta}$, inequality (6.11) holds as well, and as we have already seen, this means that the matrix M is not $\mathcal{P}_*(\kappa)$ and the vector $\Delta\mathbf{x}$ is a certificate for this fact. ■

The following lemma claims that the algorithm is well defined.

Lemma 6.14 *At each iteration, when the value of κ is updated, the new value of $\theta_a^*(\kappa)$ satisfies the inequality $\mathbf{x}(\bar{\theta})^T\mathbf{s}(\bar{\theta}) \leq (1 - 0.25\,\nu\,\theta_a^*(\kappa))\mathbf{x}^T\mathbf{s}$.*

Proof. In the proof of Theorem 6.11 we use the $\mathcal{P}_*(\kappa)$ property only for the vector $\Delta\mathbf{x}$. When parameter κ is updated, we choose the new value in such a way that the inequality in the definition of $\mathcal{P}_*(\kappa)$-matrices (2.1) would hold for vector $\Delta\mathbf{x}$. Therefore, the new point defined by the updated value of step size $\theta_a^*(\kappa)$ is feasible and it is in the τ-neighbourhood of the central path. Thus the new value of $\theta_a^*(\kappa)$ is considered in the definition of $\bar{\theta}$, so $\mathbf{x}(\bar{\theta})^T\mathbf{s}(\bar{\theta}) \leq (1 - 0.25\,\nu\,\theta_a^*(\kappa))\,\mathbf{x}^T\mathbf{s}$. ■

Now we are ready to state the complexity result for the modified affine scaling algorithm for general LCPs in case an initial interior point is given.

Theorem 6.15 *Let $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{F}^+$ such that $\delta_a(\mathbf{x}^0\mathbf{s}^0) \leq \tau = \sqrt{2}$. Then after at most*

$$\begin{cases} \mathcal{O}\left(\frac{n(1+4\hat{\kappa})}{1-2^{-\rho}} \log \frac{(\mathbf{x}^0)^T\mathbf{s}^0}{\varepsilon}\right), & \text{if } 0 < \rho \leq 1 \text{ and } n \geq 4 \\ \mathcal{O}\left(n(1+4\hat{\kappa}) \log \frac{(\mathbf{x}^0)^T\mathbf{s}^0}{\varepsilon}\right), & \text{if } \rho = 1 \text{ and } n \geq 4 \\ \mathcal{O}\left(2^{2\rho-2} n(1+4\hat{\kappa}) \log \frac{(\mathbf{x}^0)^T\mathbf{s}^0}{\varepsilon}\right), & \text{if } 1 < \rho \text{ and } n \text{ sufficiently large} \end{cases}$$

iterations the affine scaling algorithm either yields a vector $(\hat{\mathbf{x}}, \hat{\mathbf{s}})$ *such that* $\hat{\mathbf{x}}^T\hat{\mathbf{s}} \leq \varepsilon$ *and* $\delta_a(\hat{\mathbf{x}}\hat{\mathbf{s}}) \leq \tau$*, or it gives a polynomial size certificate that the matrix is not* $\mathcal{P}_*(\tilde{\kappa})$*, where* $\hat{\kappa} \leq \tilde{\kappa}$ *is the largest value of parameter* κ*.*

Proof. The algorithm at each iteration either takes a step, or detects that the matrix is not $\mathcal{P}_*(\tilde{\kappa})$ and stops. If we take a Newton step, then by the definition of the algorithm and by Lemma 6.14 the decrease of the complementarity gap is at least $0.25\,\nu\,\theta_a^*(\kappa)\,\mathbf{x}^T\mathbf{s}$. One can see from the definition of $\theta_a^*(\kappa)$ that a larger κ means a smaller $\theta_a^*(\kappa)$, so a smaller lower bound on the decrease of the complementarity gap. Therefore, if the algorithm stops with an ε-optimal solution, then each Newton step decreases the complementarity gap by more than $0.25\,\nu\,\theta_a^*(\tilde{\kappa})\,\mathbf{x}^T\mathbf{s}$. It means that after at most as many steps as in the original method the complementarity gap decreases below ε – when we realize sufficient decrease of the complementarity gap according to the $\mathcal{P}_*(\tilde{\kappa})$-property for each vector during the algorithm – or at an earlier iteration the lack of $\mathcal{P}_*(\tilde{\kappa})$-property is detected. This observation, combined with the complexity theorem of the original algorithm (see Theorem 6.12), proves our statement. ■

At the end of this subsection let us note that the case $\rho = 0$ can be treated analogously.

6.1.3 The predictor-corrector interior point algorithm

In this section we modify the algorithm proposed in [66]. In this predictor-corrector algorithm we take affine and centering steps alternately. The original predictor corrector algorithm by Potra et al. is as follows:

Predictor-corrector algorithm for $\mathcal{P}_*(\kappa)$-matrix LCPs

---

Input:

an accuracy parameter $\varepsilon > 0$;
a proximity parameter $\gamma \in (0, 1)$;
$t = (1 - \gamma)/[(1 + 4\kappa)n + 1]$;
an initial point $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{D}(\gamma)$;

begin
$\mathbf{x} := \mathbf{x}^0$, $\mathbf{s} := \mathbf{s}^0$, $\mu := (\mathbf{x}^0)^T\mathbf{s}^0/n$;
while $\mathbf{x}^T\mathbf{s} \geq \varepsilon$ **do**
Predictor step

calculate the affine Newton direction $(\Delta\mathbf{x}, \Delta\mathbf{s})$ by solving (3.11) with $\mathbf{a} = -\mathbf{xs}$;

$\bar{\theta} = \sup\left\{\hat{\theta} > 0 : \ (\mathbf{x}(\theta), \mathbf{s}(\theta)) \in \mathcal{D}\big((1-\mathrm{t})\gamma\big), \ \forall\, \theta \in [0, \hat{\theta}]\right\}$;

$\bar{\mathbf{x}} = \mathbf{x}(\bar{\theta}), \ \bar{\mathbf{s}} = \mathbf{s}(\bar{\theta}), \ \bar{\mu} = \bar{\mathbf{x}}^T\bar{\mathbf{s}}/n$;

if $\bar{\mu} = 0$ **then** STOP: $(\bar{\mathbf{x}}, \bar{\mathbf{s}})$ is an optimal solution;

if $(\bar{\mathbf{x}}, \bar{\mathbf{s}}) \in \mathcal{D}(\gamma)$ **then**

$\mathbf{x} = \bar{\mathbf{x}}, \ \mathbf{s} = \bar{\mathbf{s}}, \ \mu = \bar{\mu}$ and go to *Predictor step*;

Corrector step

calculate the centering Newton direction $(\Delta\bar{\mathbf{x}}, \Delta\bar{\mathbf{s}})$ by solving (3.11) with $\mathbf{a} = \mu\mathbf{e} - \bar{\mathbf{x}}\bar{\mathbf{s}}$;

$\theta^+ = \operatorname{argmin}\{\bar{\mu}(\theta) : \ (\bar{\mathbf{x}}(\theta), \bar{\mathbf{s}}(\theta)) \in \mathcal{D}(\gamma)\}$;

$\mathbf{x}^+ = \bar{\mathbf{x}} + \theta^+\Delta\bar{\mathbf{x}}, \ \mathbf{s}^+ = \bar{\mathbf{s}} + \theta^+\Delta\bar{\mathbf{s}}, \ \mu^+ = (\mathbf{x}^+)^T\mathbf{s}^+/n$;

$\mathbf{x} = \mathbf{x}^+, \ \mathbf{s} = \mathbf{s}^+, \ \mu = \mu^+$;

end

end.

Potra and Liu [66] determined the maximum feasible predictor step length as the minimum of $n+1$ real numbers. This is a direct deduction from the definition of the new point without any assumptions on the matrix of the problem. The first number is the upper bound which comes from the feasibility of the new point, and the other n bounds warrant the new point to be in the appropriate neighbourhood[1].

Lemma 6.16 (From expressions (3.16), (3.17) in [66]) *Let M be an arbitrary matrix, $(\mathbf{x}, \mathbf{s}) \in \mathcal{D}(\gamma)$, $(\Delta\mathbf{x}, \Delta\mathbf{s})$ be the predictor direction in the predictor-corrector algorithm and let the predictor step length be*

$$\bar{\theta} = \sup\left\{\hat{\theta} > 0 : \ (\mathbf{x}(\theta), \mathbf{s}(\theta)) \in \mathcal{D}\big((1-\mathrm{t})\gamma\big), \ \forall\, \theta \in [0, \hat{\theta}]\right\}.$$

Furthermore, let $\bar{\theta}_0 = \dfrac{2}{1 + \sqrt{1 - 4\mathbf{e}^T\mathbf{r}/n}}$ *and*

$$\bar{\theta}_i = \begin{cases} \infty & \text{if } \ \Delta_i \le 0 \\ 1 & \text{if } \ r_i - (1-t)\gamma\mathbf{e}^T\mathbf{r}/n = 0 \\ \frac{2(v_i^2 - (1-t)\gamma)}{v_i^2 - (1-t)\gamma + \sqrt{\Delta_i}} & \text{if } \ \Delta_i > 0 \text{ and } r_i - (1-t)\gamma\mathbf{e}^T\mathbf{r}/n \neq 0, \end{cases}$$

where

$$\Delta_i = (v_i^2 - (1-t)\gamma)^2 - 4(v_i^2 - (1-t)\gamma)\left(r_i - (1-t)\gamma\mathbf{e}^T\mathbf{r}/n\right), \quad \text{for each } 0 < i \le n.$$

Then we have

$$\bar{\theta} = \min\left\{\bar{\theta}_i : \ 0 \le i \le n\right\}.$$

[1] The vector $\mathbf{r}$ is a notation of scaling, see (3.12).

Since the current point is in the γ neighbourhood of the central path, the following lower bound can be proved on $\bar{\theta}_i$ for all $1 \leq i \leq n$.

Lemma 6.17 (From the proof of Theorem 3.3 in [66]) *Let the assumptions of Lemma 6.16 hold, and $\bar{\theta}_i$, $1 \leq i \leq n$ be as it is given in Lemma 6.16, then*

$$\bar{\theta}_i \geq \frac{2}{1 + \sqrt{1 + (t\gamma)^{-1}(4\|\mathbf{r}\|_\infty + 4\mathbf{e}^T\mathbf{r}/n)}}.$$

By the definition of the proximity measure δ_c and considering the fact that the point given by a predictor step is in the $(1-t)\gamma$ neighbourhood, the following upper bound can be verified on the proximity measure after a predictor step.

Lemma 6.18 (From the proof of Theorem 3.3 in [66]) *Let M be an arbitrary matrix and let the point after the predictor step in the predictor-corrector algorithm satisfy $(\bar{\mathbf{x}}, \bar{\mathbf{s}}) \in \mathcal{D}((1-t)\gamma)$. Then*

$$\delta_c(\bar{\mathbf{x}}\,\bar{\mathbf{s}}, \bar{\mu})^2 \leq \frac{1 - (1-t)\gamma}{(1-t)\gamma}\, n.$$

When the matrix of the LCP is a $\mathcal{P}_*(\kappa)$-matrix with a known κ value, the following feasible predictor and corrector step length can be determined.

Lemma 6.19 (From Theorem 3.3 in [66]) *Let M be a $\mathcal{P}_*(\kappa)$-matrix and $(\mathbf{x}, \mathbf{s}) \in \mathcal{D}(\gamma)$. Then the predictor step length satisfies*

$$\theta_p^*(\kappa) := \frac{2\sqrt{(1-\gamma)\gamma}}{(1+4\kappa)n + 2} \leq \sup\left\{\hat{\theta} > 0 : \ \ (\mathbf{x}(\theta), \mathbf{s}(\theta)) \in \mathcal{D}\big((1-\mathrm{t})\gamma\big),\ \forall\, \theta \in [0, \hat{\theta}]\right\},$$

and the corrector step length

$$\theta_c^*(\kappa) := \frac{2\gamma}{(1+4\kappa)n + 1}$$

determines a point in the $\mathcal{D}(\gamma)$ neighbourhood, i.e., $(\bar{\mathbf{x}}(\theta_c^*(\kappa)), \bar{\mathbf{s}}(\theta_c^*(\kappa))) \in \mathcal{D}(\gamma)$, where $(\bar{\mathbf{x}}, \bar{\mathbf{s}}) = (\mathbf{x}(\theta_p^*(\kappa)), \mathbf{s}(\theta_p^*(\kappa))) \in \mathcal{D}((1-t)\gamma)$.

Using the predictor and corrector step length given in Lemma 6.19, a lower bound can be provided on the decrease of the complementarity gap after each iteration.

Lemma 6.20 (From Theorem 3.3 in [66]) *Let M be an arbitrary matrix, $(\mathbf{x}, \mathbf{s}) \in \mathcal{D}(\gamma)$, $\mu_g = \mathbf{x}^T\mathbf{s}/n$, the definition of parameters $\theta_p^*(\kappa)$ and $\theta_c^*(\kappa)$ be the same as in Lemma 6.19, $\bar{\theta}$ be the predictor and θ^+ be the corrector step length, $(\Delta\mathbf{x}, \Delta\mathbf{s})$ be the predictor and $(\Delta\bar{\mathbf{x}}, \Delta\bar{\mathbf{s}})$ the corrector Newton direction in the predictor-corrector algorithm. If $\bar{\theta} \geq \theta_p^*(\kappa)$ and the step*

length $\theta_c^*(\kappa)$ *determines a point in the* $\mathcal{D}(\gamma)$ *neighbourhood, i.e.,* $(\bar{\mathbf{x}}(\theta_c^*(\kappa)), \bar{\mathbf{s}}(\theta_c^*(\kappa))) \in \mathcal{D}(\gamma)$, *where* $(\bar{\mathbf{x}}, \bar{\mathbf{s}}) = (\mathbf{x}(\bar{\theta}), \mathbf{s}(\bar{\theta}))$, *then*

$$\mu_g^+ \leq \left(1 - \frac{3\sqrt{(1-\gamma)\gamma}}{2((1+4\kappa)n+2)}\right)\mu_g,$$

where $\mu_g^+ = (\mathbf{x}^+)^T\mathbf{s}^+/n = \bar{\mathbf{x}}(\theta^+)^T\bar{\mathbf{s}}(\theta^+)/n$.

The complexity of the predictor-corrector IPM for LCPs with a $\mathcal{P}_*(\kappa)$-matrix is a direct consequence of the previous lemma.

Theorem 6.21 (Corollary 3.4 in [66]) *Let* M *be a* $\mathcal{P}_*(\kappa)$*-matrix and* $(\mathbf{x}^0, \mathbf{s}^0)$ *be a feasible interior point such that* $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{D}(\gamma)$. *Then in at most*

$$\mathcal{O}\left((1+\kappa)n \log \frac{(\mathbf{x}^0)^T\mathbf{s}^0}{\varepsilon}\right)$$

steps the predictor-corrector algorithm produces a point $(\hat{\mathbf{x}}, \hat{\mathbf{s}})$ *such that* $(\hat{\mathbf{x}}, \hat{\mathbf{s}}) \in \mathcal{D}(\gamma)$ *and* $\hat{\mathbf{x}}^T\hat{\mathbf{s}} \leq \varepsilon$.

In a predictor step $\theta_p^*(\kappa)$ (see the definition in Lemma 6.19) is a theoretical feasible step length if the matrix M is $\mathcal{P}_*(\kappa)$. Therefore, if the maximal feasible step length is smaller than $\theta_p^*(\kappa)$, then the matrix is not $\mathcal{P}_*(\kappa)$ with the actual value of κ, so κ should be increased. In a corrector step we return to the smaller $\mathcal{D}(\gamma)$ neighbourhood with the step size $\theta_c^*(\kappa)$ (see the definition in Lemma 6.19) if the matrix is $\mathcal{P}_*(\kappa)$. Accordingly, if the new point with step length $\theta_c^*(\kappa)$ is not in $\mathcal{D}(\gamma)$, then the matrix M is not $\mathcal{P}_*(\kappa)$ with the actual value of κ, so κ should be updated. Similarly to the previous two algorithms, if in a predictor or corrector step the new value of κ is not defined by (6.1), then the matrix is not $\mathcal{P}_*$ and the current Newton direction is a certificate for it. Furthermore, if the new value of κ is larger than $\tilde{\kappa}$, then the matrix is not $\mathcal{P}_*(\tilde{\kappa})$ and the Newton direction is a certificate for it. Of course, in this algorithm we also check the existence and uniqueness of the Newton direction in each iteration. The modified algorithm is as follows:

Modified predictor-corrector algorithm

Input:

an upper bound $\tilde{\kappa} > 0$ on the value of κ;
an accuracy parameter $\varepsilon > 0$;
a proximity parameter $\gamma \in (0, 1)$;
an initial point $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{D}(\gamma)$;

begin

$\mathbf{x} := \mathbf{x}^0$, $\mathbf{s} := \mathbf{s}^0$, $\mu := (\mathbf{x}^0)^T\mathbf{s}^0/n$, $\kappa := 0$;

while $\mathbf{x}^T\mathbf{s} \geq \varepsilon$ **do**

Predictor step

$t = \frac{1-\gamma}{(1+4\kappa)n+1}$;

calculate the affine Newton direction $(\Delta\mathbf{x}, \Delta\mathbf{s})$ by solving (3.11) with $\mathbf{a} = -\mathbf{xs}$;

if (the Newton direction does not exists, or it is not unique) **then**

return the matrix is not $\mathcal{P}_0$; *% see Corollary 3.10*

$\bar{\theta} = \sup\left\{\hat{\theta} > 0 : \ \ (\mathbf{x}(\theta), \mathbf{s}(\theta)) \in \mathcal{D}\big((1-\mathrm{t})\gamma\big), \ \forall\, \theta \in [0, \hat{\theta}]\right\}$;

if $\big(\bar{\theta} < \theta_p^*(\kappa)\big)$ **then**

calculate $\kappa(\Delta\mathbf{x})$; *% see (6.1)*

if ($\kappa(\Delta\mathbf{x})$ is not defined) **then**

return the matrix is not $\mathcal{P}_*$; *% see Lemma 6.2*

if $(\kappa(\Delta\mathbf{x}) > \tilde{\kappa})$ **then**

return the matrix is not $\mathcal{P}_*(\tilde{\kappa})$; *% see Lemma 6.1*

$\kappa = \kappa(\Delta\mathbf{x})$;

update $\theta_p^*(\kappa)$ and $\theta_c^*(\kappa)$;

$\bar{\mathbf{x}} = \mathbf{x}(\bar{\theta})$, $\bar{\mathbf{s}} = \mathbf{s}(\bar{\theta})$, $\bar{\mu} = \bar{\mathbf{x}}^T\bar{\mathbf{s}}/n$;

if $\bar{\mu} = 0$ **then** STOP: $(\bar{\mathbf{x}}, \bar{\mathbf{s}})$ is an optimal solution;

if $(\bar{\mathbf{x}}, \bar{\mathbf{s}}) \in \mathcal{D}(\gamma)$ **then**

$\mathbf{x} = \bar{\mathbf{x}}$, $\mathbf{s} = \bar{\mathbf{s}}$, $\mu = \bar{\mu}$ and go to *Predictor step*;

Corrector step

calculate the centering Newton direction $(\Delta\bar{\mathbf{x}}, \Delta\bar{\mathbf{s}})$ by solving (3.11) with $\mathbf{a} = \mu\mathbf{e} - \bar{\mathbf{x}}\bar{\mathbf{s}}$;

if (the Newton direction does not exists, or it is not unique) **then**

return the matrix is not $\mathcal{P}_0$; *% see Corollary 3.10*

if $\big((\bar{\mathbf{x}}(\theta_c^*(\kappa)), \bar{\mathbf{x}}(\theta_c^*(\kappa))) \notin \mathcal{D}(\gamma)\big)$

calculate $\kappa(\Delta\bar{\mathbf{x}})$; *% see (6.1)*

if ($\kappa(\Delta\bar{\mathbf{x}})$ is not defined) **then**

return the matrix is not $\mathcal{P}_*$; *% see Lemma 6.2*

if $(\kappa(\Delta\mathbf{x}) > \tilde{\kappa})$ **then**

return the matrix is not $\mathcal{P}_*(\tilde{\kappa})$; *% see Lemma 6.1*

$\kappa = \kappa(\Delta\bar{\mathbf{x}})$;

update $\theta_p^*(\kappa)$ and $\theta_c^*(\kappa)$;

$\theta^+ = \operatorname{argmin}\left\{\bar{\mu}(\theta) : \ \ (\bar{\mathbf{x}}(\theta), \bar{\mathbf{s}}(\theta)) \in \mathcal{D}(\gamma)\right\}$;

$\mathbf{x}^+ = \bar{\mathbf{x}} + \theta^+\Delta\bar{\mathbf{x}}$, $\mathbf{s}^+ = \bar{\mathbf{s}} + \theta^+\Delta\bar{\mathbf{s}}$, $\mu^+ = (\mathbf{x}^+)^T\mathbf{s}^+/n$;

$\mathbf{x} = \mathbf{x}^+$, $\mathbf{s} = \mathbf{s}^+$, $\mu = \mu^+$;

end
end.

The following lemmas show that if $\theta_p^*(\kappa)$ or $\theta_c^*(\kappa)$ is not a feasible step length (for the definitions see Theorem 6.19), then the matrix is not a $\mathcal{P}_*(\kappa)$-matrix.

Lemma 6.22 *If there exists an index i $(0 \leq i \leq n)$ such that*

$$\bar{\theta}_i \leq \theta_p^*(\kappa),$$

then the matrix M is not a $\mathcal{P}_*(\kappa)$-matrix and the affine Newton direction is a certificate for this.

Proof. For any $\kappa \geq 0$ and $n \geq 1$

$$\theta_p^*(\kappa) < \frac{2}{1+\sqrt{1+4\kappa}},$$

therefore if $\bar{\theta}_0 \leq \theta_p^*(\kappa)$, then by the definition of $\bar{\theta}_0$ one has

$$\frac{2}{1+\sqrt{1-4\mathbf{e}^T\mathbf{r}/n}} = \bar{\theta}_0 < \frac{2}{1+\sqrt{1+4\kappa}},$$

implying $\mathbf{e}^T\mathbf{r}/n < -\kappa$, thus $\sum_{i\in I}\Delta x_i \Delta s_i < -\kappa n\mu = -\kappa\,\mathbf{x}^T\mathbf{s}$. Therefore, by Lemma 6.3 the matrix M is not a $\mathcal{P}_*(\kappa)$-matrix and the affine Newton direction $\Delta\mathbf{x}$ is a certificate for this.

If $\bar{\theta}_i \leq \theta_p^*(\kappa)$, where $0 < i \leq n$, then let us consider the following inequality, which was proved by Potra and Liu in [66] on p.158:

$$\sqrt{(1-\gamma)\gamma} + \sqrt{((1+4\kappa)n+1)^2 + \gamma(1-\gamma)} < (1+4\kappa)n + 2. \tag{6.14}$$

Using Lemma 6.17, Lemma 3.13 and the definition of t, one has

$$\begin{aligned}
\frac{2\sqrt{(1-\gamma)\gamma}}{(1+4\kappa)n+2} &= \theta_p^*(\kappa) \geq \bar{\theta}_i \geq \frac{2}{1+\sqrt{1+(t\gamma)^{-1}(4\|\mathbf{r}\|_\infty + 4\mathbf{e}^T\mathbf{r}/n)}} \\
&\geq \frac{2}{1+\sqrt{1+(t\gamma)^{-1}(4\|\mathbf{r}\|_\infty + 1)}} \\
&= \frac{2\sqrt{(1-\gamma)\gamma}}{\sqrt{(1-\gamma)\gamma} + \sqrt{((1+4\kappa)n+1)(4\|\mathbf{r}\|_\infty+1)+\gamma(1-\gamma)}}.
\end{aligned} \tag{6.15}$$

From inequality (6.15) and (6.14) we get

$$4\|\mathbf{r}\|_\infty + 1 > (1+4\kappa)n + 1. \tag{6.16}$$

Since $(\Delta\mathbf{x}, \Delta\mathbf{s})$ is a solution of the system (3.11) with $\mathbf{a} = -\mathbf{xs}$, and using inequality (6.16) with $\mu\, n = \mathbf{x}^T\mathbf{s}$, one has

$$\|\Delta\mathbf{x}\Delta\mathbf{s}\|_\infty > \frac{(1+4\kappa)}{4}\,\mathbf{x}^T\mathbf{s} = \frac{1+4\kappa}{4}\left\|\frac{\mathbf{a}}{\sqrt{\mathbf{xs}}}\right\|^2,$$

so by the second statement of Lemma 6.3 one has $M \notin \mathcal{P}_*(\kappa)$, and $\Delta\mathbf{x}$ is a certificate for this. ■

Now let us analyze the corrector step.

Lemma 6.23 *If $\theta_c^*(\kappa)$ is such a corrector step length that $(\bar{\mathbf{x}}(\theta_c^*(\kappa)), \bar{\mathbf{s}}(\theta_c^*(\kappa))) \notin \mathcal{D}(\gamma)$, then the matrix M is not a $\mathcal{P}_*(\kappa)$-matrix and the corrector Newton direction is a certificate for this.*

Proof. Note that

$$\bar{\mathbf{x}}(\theta)\bar{\mathbf{s}}(\theta) = (1-\theta)\bar{\mathbf{x}}\bar{\mathbf{s}} + \theta\bar{\mu}\,\mathbf{e} + \theta^2\Delta\bar{\mathbf{x}}\Delta\bar{\mathbf{s}}$$

and

$$\bar{\mu}(\theta) = \bar{\mu} + \theta^2\frac{\Delta\bar{\mathbf{x}}^T\Delta\bar{\mathbf{s}}}{n}.$$

From Lemma 3.12 and Lemma 6.18 we get

$$\begin{aligned}\Delta\bar{\mathbf{x}}^T\Delta\bar{\mathbf{s}} &\leq \sum_{\mathcal{I}_+}\Delta\bar{x}_i\Delta\bar{s}_i \leq \frac{1}{4}\left\|\frac{\bar{\mu}\mathbf{e} - \bar{\mathbf{x}}\bar{\mathbf{s}}}{\sqrt{\bar{\mathbf{x}}\bar{\mathbf{s}}}}\right\|^2 = \frac{1}{4}\,\bar{\mu}\left\|\sqrt{\frac{\bar{\mathbf{x}}\bar{\mathbf{s}}}{\bar{\mu}}} - \sqrt{\frac{\bar{\mu}}{\bar{\mathbf{x}}\bar{\mathbf{s}}}}\right\|^2 \\ &\leq \frac{1}{4}\,\bar{\mu}\,\frac{1-(1-t)\gamma}{(1-t)\gamma}\,n,\end{aligned}$$

therefore

$$\bar{\mu}(\theta) \leq \left(1 + \frac{1-(1-t)\gamma}{4(1-t)\gamma}\,\theta^2\right)\bar{\mu}. \tag{6.17}$$

Since $\theta_c^*(\kappa)$ is an infeasible step length, there exists an index i such that

$$\bar{x}(\theta_c^*(\kappa))_i\;\bar{s}(\theta_c^*(\kappa))_i < \gamma\bar{\mu}(\theta_c^*(\kappa)), \quad \text{namely}$$

$$(1-\theta_c^*(\kappa))\bar{x}_i\bar{s}_i + \theta_c^*(\kappa)\bar{\mu}\; + (\theta_c^*(\kappa))^2\Delta\bar{x}_i\Delta\bar{s}_i < \gamma\bar{\mu}(\theta_c^*(\kappa)).$$

The predictor point $(\bar{\mathbf{x}}, \bar{\mathbf{s}}) \in \mathcal{D}((1-t)\gamma)$, so $\bar{x}_i\bar{s}_i \geq (1-t)\gamma\bar{\mu}$. Furthermore, by inequality (6.17) one has

$$(1-\theta_c^*(\kappa))(1-t)\gamma\bar{\mu} + \theta_c^*(\kappa)\bar{\mu} + (\theta_c^*(\kappa))^2\Delta\bar{x}_i\Delta\bar{s}_i < \gamma\left(1 + \frac{1-(1-t)\gamma}{4(1-t)\gamma}\,(\theta_c^*(\kappa))^2\right)\bar{\mu},$$

which implies

$$(\theta_c^*(\kappa))^2\frac{\Delta\bar{x}_i\Delta\bar{s}_i}{\bar{\mu}} < t\gamma - \theta_c^*(\kappa)(1-(1-t)\gamma) + \frac{1-(1-t)\gamma}{4(1-t)}(\theta_c^*(\kappa))^2. \tag{6.18}$$

One can check the following equality by substituting the values of t and $\theta_c^*(\kappa)$

$$0 \le \frac{(1-\gamma)\gamma^2}{((1+4\kappa)n+1)^2} = -t\gamma + \theta_c^*(\kappa)(1-(1-t)\gamma) - \frac{1-(1-t)\gamma}{4(1-t)\gamma}\big[(1+4\kappa)n+\gamma\big](\theta_c^*(\kappa))^2.$$

Therefore,

$$-\frac{1-(1-t)\gamma}{4(1-t)\gamma}\,(1+4\kappa)\,n\,(\theta_c^*(\kappa))^2 \ge t\gamma - \theta_c^*(\kappa)(1-(1-t)\gamma) + \frac{1-(1-t)\gamma}{4(1-t)}\,(\theta_c^*(\kappa))^2.$$

Combining this with inequality (6.18) and then considering Lemma 6.18, we get

$$\Delta\bar{x}_i\Delta\bar{s}_i < -\frac{1-(1-t)\gamma}{4(1-t)\gamma}\,(1+4\kappa)\,n\bar{\mu} \le -\frac{(1+4\kappa)\bar{\mu}}{4}\left\|\sqrt{\frac{\bar{\mathbf{x}}\bar{\mathbf{s}}}{\bar{\mu}}} - \sqrt{\frac{\bar{\mu}}{\bar{\mathbf{x}}\bar{\mathbf{s}}}}\right\|^2. \tag{6.19}$$

Since $(\Delta\bar{\mathbf{x}}, \Delta\bar{\mathbf{s}})$ is a solution of the system (3.11) with $\mathbf{a} = \bar{\mu}\mathbf{e} - \bar{\mathbf{x}}\bar{\mathbf{s}}$, using inequality (6.19), one get

$$\|\Delta\bar{\mathbf{x}}\Delta\bar{\mathbf{s}}\|_\infty > \frac{(1+4\kappa)\bar{\mu}}{4}\left\|\sqrt{\frac{\bar{\mathbf{x}}\bar{\mathbf{s}}}{\bar{\mu}}} - \sqrt{\frac{\bar{\mu}}{\bar{\mathbf{x}}\bar{\mathbf{s}}}}\right\|^2 = \frac{1+4\kappa}{4}\left\|\frac{\mathbf{a}}{\sqrt{\bar{\mathbf{x}}\bar{\mathbf{s}}}}\right\|^2.$$

Thus, by the second statement of Lemma 6.3, the matrix M is not a $\mathcal{P}_*(\kappa)$-matrix and the corrector Newton direction $\Delta\bar{\mathbf{x}}$ is a certificate for this. ■

The following lemma proves that the predictor-corrector algorithm is well defined.

Lemma 6.24 *At each iteration when the value of κ is updated, the new value of $\theta_p^*(\kappa)$ satisfies inequality $\bar{\theta} \ge \theta_p^*(\kappa)$ and the new point $(\bar{\mathbf{x}}(\theta_c^*(\kappa)), \bar{\mathbf{s}}(\theta_c^*(\kappa)))$ determined by the new value of the corrector step size $\theta_c^*(\kappa)$ is in the $\mathcal{D}(\gamma)$ neighbourhood.*

Proof. In the proof of Lemma 6.19 we use the $\mathcal{P}_*(\kappa)$ property only for the vector $\Delta\mathbf{x}$ or $\Delta\bar{\mathbf{x}}$. When parameter κ is updated we choose the new value in such a way that the inequality in the definition of $\mathcal{P}_*(\kappa)$-matrices (2.1) holds for the vectors $\Delta\mathbf{x}$ and $\Delta\bar{\mathbf{x}}$. Therefore, the new value of $\theta_p^*(\kappa)$ satisfies the inequality $\bar{\theta} \ge \theta_p^*(\kappa)$ and the new value of $\theta_c^*(\kappa)$ determines a point in the $\mathcal{D}(\gamma)$ neighbourhood. ■

Now we are ready to state the complexity result for the modified predictor-corrector algorithm for general LCPs in case an initial interior point is available.

Theorem 6.25 *Let $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{F}^+$ such that $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{D}(\gamma)$. Then after at most*

$$\mathcal{O}\left((1+\hat{\kappa})n\log\frac{(\mathbf{x}^0)^T\mathbf{s}^0}{\varepsilon}\right)$$

steps, where $\hat{\kappa} \leq \tilde{\kappa}$ is the largest value of parameter κ throughout the algorithm, the predictor-corrector algorithm generates a point $(\hat{\mathbf{x}}, \hat{\mathbf{s}})$, such that $\hat{\mathbf{x}}^T\hat{\mathbf{s}} \leq \varepsilon$ and $(\hat{\mathbf{x}}, \hat{\mathbf{s}}) \in \mathcal{D}(\gamma)$ or provides a certificate that the matrix is not $\mathcal{P}_*(\tilde{\kappa})$.

Proof. We follow the proof of the previous complexity theorem (see Theorem 6.15). If we take a predictor and a corrector step, then by Theorem 6.20 and Lemma 6.24 the decrease of the complementarity gap is at least

$$\frac{3\sqrt{(1-\gamma)\gamma}}{2((1+4\kappa)n+2)}\frac{\mathbf{x}^T\mathbf{s}}{n}.$$

This expression is a decreasing function of κ, so at each iteration when we make a predictor and a corrector step, the complementarity gap decreases at least by

$$\frac{3\sqrt{(1-\gamma)\gamma}}{2((1+4\tilde{\kappa})n+2)}\frac{\mathbf{x}^T\mathbf{s}}{n}.$$

We take at most as many iterations as in the original predictor-corrector IPM with a $\mathcal{P}_*(\hat{\kappa})$-matrix. Thus, referring to the complexity theorem of the original algorithm (see Theorem 6.21) we have proved the theorem. ■

6.1.4 An EP theorem for LCPs based on interior point algorithms

We know that if we assume $\mathcal{F}^+ \neq \emptyset$ and the matrix of the LCP is sufficient, then the LCP has a solution by Corollary 3.4. According to this result, by making use of the complexity theorem of the previous sections (Theorem 6.9, 6.15, 6.25) and the rounding procedure of [36], we can now present the following EP type theorem. We assume that the data are rational (since problems are solved with computers, this is reasonable), ensuring polynomial encoding size of certificates and polynomial complexity of the algorithm.

Theorem 6.26 *Let an arbitrary matrix $M \in \mathbb{Q}^{n\times n}$, a vector $\mathbf{q} \in \mathbb{Q}^n$, a point $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{F}^+$ with $\delta_c(\mathbf{x}^0\mathbf{s}^0, \mu^0) \leq \tau$ and $\varepsilon > 0$ be given. Then one can verify in polynomial time that at least one of the following statements holds*

(1) the LCP has an ε-optimal solution[2] $(\mathbf{x}, \mathbf{s})$ whose encoding size is polynomially bounded.

(2) the matrix M is not in the class of $\mathcal{P}_*(\tilde{\kappa})$ and there is a certificate whose encoding size is polynomially bounded.

[2]Here we preserve the designation of linear programming theory, namely, a point $(\mathbf{x}, \mathbf{s})$ is called ε-optimal solution of the LCP, if it is a feasible solution and $\mathbf{x}^T\mathbf{s} \leq \varepsilon$.

6.2 Solving general LCPs without having an initial interior point

When for an LCP no initial interior points are known, then we have two possibilities: (i) we apply an infeasible interior point algorithm, or (ii) we use an embedding technique (for example see Section 3.7). Let us consider here the second one.

We dealt with the dual of the LCP in Chapter 4. Based on Lemma 4.3, let us approach the problem from the dual side. First, we try to solve the feasibility problem of $DLCP$. It is a linear optimization problem, therefore we can solve it in polynomial time. We have the following cases:

(a) $\mathcal{F}_D \neq \emptyset$ and $\mathbf{uz} = \mathbf{0}$ holds for the computed $(\mathbf{u}, \mathbf{z}) \in \mathcal{F}_D$: then we solved the $DLCP$.

(b) $\mathcal{F}_D \neq \emptyset$ and for the computed $(\mathbf{u}, \mathbf{z}) \in \mathcal{F}_D$: $\mathbf{uz} \neq \mathbf{0}$ holds, then by Lemma 4.3 we know that M is not a row sufficient matrix, therefore it is not a sufficient matrix either and vector $\mathbf{z}$ is a certificate for this.

(c) $\mathcal{F}_D = \emptyset$, then the $DLCP$ has no solution.

In cases (a) and (b) we have solved the LCP in the sense of Theorem 4.1. In case (c) we try to solve the embedded problem LCP' with using one of the above presented modified IPMs. The modified algorithm either shows that matrix M' – and thus by Lemma 3.15 matrix M as well – is not in the class of $\mathcal{P}_*(\tilde{\kappa})$ or solves the problem (LCP'). In the latter case we have two subcases:

(i) $\tilde{\mathbf{x}} = \mathbf{0}$, then by Lemma 3.15 the LCP has a solution.

(ii) $\tilde{\mathbf{x}} \neq \mathbf{0}$.

When $\tilde{\mathbf{x}} \neq \mathbf{0}$ and $\mathcal{F}_D = \emptyset$, if matrix M is sufficient, then it is also column sufficient, so the LCP has no solution by Lemma 3.15. But this contradicts the Fukuda-Terlaky LCP duality theorem [16, 25, 26], therefore in this case matrix M can not be sufficient and the vector $\tilde{\mathbf{x}}$ is an indirect certificate for this.

We can state our main result combining the dual side approach with the complexity result Theorem 6.26 (an interior point of the problem (LCP') is known by construction).

Theorem 6.27 *Let an arbitrary matrix $M \in \mathbb{Q}^{n\times n}$, a vector $\mathbf{q} \in \mathbb{Q}^n$, and $\varepsilon > 0$ be given. Then one can verify in polynomial time that at least one of the following statements holds*

(1) the LCP problem (1.1) has an ε-optimal solution $(\mathbf{x}, \mathbf{s})$ whose encoding size is polynomially bounded.

(2) the DLCP problem (4.1) has a feasible complementary solution $(\mathbf{u}, \mathbf{z})$ *whose encoding size is polynomially bounded.*

(3) matrix M *is not in the class* $\mathcal{P}_*(\tilde{\kappa})$.

Theorem 6.27 is a generalization of Theorem 6.26. Since the interior point assumption is eliminated, it can occur that the LCP has no solution while matrix M is sufficient. This is the second statement of Theorem 6.27. On the other hand, as we see in case (ii) in the dual side approach, when the matrix is not sufficient, we only have an indirect certificate $\tilde{\mathbf{x}}$. This is the reason why in the last case of Theorem 6.27 we can not ensure an explicit certificate. Therefore, Theorem 6.27 is stronger than Theorem 6.26, because the interior point assumption is eliminated, however only an indirect certificate is provided in the last case.

It is interesting to note that Theorem 6.27 and Theorem 4.1 (a result of Fukuda et al. [25]) are different in two aspects: first, our statement (3) is weaker in some cases than theirs (there is no direct certificate in one case), but on the other hand, our constructive proof is based on polynomial time algorithms and a polynomial size certificate is provided in all the other cases in polynomial time.

6.3 Computational results

We would like to demonstrate the efficiency of the modified interior point algorithms. Unfortunately, there is not such a set of test problems for LCP algorithms as in the case of linear programming problems (like NETLIB). Ye and his colleagues [88] published computational results of their algorithms for a special class of LCPs, which comes from an economical problem, the Leontief economy equilibrium problem, as we introduced in Section 1.2.5. In this section we summarize our computational experiences of two modified IPMs (the long-step path-following and the predictor-corrector algorithm) and compare our computational results with Ye's results.

Ye et al. dealt with the following LCP:

$$A^T\mathbf{u} + \mathbf{v} = \mathbf{e}, \quad \mathbf{uv} = \mathbf{0}, \quad \mathbf{u}, \mathbf{v} \geq \mathbf{0}, \tag{6.20}$$

where A is a nonnegative matrix, the Leontief coefficient matrix. It is easy to see, that $\mathbf{u} = \mathbf{0}$, $\mathbf{v} = \mathbf{e}$ is a trivial solution of the problem. However, this solution has no meaning in economy, because $\mathbf{u}$ represents the utility of traders. Therefore, the following system has to be solved:

$$A^T\mathbf{u} + \mathbf{v} = \mathbf{e}, \quad \mathbf{uv} = \mathbf{0}, \quad \mathbf{u}, \mathbf{v} \geq \mathbf{0}, \quad \mathbf{u} \neq \mathbf{0}, \tag{6.21}$$

which is equivalent with the Leontief economy equilibrium problem if A is a positive matrix (see Theorem 1.1).

Ye et al. proved that the solvability of this problem is equivalent with the co-positive plus property of the matrix $-A$, if the matrix A is symmetric.

Theorem 6.28 ([88]) *Let A be a real symmetric matrix. Then, it is $\mathbb{NP}$-complete to decide whether or not the problem (6.21) has a solution.*

On the other hand, if the matrix A is nonnegative, namely all elements are nonnegative, then the problem has a solution.

Ye and his colleagues constructed two algorithms to solve the problem (6.21) with a symmetric nonnegative matrix. HOMOTOPY is a homotopy based interior point path-following algorithm [17], while QP is a quadratic programming based potentional reduction algorithm, which is a fully polynomial approximation schema (FPTAS) [88].

Ye et al. proved that the problem (6.21) can be reformulated as an LCP in the following way:

$$-M\mathbf{x}+\mathbf{s}=\mathbf{q}, \quad \mathbf{xs}=\mathbf{0}, \quad \mathbf{x},\mathbf{s}\geq\mathbf{0}, \tag{6.22}$$

where

$$M=\begin{pmatrix} A^T & -\mathbf{e} \\ -\mathbf{e}^T & 0 \end{pmatrix} \qquad \text{and} \qquad \mathbf{q}=\begin{pmatrix} \mathbf{0} \\ -1 \end{pmatrix}.$$

They randomly generated sparse symmetric matrices, whose elements are uniform in $[0,1]$. They also tried to solve the test problems in form (6.20) with solvers PATH and MILES, which use a Lemke type algorithm. In their experiments, these solvers always returned the trivial solution. Furthermore, MILES could solve none of their test problems in form (6.22), and PATH could only solve a few problems.

We could not acquire this test set of matrices, therefore we also generated such matrices; for each size (10, 20, 40, 60, 80, 100, 200) we created 10 matrices. For each matrix we solve the problem from 1000 randomly generated initial points. The special structure of the problem enables us to neglect the embedding and directly construct an initial point. Since the matrix A is nonnegative, if the first n element of the vector $\mathbf{x}$ are arbitrary positive numbers and the last coordinate of the vector $\mathbf{x}$ is large enough, then the corresponding vector $\mathbf{s}=\mathbf{q}+M\mathbf{x}$ will also be positive.

The two modified IPMs are coded in MATLAB and were run on a desktop computer (1.54 GHz CPU, Windows XP and 1 GB RAM). We set the following parameter values:

- For both IPMs we choose $\tilde{\kappa}=100$, so we either solve the problem, or give a certificate, that M is not a $\mathcal{P}_*(\tilde{\kappa})$-matrix.

- In the modified predictor-corrector IPM the proximity parameter γ is chosen such that the initial point $(\mathbf{x}^0, \mathbf{s}^0)$ will be in the neighbourhood $\mathcal{D}(\gamma)$, i.e., $\gamma = n/2 \min\left(x_i^0 s_i^0\right) / \left(\mathbf{x}^0\right)^T \mathbf{s}$.
- We use a larger centrality parameter in the modified long-step path-following IPM than the theoretical value, because the computational experience shows a better efficiency for a larger value (the same as for linear programming problems), so we choose $\tau = 500$.
- The barrier update parameter in the long-step path-following IPM: $\gamma = 0.5$.
- The step length in the long-step path-following IPM is determined by line search (at most 20 iterations).
- The accuracy parameter ε is 10^{-8} in both modified IPMs.
- We allow maximum 1000 iterations per run, because as experience shows, when a solution is found, the number of iterations is less than 1000 (see the tables below).

The following tables show the efficiency of the algorithms. The first table contains the average computing time of runs in seconds (mean T), the average number of iterations (mean it), the maximum time of runs (max T), the maximum number of iterations (max it) for each dimension for the modified long-step path-following IPM (LS). Here we consider all runs, so each element of the table is computed from 1000 runs for 10 matrices, i.e., from 10000 value.

n	mean T	mean it	max T	max it
10	0.0143	36.93	0.343	48
20	0.0231	35.87	0.765	49
40	0.0542	32.16	7.000	50
60	0.1119	28.33	8.360	50
80	0.2248	27.59	19.922	51
100	0.4093	24.75	23.313	51
200	3.4648	22.49	470.969	52

Table 6.1: Results of the LS considering all runs.

Table 6.2 contains the previous values, but only for the successful runs, namely when the algorithm returns a solution. Furthermore, "# sol" means how many times we have found a solution from the 1000 runs, and "# diff solution" is the number of the different solutions we found. Both of them are an average value for the appropriate ten matrices for

n	#sol	#diff sol	s.mean T	s.mean it	s.max T	s.max it	mean sup
10	856.0	11.6	0.0139	39.58	0.328	48	3.22
20	728.5	79.8	0.0208	41.36	0.313	49	5.20
40	493.3	228.9	0.0450	42.94	2.094	50	7.81
60	312.7	241.6	0.0860	43.57	3.265	50	10.02
80	261.9	217.3	0.1651	44.33	5.766	51	12.18
100	152.9	148.4	0.3036	44.63	19.687	51	14.15
200	19.4	19.4	3.1757	45.76	34.875	52	26.07

Table 6.2: Results of the LS for runs which return a solution.

each dimension. The last column of the table (mean sup) contains the average support size of the found solutions.

Similarly, the computational efficiency of the modified predictor-corrector IPM (PC) is presented by the following two tables. (Here an iteration consists of a predictor and a corrector step.)

n	mean T	mean it	max T	max it
10	0.0112	14.14	0.641	913
20	0.0243	17.67	1.078	803
40	0.0711	23.01	2.203	688
60	0.1715	27.24	6.297	1000
80	0.4224	32.55	13.453	1000
100	0.7378	34.75	21.859	1000
200	8.1261	54.62	155.343	1000

Table 6.3: Results of the PC considering all runs.

When we compare the above tables, we can see that the LS algorithm is faster for problems with size more than 20, if we account all runs. There is an interesting phenomena: the average iteration number decreases for the LS algorithm when the dimension increases. The reason of this is maybe that for a larger problem we may find with larger probability a bad point, namely a point where the LS algorithm stops with a certificate, that the matrix is not $\mathcal{P}_*(\tilde{\kappa})$. For larger sizes, the PC algorithm always has some bad cases, when the iteration number reaches the threshold, but it happens only for a few instances, as the average iteration number is much smaller.

n	#sol	#diff sol	s.mean T	s.mean it	s.max T	s.max it	mean sup
10	839.2	10.6	0.0117	14.89	0.188	240	3.09
20	707.4	67.3	0.0274	20.13	0.485	356	5.03
40	509.1	197.1	0.0916	29.83	2.203	688	7.21
60	335.6	241.3	0.2481	39.68	3.157	479	9.09
80	282.1	225.3	0.6225	48.38	5.344	436	10.79
100	166.5	159.5	1.2625	59.98	14.844	621	12.26
200	28.6	28.6	16.2489	110.22	51.282	346	22.49

Table 6.4: Results of the PC for runs which return a solution.

When the dimension is over 40, the PC algorithm returns a solution in a bit more cases than the LS, and the ratio of number of different solutions is almost the same for both algorithms. On average, the LS algorithm is faster if a solution is found. The average necessary iteration number to find a solution increases slightly for the LS algorithm, while for the PC algorithm it increases rapidly. The average support size of the found solutions is almost the same for the two algorithms.

The computational results show that one run of the modified interior point methods is faster for each dimension than of HOMOTOPY, or QP. But this comparison is not absolutely correct, because the modified IPMs do not give a solution in all runs. Therefore, a more sophisticated comparison is when we evaluate the average time to determine a solution of the problem. Accordingly, we divide time sum for all runs with the number of "successful" runs, namely when a solution is given. In this case we can see that the modified algorithms are better from the point of efficiency if the dimension is less than 80 (see Table 6.5)[3].

To summarize our computational experience, the modified IPMs find a solution less often as the problem size increases. They are suitable for solving small scale problems (dimension is less than 100). They can be used to solve larger problems, but the computational time is worse by orders of magnitude than for QP or HOMOTOPY. The most important advantage of the modified IPMs is that they can determine different solutions – because of randomly generated initial points – in contrast to HOMOTOPY and QP. Furthermore, we do not use the special structure of the problem, only at generating initial points, but it can be substituted by an embedding.

Finally, let us note, that the generated matrices are not sufficient matrices, moreover not even $\mathcal{P}_0$-matrices, so the solution sets of the problems are not convex and not even connected.

[3]Ye and his colleagues coded their algorithms (HOMOTOPY and QP) also in MATLAB, but run their solvers on a desktop computer with a 2.8 GHz CPU.

n	PC	LS	QP	HM	PATH
10	0.01	0.02			
20	0.03	0.03	0.1	0.2	0.1004
40	0.14	0.11	0.1	0.4	0.3406
60	0.51	0.36	0.1	0.8	fail
80	1.50	0.86	0.2	1.4	
100	4.43	2.68	0.3	2.2	
200	284.13	178.60	1.2	14	

Table 6.5: The average computational time for all algorithms.

Furthermore, it is an interesting observation, that the value of κ is almost always 0 through the runs when the algorithm returns a solution.

In Chapter 7 we discuss some ideas how the modified algorithms could be improved.

Chapter 7

Open problems

Here we collect some open problems related to the topic of the thesis.

- The set $\mathcal{P}_0 \setminus \mathcal{P}$ can be considered as the boundary of the set $\mathcal{P}_0$ (see the seventh statement of Lemma 2.5). There are sufficient and not sufficient matrices as well. Furthermore, the boundary of the $\mathcal{P}_*$ matrix class is a subset of $\mathcal{P}_0 \setminus \mathcal{P}$. Is there a good characterization of sufficient matrices which are not $\mathcal{P}$-matrices, namely which are on the boundary of $\mathcal{P}_*$?
  Kojima et al. gave an upper bound on the handicap of a matrix in $\mathcal{P}$ (see Theorem 2.10). Is there a reasonable upper bound on the handicap of a sufficient but not a $\mathcal{P}$ matrix?
- Conjecture: $\hat{\kappa}(M)$ is a continuous function of the entries of M for sufficient matrices.
- Conjecture: The determination of the handicap of a sufficient matrix is $\mathbb{NP}$-hard problem.

- We have seen that if the matrix M is sufficient and the interior point assumption holds, then the complementarity level sets of the LCP are nonempty and compact (Proposition 3.3). Furthermore, the existence and uniqueness of the central path is a corollary of the compactness and nonemptyness of the complementarity level sets. Is the reverse is true? Namely, if the central path exists and it is unique, then should the complementarity level sets be nonempty and compact?
- Give an EP type theorem for the existence and uniqueness of the central path, namely for Theorem 3.6.
- Give a classification of LCPs (with non sufficient matrices) according to the number of central paths.

- Construct an LCP, which has no central path, if it is possible. (At the end of this chapter we present a problem, which has a solution without a central path, namely, there is no central path which tends to this solution.)

- We proved that the DLCP can be solved in polynomial time if the matrix M is row sufficient (Corollary 4.4). Contrarily, for the primal problem, so for the LCP there is no similar result. Although, when the matrix M is sufficient, there are polynomial time algorithms, but their complexity depends on the handicap of the matrix M, which can be arbitrary large for sufficient matrices, and it can be infinite for a row sufficient matrix. Is there a matrix class – wider than PSD – for which the LCP can be solved in polynomial time with interior point methods?

- The Rounding procedure – from an ε-optimal solution determines an exact solution of the LCP – is proved only for the sufficient matrix class. Generalize it for LCPs with arbitrary matrices, and give an EP type theorem. Based on such a result, the "ε-optimal solution" could be changed to "optimal solution" in Theorem 6.26 and 6.27.

- Possibilities to improve the efficiency of the modified interior point methods:
  - Generate initial points in a more sophisticated way.
  - Utilize information of the previous runs.
  - What can we do if the algorithm terminates with a certificate? Does the algorithm have to stop by all means if it finds a certificate? Can we find another point or Newton direction in this case, which is suitable to continue the algorithm in a proper way? (Note that we use the special properties of the matrix only locally.) Maybe a proper perturbation of the problem can help.

We finish the thesis with a small example, which illustrates well the difficulties of the modified interior point methods at the solution of LCPs with arbitrary matrices . Let us consider the following LCP

$$\left.\begin{aligned} x + y &\leq 1 \\ y &\leq 1 \\ x, y &\geq 0 \\ x\,y &= \mathbf{0} \end{aligned}\right\} \tag{7.1}$$

It is easy to see, that this problem has three solutions: $(0,0)$, $(1,0)$ and $(0,1)$, and the second condition is redundant. Figure 7 shows that there are only two central paths. The

blue curve tends to the solution $(0,0)$, while the green one to the point $(1,0)$. We try to solve the problem with the modified predictor-corrector interior point algorithm from different initial points. Figure 7 illustrates how the algorithm terminates starting from a given point. When we start from a point in the blue area, the algorithm finds the solution $(0,0)$. Similarly, the algorithm determines the solution (1,0) starting from the green area. When the initial point is on the red line, the algorithm terminates with a certificate, that the matrix is not $\mathcal{P}_0$, and finally, starting from a red point (beside the red lines) we get a certificate that the matrix is not sufficient. Thus the algorithm never finds the third solution $(0,1)$.

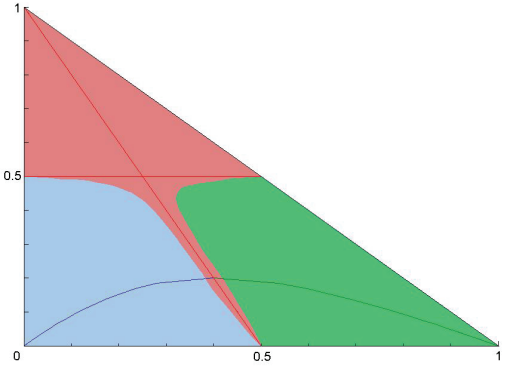


Figure 7.1: Regions of attraction for modified predictor-corrector algorithm on problem (7.1).

Let us perturb the problem (7.1). We decrease the right hand side of the second condition to 0.9, thus the second condition is not redundant any more. The perturbed problem has four solutions: $(0,0)$, $(1,0)$, $(0,0.9)$ $(0.1,0.9)$, and there are four central paths. However, the two new solutions have very short central paths, which lie very close to the boundary of the set of feasible solutions. It means, that these two solutions could be determined by the modified predictor-corrector algorithm, but the initial point has to be very close to these solutions.

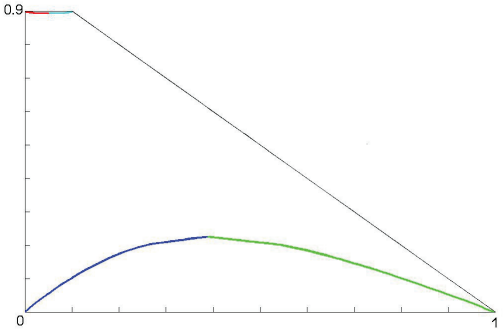


Figure 7.2: Central paths for the perturbed problem of (7.1).

Summary

The thesis is concerned with the interior point methods for linear complementarity problems (LCP). The LCP is an $\mathbb{NP}$-complete problem, when we have no information about the coefficient matrix of the problem. In the literature there are efficient algorithms to solve the LCP, but only when the matrix of the problem belongs to a certain special matrix class. An arbitrary LCP problem can not be solved in polynomial time, not even if the coefficient matrix belongs to the sufficient matrix class, since it can not be decided in polynomial time.

Our aim was to construct an algorithm, which provides some kind of information about the LCP problem in polynomial time – it gives a solution in the best case. We take well-known interior point methods (for LCPs with $\mathcal{P}_*(\kappa)$-matrices) for the basis of the new algorithm. The modified algorithms are polynomial time methods. Our motivation was not only to handle the LCP problems in applications efficiently, but it was also theoretical. Based on the modified interior point methods we gave constructive proofs for new EP theorems.

The thesis can be divided into two parts. In the first part we mainly collect well-known results according to the LCPs and interior point methods. At the end of the Introduction we collect some well known problems, which can be reformulated as an LCP to illustrate how important an efficient algorithm for handling LCPs with arbitrary matrices will be in practice. Then we deal with some matrix classes related to LCPs, which are important for our purposes. Finally, we give an overview of the theory of interior point methods.

In the second part of the thesis first the dual of the LCP is considered. We show that the dual LCP can be solved in polynomial time if the matrix is row sufficient, and we give an EP type theorem based on this complexity result.

We generalize the Mizuno–Todd–Ye predictor-corrector algorithm for LCPs with $\mathcal{P}_*(\kappa)$-matrices and show that the algorithm preserves its nice property, that is, if the two used neighbourhoods are well chosen, then after each predictor step we can return to the smaller neighbourhood of the central path with only one corrector step. Then we take a further step in the generalization of interior point methods. As we have already pointed out, usually we do not know anything about the matrix of a real life problem, moreover, in most cases it is not a $\mathcal{P}_*(\kappa)$-matrix. Therefore, we construct modified interior point methods (a long step path-following, an affine scaling and a predictor-corrector one), which can handle any LCPs. These algorithms either solve the problem or its dual (in the latter case proving that the problem has no solution), or give a polynomial size certificate that the matrix is not a $\mathcal{P}_*(\tilde{\kappa})$-matrix with an arbitrary but a priori chosen $\tilde{\kappa}$. We present preliminary computational results of the modified long-step path-following and predictor-corrector interior point methods on a special variant of the Arrow–Debreu market equilibrium problem.

Összefoglalás

A doktori értekezés a lineáris komplementaritási feladatokra (LCP) kidolgozott belsőpontos algoritmusokkal foglalkozik. Az LCP feladat egy $\mathbb{NP}$-teljes probléma ha nem áll rendelkezésünkre semmilyen információ a feladat mátrixáról. Csak abban az esetben ismertek hatékony megoldó módszerek, ha a feladat mátrixa valamilyen speciális tulajdonsággal bír. Egy tetszőleges LCP feladatot nem tudunk polinom időben megoldani, még abban az esetben sem, ha az együtthatómátrix elégséges, ugyanis ezt nem tudjuk polinom időben leellenőrizni.

Célunk olyan algoritmus megalkotása volt, mely polinom időben szolgáltat információt az LCP feladatról – a legjobb esetben egy megoldását kapjuk. Az új algoritmus jól ismert, $\mathcal{P}_*(\kappa)$ mátrixú LCP feladatokra kidolgozott belsőpontos algoritmusokon alapszik. A módosított algoritmusok polinomiálisak. A motivációnk nem csak a gyakorlati problémákból származó LCP feladatok hatékony kezelése, hanem elméleti jellegű is. A módosított belsőpontos algoritmusok segítségével konstruktív bizonyítást adunk új EP típusú tételekre.

A doktori dolgozat két részre bontható. Az első részben főként az LCP feladatokkal és a belsőpontos algoritmusokkal kapcsolatos már ismert eredményeket gyűjtöttük össze. A Bevezetés végén néhány jól ismert, LCP feladatra visszvezethető feladatot mutatunk be, ezzel is alátámasztva egy hatékony, tetszőleges LCP feladatokat kezelő algoritmus fontosságát. Ezt követően néhány, az LCP feladatokkal kapcsolatos, számunkra fontos mátrixosztályt tárgyalunk. Végül a belsőpontos algoritmusok elméletéről adunk egy áttekintést.

A dolgozat második felében először az LCP feladat duálját vizsgáljuk. Megmutatjuk, hogy a duál LCP feladat polinom időben megoldható, ha a feladat mátrixa sorelégséges. Továbbá ezen komplexitási eredményre támaszkodva egy EP típusú tételt fogalmazunk meg.

A Mizuno–Todd–Ye prediktor-korrektor algoritmust $\mathcal{P}_*(\kappa)$ mátrixú LCP feladatokra általánosítjuk, megmutatjuk, hogy az algoritmus megőrzi azon szép tulajdonságát, miszerint a két környezet megfelelő összehangolása esetén minden prediktor lépés után egyetlen korrektor lépéssel visszatérhetünk a szűkebb környezetbe. Ezután egy további lépést teszünk a belsőpontos algoritmusok általánosításában. Egy valós gyakorlati probléma esetén általában nem tudunk semmit a feladat mátrixáról, sőt legtöbbször az nem is $\mathcal{P}_*(\kappa)$ mátrix. Ezért úgy módosítottuk a belsőpontos algorimusokat (hosszúlépéses útkövető, affin skálázású, illetve prediktor-korrektor algoritmusokat), hogy tetszőleges LCP feladatokon fussanak. Ezen algoritmusok vagy megoldják a feladatot, vagy megoldják a duál feladatot (bizonyítva ezzel, hogy az eredeti feladatnak nincs megoldása), vagy pedig egy polinom méretű tanút szolgáltatnak, mely igazolja, hogy a mátrix nem $\mathcal{P}_*(\tilde{\kappa})$ mátrix, ahol $\tilde{\kappa}$ tetszőleges, de előre rögzített. Végül ismertetjük a módosított hosszúlépéses útkövető, illetve prediktor-korrektor algoritmusok futási eredményeit az Arrow–Debreu piaci egyensúlyi probléma egy speciális feladatán.

Index

Bibliography

[1] A. Al-Nowaihi. P-matrices: An equivalent characterization. *Journal of Algebra*, 112:385–387, 1988.

[2] K. J. Arrow and G. Debreu. Existence of an equilibrium for competitive economy. *Econometrica*, 22:265–290, 1954.

[3] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. *Nonlinear programming*. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, third edition, 2006. Theory and algorithms.

[4] R. Bentzel. Presentation Speech for the 1972 Nobel Prize in Economic Sciences. http://nobelprize.org/nobel_prizes/economics/laureates/1972/presentation-speech.html.

[5] K. Cameron and J. Edmonds. Existentially polytime theorems. In *Polyhedral combinatorics (Morristown, NJ, 1989)*, volume 1 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 83–100. Amer. Math. Soc., Providence, RI, 1990.

[6] M. Cao and M. C. Ferris. P_c-matrices and the linear complementarity problem. *Linear Algebra and Its Applications*, 246:299–312, 1996.

[7] L. Chen, Y. Ye, and J. Zhang. A note on equilibrium pricing as convex optimization. in WINE´07 (http://www.stanford.edu/∼yyye/convexpricingms.pdf).

[8] S. J. Chung. NP-completeness of the linear complementarity problem. *Journal of Optimization Theory and Applications*, 60(3):393–399, March 1989.

[9] B. Codenotti, A. Saberi, K. Varadarajan, and Y. Ye. The complexity of equilibria: Hardness results for economies via a correspondence with games. *Theoretical Computer Science*, 408:188–198, 2008.

[10] R. W. Cottle. A field guide to matrix classes found in the linear complementarity problem. http://www.stanford.edu/dept/MSandE/people/emeriti/cottle/fieldguide3.PDF.

[11] R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem*. Computer Science and Scientific Computing. Academic Press, Inc., Boston, 1992.

[12] R. W. Cottle, J.-S. Pang, and V. Venkateswaran. Sufficient matrices and the linear complementarity problem. *Linear Algebra and Its Applications*, 114/115:231–249, 1989.

[13] R.W. Cottle. *Nonlinear programs with positively bounded Jacobians*. PhD thesis, Department of Mathematics, University of California, Berkeley, 1964.

[14] G. E. Coxson. The P-matrix problem is co-NP-complete. *Mathematical Programming*, 64:173–178, 1994.

[15] Zs. Csizmadia. *New pivot based methods in linear optimization, and an application in the petroleum industry*. PhD thesis, Department of Operations Research, Eötvös Loránd University of Sciences, Budapest, Hungary, 2007.

[16] Zs. Csizmadia and T. Illés. New criss-cross type algorithms for linear complementarity problems with sufficient matrices. *Optimization Methods and Software*, 2:247–266, 2006.

[17] C. Dang, Y. Ye, and Z. Zhu. An interior-point path-following algorithm for computing a Leontief economy equilibrium. Technical report, Management Science and Engineering, Stanford University, March 2008. http://www.stanford.edu/∼yyye/Homotopy5.pdf.

[18] D. den Hertog, C. Roos, and T. Terlaky. The linear complementarity problem, sufficient matrices and the criss-cross method. *Linear Algebra and Its Applications*, 187:1–14, 1993.

[19] I.I. Dikin. Iterative solution of problems of linear and quadratic programming. *Doklady Academii Nauk SSSR*, 174:747–748, 1967.

[20] E. Eisenberg. Aggregation of utility functions. *Management Sciences*, 7(4):337–350, 1961.

[21] E. Eisenberg and D. Gale. Consensus of subjective probabilities: The pari-mutuel method. *Annals Of Mathematical Statistics*, 30:165–168, 1959.

[22] L. Fang. On the spectra of P- and P_0-matrices. *Linear Algebra and Its Applications*, 119:1–25, 1989.

[23] M. Fiedler and V. Pták. On matrices with non-positive off-diagonal elements and positive principal minors. *Czechoslovak Mathematical Journal*, 12:382–400, 1962.

[24] M. Fiedler and V. Pták. Some generalizations of positive definiteness and monotonicity. *Numerische Mathematik*, 9:163–172, 1966.

[25] K. Fukuda, M. Namiki, and A. Tamura. EP theorems and linear complementarity problems. *Discrete Applied Mathematics*, 84:107–119, 1998.

[26] K. Fukuda and T. Terlaky. Linear complementarity and oriented matroids. *Journal of the Operations Research Society of Japan*, 35:45–61, 1992.

[27] D. Gale. *The Theory of Linear Economic Models*. McGraw-Hill, New York, 1960.

[28] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, New York, 1979.

[29] A.J. Goldman and A.W. Tucker. Theory of linear programming. In H.W. Kuhn and A.W. Tucker, editors, *Linear Inequalities and Related Systems*, number 38 in Annals of Mathematical Studies, pages 53–97, Princeton, New Jersey, 1956. Princeton University Press.

[30] R.C. Griffin, J.M. Montgomery, and Rister M.E. Selecting functional form in production function analysis. *Western Journal of Agricultural Economics*, 12(2):216–227, 1987.

[31] S. M. Guu. Sign reversing and matrix classes. *Journal of Optimization Theory and Applications*, 89(2):373–387, May 1996.

[32] S. M. Guu and R. W. Cottle. On a subclass of P_0. *Linear Algebra and Its Applications*, 223/224:325–335, 1995.

[33] D. Hershkowitz. On the spectra of matrices having nonnegative sums of principal minors. *Linear Algebra and Its Applications*, 55:81–86, 1983.

[34] D. Hershkowitz and A. Berman. Localization of the spectra of P- and P_0-matrices. *Linear Algebra and Its Applications*, 52/53:383–397, 1983.

[35] R.A. Horn and C.R. Johnson. *Matrix analysis*. Cambridge University Press, 1990.

[36] T. Illés, J. Peng, C. Roos, and T. Terlaky. A strongly polynomial rounding procedure yielding a maximally complementary solution for $P_*(\kappa)$ linear complementarity problems. *SIAM Journal on Optimization*, 11(2):320–340, 2000.

[37] T. Illés, C. Roos, and T. Terlaky. Polynomial affine–scaling algorithms for $P_*(\kappa)$ linear complementarity problems. In P. Gritzmann, R. Horst, E. Sachs, and R. Tichatschke,

editors, *Recent Advances in Optimization*, number 452 in Lecture Notes in Economics and Mathematical Systems, pages 119–137. Proceedings of the 8-th French-German Conference on Optimization, Trier, July 21-26, 1996, Springer, 1997.

[38] T. Illés, C. Roos, and T. Terlaky. Simple approach for the interior point theory of linear complementarity problems with $\mathcal{P}_*$ matrices. Unpublished manuscript, 1998.

[39] T. Illés and M. Wenzel. A survey on properties of P- and P_*-matrices. Unpublished manuscript, 1999.

[40] T. Illés and M. Nagy. A Mizuno–Todd–Ye type predictor-corrector algorithm for sufficient linear complementarity problems. *European Journal of Operation Research*, 181(3):1097–1111, 2007.

[41] T. Illés, M. Nagy, and T. Terlaky. Interior point algorithms. In A. Iványi, editor, *Algorithms of Informatics II.*, pages 1230–1297. ELTE Eötvös Kiadó, 2005. [In Hungarian].

[42] T. Illés, M. Nagy, and T. Terlaky. Polynomial interior point algorithms for general linear complementarity problems. 2008. Submitted.

[43] T. Illés, M. Nagy, and T. Terlaky. An EP theorem for dual linear complementarity problems. *Journal of Optimization Theoryan Applications*, 140(2):233–238, 2009. http://www.springerlink.com/content/w3x6401vx82631t2/fulltext.pdf.

[44] T. Illés, M. Nagy, and T. Terlaky. A polynomial path-following interior point algorithm for general linear complementarity problems. *Journal of Global Optimization*, In print, accepted: 26 August 2008.
Online version: http://www.springerlink.com/content/547557t2077t2525/.

[45] B. Jansen, C. Roos, and T. Terlaky. A family of polynomial affine scaling algorithms for positive semi-definite linear complementarity problems. *SIAM Journal on Optimization*, 7:126–140, 1996.

[46] M. J. M. Jansen. On the structure of the solution set of a linear complementarity problem. *Cahiers du Centre d'Études de Recherche Opérationnelle*, 1-2:41–48, 1983.

[47] J. Ji, F. A. Potra, and R. Sheng. A predictor-corrector method for the P_*-matrix LCP from infeasible starting points. *Optimization Methods and Software*, 6:109–126, 1995.

[48] C. Jones and M. S. Gowda. On the connectedness of solution sets in linear complementarity problems. *Linear Algebra and Its Applications*, 272:33–44, 1998.

[49] W. Karush. Minima of functions of several variables with inequalities as side constraints. Master's thesis, Department of Mathematics, University of Chicago, Chicago, Illinois, 1939.

[50] R. B. Kellogg. On complex eigenvalues of M and P matrices. *Numerische Mathematik*, 19:170–175, 1972.

[51] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise. *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, volume 538 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1991.

[52] M. Kojima, S. Mizuno, and A. Yoshise. A polynomial-time algorithm for a class of linear complementarity problems. *Mathematical Programming*, 44:1–26, 1989.

[53] C.E. Lemke and J.T. Howson. Equilibrium points of bimatrix games. *SIAM Journal on Applied Mathematics*, 12:413–423, 1964.

[54] N. Megiddo. Pathways to the optimal set in linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming: Interior Point and Related Methods*, pages 131–158. Springer Verlag, New York, 1989. Identical version in Proceedings of the 6th Mathematical Programming Symposium of Japan, Nagoya, Japan, pp. 1-35, 1986.

[55] J. Miao. A quadratically convergent $O((\kappa+1)\sqrt{n}L)$-iteration algorithm for the $P_*(\kappa)$-matrix linear complementarity problem. *Mathematical Programming*, 69:355–368, 1995.

[56] S. Mizuno, M. J. Todd, and Y. Ye. On adaptive-step primal-dual interior-point algorithms for linear programming. *Mathematics of Operations Research*, 18:964–981, 1993.

[57] K.-G. Mäler. Presentation Speech for the 1983 Nobel Prize in Economic Sciences. (http://nobelprize.org/nobel_prizes/economics/laureates/1983/presentation-speech.html.

[58] G. S. R. Murty, T. Parthasarathy, and M. Sabatini. Lipschitzian Q-matrices are P-matrices. *Mathematical Programming*, 74:55–58, 1996.

[59] G. S. R. Murty, T. Parthasarathy, and B. Sriparna. Constructive characterization of Lipschitzian Q_0-matrices. *Linear Algebra and Its Applications*, 252:323–337, 1997.

[60] K. G. Murty. *Linear and Combinatorial Programming*. John Wiley & Sons, Inc., New York-London-Sydney, 1976.

[61] K. G. Murty. *Linear Complementarity, Linear and Nonlinear Programming*. Heldermann Verlag, Berlin, 1988.

[62] K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39:117–129, 1987.

[63] J. Peng, C. Roos, and T. Terlaky. New complexity analysis of primal–dual Newton methods for $P_*(\kappa)$ linear complementarity problems. In J.B.G. Frenk, C. Roos, T. Terlaky, and S. Zhang, editors, *High Performance Optimization Techniques*, pages 245–266, Kluwer Academic Publishers, Dordrecht, 1999.

[64] S. Poljak and J. Rohn. Checking robust nonsingularity is NP-hard. *Mathematics of Control, Signals and Systems*, 6:1–9, 1993.

[65] F. A Potra. The Mizuno-Todd-Ye algorithm in a larger neighborhood of the central path. *European Journal of Operational Research*, 143:257–267, 2002.

[66] F. A. Potra and X. Liu. Predictor-corrector methods for sufficient linear complementarity problems in a wide neighborhood of the central path. *Optimization Methods & Software*, 20(1):145–168, 2005.

[67] F. A. Potra and R. Sheng. Predictor-corrector algorithm for solving $P_*(\kappa)$-matrix LCP from arbitrary positive starting points. *Mathematical Programming*, 76(1):223–244, 1996.

[68] F. A. Potra and R. Sheng. A large step infeasible interior point method for the P_*-matrix LCP. *SIAM Journal on Optimization*, 7(2):318–335, 1997.

[69] F. A. Potra and R. Sheng. Superlinearly convergent infeasible interior point algorithm for degenerate LCP. *Journal of Optimization Theory and Application*, 97(2):249–269, 1998.

[70] C. Roos, T. Terlaky, and J.-Ph Vial. *Theory and Algorithms for Linear Optimization, An Interior Point Approach*. Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, New York, 1997. Second edition: *Interior Point Methods for Linear Optimization*, Springer, New York, 2006.

[71] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, Inc., 1976.

[72] S. Sahni. Computationally related problems. *SIAM Journal on Computing*, 3:262–279, 1974.

[73] Gy. Sonnevend. An "analytical center" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. *Lecture Notes in Control and Information Sciences*, 84:866–876, 1985.

[74] Gy. Sonnevend. A new method for solving a set of linear (convex) inequalities and its applications. Technical report, Department of Numerical Analysis, Institute of Mathematics, Eötvös Loránd University, Budapest, 1985.

[75] Gy. Sonnevend, J. Stoer, and G. Zhao. On the complexity of following the central path of linear programs by linear extrapolation. *Methods of Operations Research*, 63:19–31, 1989.

[76] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. Number 12 in Texts in Applied Mathematics. Springer Verlag, New York, 3rd edition, 2002.

[77] Terlaky. T. An easy way to teach interior point methods. *European Journal of Operational Research*, 130:1–19, 2001.

[78] P. Tseng. Co-NP-completeness of some matrix classification problems. *Mathematical Programming*, 88:183–192, 2000.

[79] H. Väliaho. Criteria for sufficient matrices. *Linear Algebra and Its Applications*, 233:109–129, 1996.

[80] H. Väliaho. P_*-matrices are just sufficient. *Linear Algebra and Its Applications*, 239:103–108, 1996.

[81] H. Väliaho. Determining the handicap of a sufficient matrix. *Linear Algebra and Its Applications*, 253:279–298, 1997.

[82] S. A. Vavasis. Quadratic programming is in NP. *Information Processing Letters*, 36(2):73–77, 1990.

[83] V.V. Vazirani. *Approximation Algorithms*. Springer-Verlag, Berlin, 2001.

[84] L. Walras. *Elements of Pure Economics, or the Theory of Social Wealth*. 1874. (1899, 4th edition; 1926 revised edition, 1954 English translation).

[85] Y. Ye. Exchange market equilibria with Leontief's utility: freedom of pricing leads to rationality. *Theoretical Computer Science*, 378(2):134–142, 2007.

[86] Y. Ye. A path to the Arrow-Debreu competitive market equilibrium. *Mathematical Programming*, 111(1-2):315–348, 2008.

[87] Y. Ye and K. Anstreicher. On quadratic and $O(\sqrt{n}L)$ convergence of a predictor-corrector algorithm for LCP. *Mathematical Programming*, 62:537–551, 1993.

[88] Z. Zhu and Y. Ye. A FPTAS for computing a symmetric Leontief competitive economy equilibrium. Technical report, Management Science and Engineering, Stanford University, February 2008. http://www.stanford.edu/∼yyye/SymLeontief31.pdf.