Interior-Point Algorithms Based on Primal-Dual Entropy

by

Shen Luo

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Mathematics
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2006

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

We propose a family of search directions based on primal-dual entropy in the context of interior point methods for linear programming. This new family contains previously proposed search directions in the context of primal-dual entropy. We analyze the new family of search directions by studying their primal-dual affine-scaling and constant-gap centering components. We then design primal-dual interior-point algorithms by utilizing our search directions in a homogeneous and self-dual framework. We present iteration complexity analysis of our algorithms and provide the results of computational experiments on NETLIB problems.

Acknowledgements

I would like to express my deepest thank to my supervisor, Dr. Levent Tunçel. Without his continues guidance and support, I could not complete this thesis. I would also like to thank the two readers, Dr. Henry Wolkowicz and Dr. Chek Beng Chua, for their comments and careful reading of the draft.

Thanks my parents for their love and continues encouragement. Although they were not in Canada when I was writing the thesis, I can always feel their support.

Last, I want to express my sincere appreciation to my wife, Kun He, for her love, encouragement, and being my company in this hard but fruitful journey in Canada.

Contents

Chapter 1

Introduction

The entropy concept has played important roles in many areas such as statistical mechanics, thermodynamics and information theory. The concept we refer to in this thesis is the informational entropy. This definition was proposed in a paper of Shannon [22]. This entropy stands for a quantitative measure of the amount of uncertainty about the possible outcome of a probabilistic experiment.

Consider a probabilistic experiment having n discrete possible final states $a_1, \ldots, a_n$ with the respective discrete probabilities $p_1, \ldots, p_n$ satisfying the following conditions:
$p_i \geq 0$, $i \in \{1, 2, \ldots, n\}$, and $\sum_{i=1}^{n} p_i = 1$. The *informational entropy* is defined as below:

$$S = -k \sum_{i=1}^{n} p_i \ln(p_i)$$

where k is a positive constant depending on a suitable choice for the unit of measure and it is defined that $0 \ln 0 := 0$.

In this thesis, $\mathbf{e}$ stands for the logarithmic constant and e stands for the vector of all ones whose dimension will be clear from the context. $\ln(x)$ where $x \in \mathbb{R}^n$ stands for the vector $(\ln(x_1), \ln(x_2), \ldots, \ln(x_n))^T$. We use lower case letters such as x, s, v for vectors and upper case letters such as A, X, S, V for matrices. If x is defined, then X is the diagonal matrix with $X_{ii} = x_i$ for all i. R^n_+ stands for the n dimensional nonnegative real vector. R^n_{++} stands for the n dimensional positive real vector.

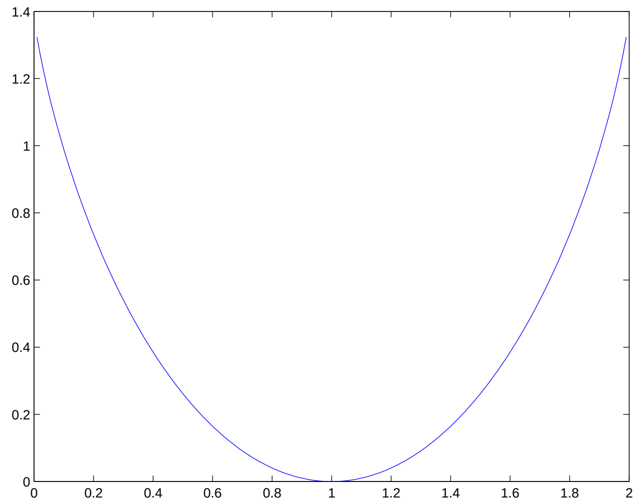


Figure 1.1: Graph of $f(x)$ in 2-dimensional case (projection on one dimension)

Lemma 1.1. *Let* $f : \{x \in \mathbb{R}^n_{++} : \sum_{i=1}^{n} x_i = n\} \to \mathbb{R}$ *be defined as:*

$$f(x) := \sum_{i=1}^{n} x_i \ln(x_i).$$

Then f *is strictly convex on* $\mathbb{R}^n_{++}$ *and its unique minimizer is* $x^* := e$.

Proof. We know $\nabla f(x) = e + \ln(x)$, $\nabla^2 f(x) = X^{-1}$ is positive definite since $x \in \mathbb{R}^n_{++}$. Therefore, f is a strictly convex function.
We consider the constrained problem:

$$\begin{aligned} &\min f(x) \\ \text{s.t.} \quad &\sum_{i=1}^{n} x_i = n, \quad (x > 0). \end{aligned}$$

The Lagrangean for this problem $\mathcal{L}(x, \lambda) = f(x) + \lambda \left(n - \sum_{i=1}^{n} x_i \right)$, and $\nabla^2_{xx} \mathcal{L}(x, \lambda) = \nabla^2 f(x) \succ 0$ for all $x \in \mathbb{R}^n_{++}$. We find that ($x = e$, $\lambda = 1$) satisfies the KKT conditions. Since $\nabla^2 \mathcal{L}(x, \lambda) \succ 0$, we know $x = e$ will be the only global minimizer for this problem by the property of strict convexity. $\square$

Shannon's Entropy has been applied to many optimization problems such as transportation problems [7], Linear Programming (LP) problems [8], some infinite dimensional convex optimization problems [5], [6], [17] and convex constrained programming problems [13]. In the paper [7] and [8], The author added an entropy constraint to the original LP problem and simplified the problem to a dual problem of Lagrange multipliers. He developed an iterative procedure to solve the Lagrange multipliers for this dual problem. In the book [9], the frame of entropy maximization technique via convex programming was discussed in detail. The entropy function is used widely in infinite dimensional convex optimization problems in the areas such as spectral estimation and crystallography [5], [6], [17]. The conjugate of the Shannon Entropy is the Burg Entropy, which has the formulation $-\int_0^\infty \ln(x)dx \equiv -\sum_{i=1}^{n} \ln(x_i)$. When we are facing some infinite dimensional convex optimization problems such as the entropy maximization problems, instead of solving directly, we can solve the dual problem which is a convex problem of finite dimensions [5]. Then the dual problem is known to be easy to solve. In the paper [13], surrogate Lagrangean technique [10] and maximum entropy criterion is used to develop several entropy based algorithms for convex constrained programming problems.

The entropy function is also used in other LP solving techniques such as the barrier term in augmented Lagrangean method. The entropy function is almost always used on only one of the primal or dual forms, i.e., it is applied to the primal problem or the dual problem separately. In this thesis, we will explore the Primal-Dual entropy for LP solving.

We consider the LP problems in the following standard form:

$$\begin{array}{lll} \text{(P)} & \text{minimize} & c^T x \\ & \text{subject to} & Ax = b, \quad x \geq 0, \end{array}$$

where c is in $\mathbb{R}^n$, A is in $\mathbb{R}^{m \times n}$, and b is in $\mathbb{R}^m$. In this thesis, all vectors are column vectors.

Without loss of generality, we always assume A has full row rank. If A does not have full row rank, then either the linear system $Ax = b$ has no solution, or $Ax = b$ has some redundant equations. We can remove those redundant equations without changing the solution set of $Ax = b$, and then we have a new matrix which is of full row rank.

We use (P) to denote the primal problem, the corresponding dual problem is:

$$\begin{array}{lll} \text{(D)} & \text{maximize} & b^T y \\ & \text{subject to} & A^T y \leq c. \end{array}$$

$s := c - A^T y$ is called the *dual slack vector*. By definition, for a feasible solution (y, s) of (D), y uniquely identifies s. Conversely, since A has full row rank, s part of a feasible solution of (D) uniquely identifies y. Sometimes, we only refer to s instead of (y, s) when we talk about a feasible solution of (D).

We define the *Primal-dual entropy* for the above problem as $\sum_{j=1}^{n} x_j s_j \ln(x_j s_j)$. And we will discuss it in the context of Interior Point Methods in the next chapter.

We use the following notations to denote various sets related to the feasible regions of (P) and (D):

$$\begin{aligned} \mathcal{F}(P) &:= \{x \in \mathbb{R}^n_+ : Ax = b\}, \\ \mathcal{F}_+(P) &:= \{x \in \mathbb{R}^n_{++} : Ax = b\}, \\ \mathcal{F}(D) &:= \{s \in \mathbb{R}^n_+ : A^T y + s = c \text{ for some } y \in \mathbb{R}^m\}, \\ \mathcal{F}_+(D) &:= \{s \in \mathbb{R}^n_{++} : A^T y + s = c \text{ for some } y \in \mathbb{R}^m\}, \\ \mathcal{F}_+ &:= \mathcal{F}_+(P) \oplus \mathcal{F}_+(D). \end{aligned}$$

We say that $\overline{x} \in \mathbb{R}^n$ is *primal feasible* if $\overline{x} \in \mathcal{F}(P)$; $\overline{x}$ is *strictly primal feasible* if $\overline{x} \in \mathcal{F}_+(P)$. Similarly, $\overline{s} \in \mathcal{F}(D)$ is called *dual feasible*; $\overline{s} \in \mathcal{F}_+(D)$ is called *strictly dual feasible.*

Next we'd like to introduce the *reparameterization of nonlinear systems* and the effects of reparameterization on Newton's Method. We will use this technique to derive new search directions in Chapter 3.

Define $F_1 : \mathbb{R}^{n_1} \rightarrow \mathbb{R}^n$, $F_2 : \mathbb{R}^{n_2} \rightarrow \mathbb{R}^n$ such that $n_1 + n_2 = n$. We will try to solve the system: $F_1(x) - F_2(y) = 0 \Leftrightarrow F_1(x) = F_2(y)$.

Define a reparameterization function $\psi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $\psi(F_1(x)) = \psi(F_2(y))$. If we select ψ as a bijection, then the two systems $F_1(x) = F_2(y)$ and $\psi(F_1(x)) = \psi(F_2(y))$ are equivalent. To apply Newton's method, we need F_1, F_2 to be continuously differentiable. We require the same of ψ as well. Depending on the choice of ψ which depends on F_1, F_2, Newton's method has different performance when applied to the original system and the new formulation.

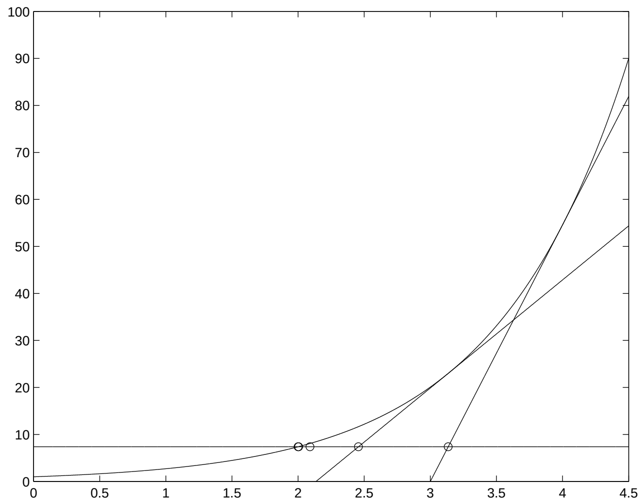


Figure 1.2: Newton's method for $f(x) = \mathbf{e}^x - \mathbf{e}^2$. The curve is $f(x) = \mathbf{e}^x$ and the horizontal line stands for $f(x) = \mathbf{e}^2$. The dots stand for the x_i we find in each iteration.

We observe that if we apply Newton's method to the new system, it may be better than applying Newton's method directly to the original system at least in some cases. We give out two examples as blew and some more examples can be found in [3].

Example 1: We choose the original system as $\mathbf{e}^x - \mathbf{e}^2 = 0$. If we solve this system by Newton's method directly, we may choose the starting point $x = 4$ and calculate the derivative of $f(x) = \mathbf{e}^x - \mathbf{e}^2$ at point x_i in every iteration until the distance $\|x_i - x_{i+1}\| \leq \epsilon$ where ϵ is the desired accuracy (here we set $\epsilon = 10^{-10}$). We need 7 iterations to find the solution $x = 2($ as shown in the Figure above). If we apply $\psi(x) = \ln(x)$ to the system $\mathbf{e}^x = \mathbf{e}^2$, then we can deduce the solution at once.

Example 2: Consider the system

$$
\begin{aligned}
&Xs = 5e, \\
&x_1 + x_2 = 10, \\
&s_1 + s_2 = 10,
\end{aligned}
$$

where $x \in \mathbb{R}^2_{++}$, $s \in \mathbb{R}^2_{++}$.

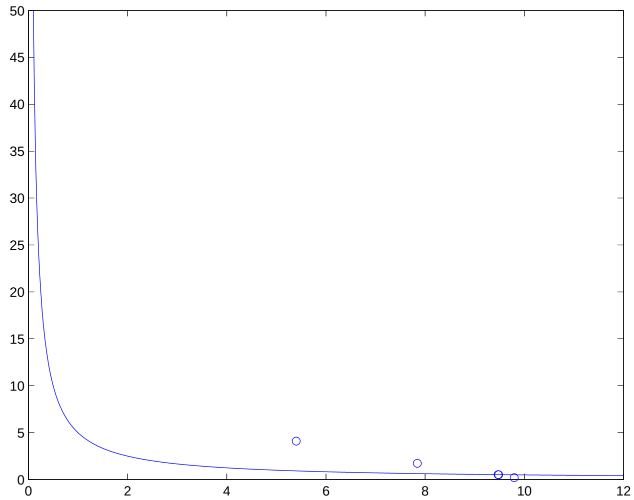


Figure 1.3: Newton's method applied to original problem (projection on the (x_2, s_2)-space).

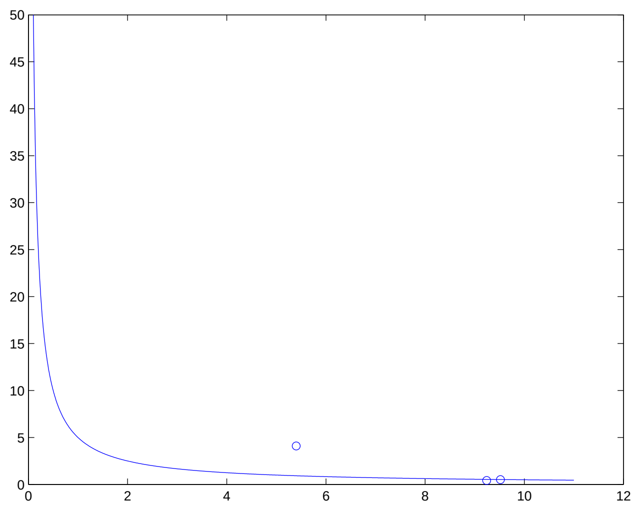


Figure 1.4: Newton's method applied to system after we reparemeterize it with $\psi(x) = \ln(x)$ (projection on the (x_2, s_2)-space).

We select the starting point $x^0 := [4.6, 5.4]^T$, $s^0 := [5.9, 4.1]^T$. If we apply Newton's method directly, then we need 5 iterations to achieve the accuracy of 10^{-4}. While if we apply Newton's method to the system after changing $Xs = 5e$ to $\ln(Xs) = \ln 5e$, we need only 3 iterations to achieve the same accuracy. We can get a direct impression from the figures below.

These examples are extremely simple and biased; however, we will show that reparameterizations can also have positive effect on much more complicated systems in this thesis.

This thesis has been structured in five chapters. The first one is the introduction, which is what we are going through now. The second chapter presents the fundamentals of Interior Point Methods which are necessary for the theories developed in the later chapters. The third chapter generalizes a class of entropic search directions and then presents some technical lemmas and complexity analysis of two existing entropic search directions when applied to different neighborhoods or models. The fourth chapter analyzes the iteration complexity for the algorithms using the generalized search directions from the entropic family. The fifth chapter presents the results of the computational experiments comparing the performance of the algorithms based on different search directions in this family within the wide neighborhood.

Chapter 2

Fundamentals of Interior Point Methods

2.1 Central Path and its Neighborhoods

Before we describe the primal-dual algorithms, we first define some of the main ingredients:

- central path and
- various neighborhoods of the central path.

We assume for the next definition that $\mathcal{F}_+ \neq \emptyset$.
The central path $\mathcal{C}$ is a set of strictly feasible points that plays an important role in primal-dual interior-point method. It is parameterized by a scalar $\mu > 0$, and each point $(x_\mu, y_\mu, s_\mu) \in \mathcal{C}$ solves the following system:

$$\text{(CP)} \qquad \begin{aligned} A^T y + s = c, &\quad s > 0, \\ Ax = b, &\quad x > 0, \\ Xs = \mu e. & \end{aligned}$$

Note that every solution (x_μ, y_μ, s_μ) of (CP) satisfies $\mu = \frac{x^T s}{n}$ (implied by $Xs = \mu e$). From now on, when (x, s) is clear from the context, μ denotes $\frac{x^T s}{n}$.

We define the *central path* as:

$$\mathcal{C} = \{(x_\mu, y_\mu, s_\mu) : \mu > 0\}.$$

It can be shown [29] that $(x_\mu, y_\mu, s_\mu) \in \mathcal{C}$ is defined uniquely for each $\mu > 0$.

As $\mu \to 0^+$, the central path converges to an optimal primal-dual solution of the original problems (P) and (D). The central path thus guides us to a solution along a route which keeps all x and s components strictly positive and decreases the pairwise products $x_j s_j$, $j = 1, 2, \ldots, n$, to zero at roughly the same rate.

For a given $\beta \geq 0$, we define the following neighborhoods of the central path:
$\mathcal{N}_2(\beta) := \{(x, s) \in \mathcal{F}_+ : \|\frac{Xs}{\mu} - e\|_2 \leq \beta\}$;
$\mathcal{N}_\infty(\beta) := \{(x, s) \in \mathcal{F}_+ : \|\frac{Xs}{\mu} - e\|_\infty \leq \beta\}$;
$\mathcal{N}_\infty^-(\beta) := \{(x, s) \in \mathcal{F}_+ : \frac{x_j s_j}{\mu} \geq 1 - \beta$, for all j $\}$.

$\mathcal{N}_2(\beta)$ is also called the *narrow neighborhood*, $\mathcal{N}_\infty^-(\beta)$ is also called the *wide neighborhood*.

For $\beta \geq \frac{1}{2}$, the following neighborhood was also used:
$\mathcal{N}_E(\beta) := \{(x, s) \in \mathcal{F}_+ : \frac{1}{2} - \beta \leq \ln(\frac{x_j s_j}{\mu}) \leq \frac{1}{2} + \beta$, for all j $\}$.

2.2 The Classical Primal-Dual Interior-Point Methods

After the groundbreaking paper of Karmarkar [12], Interior-Point Methods(IPMs) for linear programming became a very active area of research. The algorithm proposed in that paper is the first proven polynomial-time IPM for LP solving. After the paper was published, many researchers were inspired and many interesting theoretical results for IPMs burst out since then. Interior Point Methods are now among the most effective methods for solving LP problems and their various generalizations. For a survey, we refer to one recent book on this subject [29]. In this thesis, we deal with so-called primal-dual IPMs.

Among all variants of interior-point methods, symmetric primal-dual interior-point methods have good performance in practice and induced many significant theoretical results such as termination technique [14], homogeneous and self-dual model [28] and etc.. On the one hand, they have the best worst-case iteration complexity obtained so far and induced significant results in theory; on the other hand, they are easy to implement and the most efficient from a computational point of view (see, e.g. Anderson et al. [1]).

Suppose we have $(x, s) \in \mathcal{F}_+$, which is an approximate solution of (CP) with $\mu := \frac{x^T s}{n}$. We want to decrease the complementary gap $x^T s$. So, we aim for another pair $(x^+, s^+) \in \mathcal{F}_+$ such that $\mu_+ := \frac{(x^+)^T s^+}{n}$ is smaller than μ and/or (x^+, s^+) is closer to the central path than (x, s). Let $\gamma \in [0, 1]$ such that $\mu_+ = \gamma\mu$. Now, applying Newton's method to the system of equations (CP) where μ is replaced by $\gamma\mu$, we obtain the following system (CPS) of linear equations:

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} d_x \\ d_y \\ d_s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -XSe + \gamma\mu e \end{bmatrix}.$$

The displacement (d_x, d_y, d_s) is a Newton step towards the point $(x_{\gamma\mu}, y_{\gamma\mu}, s_{\gamma\mu}) \in \mathcal{C}$, at which the pairwise products $x_j s_j$ are equal to $\gamma\mu$.

For $\gamma = 1$, the equations (CPS) define a *constant-gap centering direction*, a Newton step toward the point $(x_\mu, y_\mu, s_\mu) \in \mathcal{C}$, at which all the pairwise products $x_j s_j$ are identical to μ. Along this direction, the complementary gap is kept constant. At the other extreme, the value $\gamma = 0$ gives the standard Newton step, which is known as the *primal-dual affine-scaling direction* in this special setting.

Now, we introduce the definition of the *iteration complexity* of IPMs. At the beginning of the algorithm, we are given x^0, s^0 both strictly feasible in problems (P) and (D) respectively. Define $\mu_k := \frac{\langle x^k, s^k \rangle}{n}$, we need to find out the vectors x^k and s^k by an algorithm such that both vectors are feasible and for the desired accuracy $\epsilon \in (0, 1)$, $\mu_k \leq \epsilon\mu_0$. We call $(x^k,\, s^k)$ an *ϵ-solution* if the vectors satisfy the conditions referred to in the last sentence. We call the number of iterations to achieve an ϵ-solution the *iteration complexity.* In the practical algorithms, we start from an initial feasible point (x^0, s^0) and the complementary gap $(x^k)^T(s^k)$ is reduced at each iteration. We call the termination subroutine every 3 iterations after we reach an ϵ-solution, i.e., after the complementarity gap decreases from $(x^0)^T s^0$ to $\epsilon(x^0)^T s^0$. It is proven that we can obtain an exact optimal solution in polynomially many iterations by a proper termination technique if the initial data A, b, c are rational ([23], Chapter 4; also see Section 2.5 for detail). The termination technique we use in this thesis is in [16].

We first introduce the $predictor - corrector$ algorithm proposed by Mizuno ,Todd and Ye [16]. It is proved that we can obtain an ϵ-solution in $O(\sqrt{n}\ln\frac{1}{\epsilon})$ steps by the following algorithm.

Algorithm 2.1.
Input $(A, x^0, s^0, b, c, \epsilon, \eta)$, *where* A, x^0, s^0, b, c *are defined in the* (HLP) *formulation in Section 2.4,* ϵ *is the desired tolerance,* $\beta = \frac{1}{4}$.
Define k:=0;
while $x^T s > \epsilon$
    If k is even, execute predictor step as below,
        solve the linear system (CPS) with $\gamma = 0$, *then*
        let
            $x(\alpha) := x^k + \alpha d_x > 0,$
            $s(\alpha) := s^k + \alpha d_s > 0,$
            $\alpha^\star := \max\{\alpha : (x(\alpha), s(\alpha)) \in \mathcal{N}_2(2\beta)\}.$
        Let $x^{k+1} := x(\alpha^\star)$; $s^{k+1} := s(\alpha^\star)$;
    else
        execute the corrector step as below,
        solve the linear system (CPS) with $\gamma = 1$, *then*
        let
        $x^{k+1} := x^k + d_x,$
        $s^{k+1} := s^k + d_s,$
    end
    $k := k + 1.$
repeat

2.3 The Primal-Dual IPMs Based on Entropy

In the Ph.D thesis of Li [13], the author proposed several algorithms based on entropy. The thesis applied Shannon's (informational) entropy measure and Jaynes' maximum entropy criterion in the solving of constrained optimization problems.

In Li's thesis, a constrained optimization problem is divided into a number of subproblems which are analogous to the micro-states of a statistical thermodynamic system by introducing the idea of surrogate Lagrangean [10]. Surrogate multipliers represent probabilities of the system being in each micro-state. An optimizing process is then interpreted as the transition of the system to an equilibrium state. We know that entropy of thermodynamic system should attain a maximum value in stable state. Jayne's maximum entropy criterion is employed to formulate the surrogate multipliers, i.e., using an entropy maximization model to calculate least biased probabilities in thermodynamic system.

In summary, the method proposed in [13] updates x in the feasible region and the surrogate multiplier λ separately and alternatively. It does not belong to the IPMs we are currently using. However, the following result of [13] is meaningful in the exploration of new search directions in IPMs.

Lemma 2.1. *Let* $x > 0$*,* $s > 0$ *and* $\delta(x,s) := \sum_{j=1}^{n} \frac{x_j s_j}{n\mu} \ln\left(\frac{x_j s_j}{\mu}\right)$*. Then,*

1. [13, 25] $\delta(x,s) \geq 0$*;*
2. [25] *equality holds above if and only if* $Xs = \mu e$*;*
3. [13, 25] $\sum_{j=1}^{n} x_j s_j \ln\left(\frac{x_j s_j}{\mu \mathbf{e}^{\delta}}\right) = 0.$

Proof. The first inequality is proved in [25]. The third equality is trivial by definition. We prove the second part as below.

Define $u_j := \frac{x_j s_j}{\mu}$, $\forall j \in \{1, 2, \ldots, n\}$. Then, $\delta(u) = \frac{1}{n}\sum_{j=1}^{n} u_j \ln(u_j)$. We have,

$\nabla\delta(u) = \frac{1}{n}(e + \ln(u))$, $\nabla^2\delta(u) = \frac{1}{n}U^{-1} \succ 0$. Therefore, $\delta(u)$ is strictly convex.

We consider the constrained problem:

$$\begin{aligned} & \min \delta(u) \\ \text{s.t.} \quad & \sum_{j=1}^{n} u_j = n, \quad u > 0. \end{aligned}$$

We know the Lagrangean for this problem $\mathcal{L}(u,\lambda) = \delta(u) + \lambda(n - \sum_{j=1}^{n} u_j)$, and $\nabla^2_{uu}\mathcal{L}(u,\lambda) = \nabla^2\delta(u) \succ 0$. We found that $(u = e,\ \lambda = \frac{1}{n})$ satisfies the KKT conditions. Since $\nabla^2_{uu}\mathcal{L}(u,\lambda) \succ 0$, we know $u = e$ is the unique global minimizer for this problem by the property of strict convexity. Thus the minimum value of $\delta(u)$ is achieved only at $u = e$ where $\delta(x,s) = 0$. Due to the definition of u, we know that $\frac{Xs}{\mu} = u = e$, i.e., $Xs = \mu e$. □

In the paper of Tunçel and Todd [25], a new search direction is proposed by introducing the proximity measure $\delta(x,s)$ defined above (which measures how close (x,s) is to the central path) and $\delta(x,s)$ is closely related to the formulation of Primal-Dual Entropy. From now on, when x and s are clear from the context, we use δ to stand for $\delta(x,s)$.

We use the following notations.

In the k^{th} step, $x_j := x_j^k$, $s_j := s_j^k$,

$D := X^{\frac{1}{2}}S^{-\frac{1}{2}}$, we use D to scale the primal and dual space to the v-space where $v := X^{\frac{1}{2}}S^{\frac{1}{2}}e$. This will make the arguments easier. $w := -v + \delta v - V\ln(\frac{Vv}{\mu})$.

w_p is the orthogonal projection of w on the null space of matrix AD.

$w_q = w - w_p$, $d_x := Dw_p$, $d_s := D^{-1}w_q$.

We define: $x(\alpha) := x + \alpha d_x$, $s(\alpha) := s + \alpha d_s$.

We determine the step size by using $\alpha^\star := \max\{\alpha : (x(\alpha), s(\alpha)) \in \mathcal{N}_E(\beta)\}$. Then compute the next point by $x^{k+1} := x(\alpha^\star)$, $s^{k+1} := s(\alpha^\star)$, this process will continue until the convergence criteria is satisfied.

It is proved in [25] that in the neighborhood of $\mathcal{N}_E(3/2)$, the above algorithm will terminate in $O(n\ln\frac{1}{\epsilon})$ steps to get an optimal solution. This search direction has some good properties which we will explore later. We will use w to stand for $-v + \delta v - V\ln(\frac{Vv}{\mu})$ in the context.

We know that new search directions may be deduced via some alternative description of the central path. Recall the standard KKT conditions defining the central path:

$$
\text{(CP)} \qquad \begin{aligned} A^T y + s &= c, \quad s > 0 \\ Ax &= b, \quad x > 0 \\ Xs &= \mu e. \end{aligned}
$$

Assume that we are given a continuously differentiable strictly monotone function $\Psi(x) : \mathbb{R}^n_{++} \to$

$\mathbb{R}^n$ where the strict monotone property is componentwise. Any such function determines an IPM in a natural way. If we replace the last vector equation of (CP) by its equivalent: $\Psi(Xs) = \Psi(\mu e)$, then apply Newton's method to this new system in order to approximate a point on the central path corresponding to the parameter $(1-\theta)\mu$. Then a new search direction is defined.

If we replace the last vector equation of (CP) by its equivalent: $\ln(Xs) = \ln(\mu e)$, i.e., let $\Psi(v) = \ln(v)$, then apply the Newton method to this new system to approximate a point on the central path corresponding to the parameter μ_+. This direction (d_x, d_s) coincides with the direction derived from w if we choose μ_+ such that $\ln \frac{\mu}{\mu_+} = 1 - \delta(x, s)$.

In the paper by Zhang and Li [30], another search direction is proposed and discussed within the wide neighborhood $\mathcal{N}_\infty^-(\frac{1}{2})$. This search direction also arises from the technique of introducing algebraically equivalent central path with $\Psi(v) = \ln(v)$. It has the form:

$$-\sigma v + \delta v - V \ln(\tfrac{Vv}{\mu}), \text{ with } \sigma \in (0.5, 1), \text{ and } \sigma < \min\{1, \ln \tfrac{1}{1-\beta}\}.$$

It is shown that an infeasible-start Primal-Dual interior-point algorithm based on the above search direction can achieve iteration complexity of $O(n^2 \ln \frac{1}{\epsilon})$ in a wide neighborhood when applied to monotone LCPs [30].

There are several search directions proposed with different functions $\Psi(x)$, such as the direction of Jansen et al. [11] and Nazareth [18]. For instance, we can take $\Psi(x) = -X^{-1}e$ which will yield the search direction proposed in [11] (see [25]).

In the paper of Bai, El Ghami and Roos [2], the authors gave out a large class of search directions in a similar way. They choose strictly convex differentiable $\Psi(v)$ such that $\Psi(v)$ is minimized at $v = e$ and $\Psi(e) = 0$ and also based on the properties of up to the third order derivative of $\Psi(v)$. The authors summarized the properties for different $\Psi(x)$ and proposed some new $\Psi(x)$ which can induce search directions of good performance with respect to iteration complexity. Peng, Roos and Terlaky used similar methods to derive a search direction based on a called self-regular function $\Psi(x)$([20, 21]), while they use a direction that can be characterized as a steepest decent direction in a scale space instead of using the classical Newton direction. The prototype self-regular kernel function is given by:

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

where $p \geq 1$ and $q > 1$. The best theoretical iteration complexity for a large-update algorithm based on self-regular functions is $O(n^{\frac{q+1}{2q}} \ln(\frac{1}{\epsilon}))$. A recent paper [19] used the same search direction with [30] to solve LP problems and achieved polynomial convergence. The authors also discussed the relationship between reparameterization of the central path and the self-regular functions approach [21] by showing a reparameterization function $\Psi(x) = x^{\frac{q+1}{2}}$ can induce the search direction in [21]. There are still many open problems in this class of directions based on the large number of possible functions $\Psi(v)$.

2.4 Homogeneous and Self-Dual model

The algorithm proposed in [30] is one of the infeasible-start algorithms. The other algorithms mentioned in our last section all require that we know an initial strict feasible point in the set $\mathcal{F}_+$ at the very beginning. However, we know that it is not easier to find a strictly feasible point than to find the optimal solution for an LP problem. So we need some way to guarantee that we will have a strictly feasible point in $\mathcal{F}_+$ when the algorithm initializes.

We found that by using homogeneous and self-dual model of LP problem we can have some advantage in the aspect mentioned above.

Linear system $Ax = b$ is called *homogeneous* if $b = 0$.

It is a well known skill to attack a standard form LP by solving a related homogeneous artificial LP problem such as the formulation stated in next paragraph. It is easily seen that $x = 0$ will always be feasible for the artificial homogeneous problem. In our discussion, we allow a single non-homogeneous constraint, which is often called a *normalizing constraint*.

For the problems (P) and (D), the approximate solution of the following homogeneous Linear system yields an optimal solution of (P).

$$
\begin{array}{rrrlll}
 & Ax & -bt & = 0 & x \geq 0 & (2.1) \\
-A^T y & -s & +ct & = 0 & s \geq 0 & (2.2) \\
b^T y & -c^T x & & = 0 & & (2.3)
\end{array}
$$

If $t = 1$, then

(2.1) stands for the feasibility of (P),

(2.2) stands for the feasibility of (D),

(2.3) stands for the primal-dual zero gap attainment.

If the above homogeneous LP has a solution (x, y, t) with $t > 0$, then $\frac{x}{t}$ is a primal optimal solution, $(\frac{y}{t}, \frac{s}{t})$ is a dual optimal solution.

An LP problem is called self-dual problem if its dual is equivalent to itself.
E.g.:

$$\begin{array}{lll} \text{(SD)} & \text{minimize} & \tilde{c}^T u \\ & \text{subject to} & \tilde{A}^T u \geq \tilde{b}, \quad u \geq 0 \end{array}$$

where $\tilde{A} \in \mathbb{R}^{n \times n}$ is skew-symmetric, (i.e., $\tilde{A}^T = -\tilde{A}$), and $\tilde{b} = -\tilde{c} \in \mathbb{R}^n$. Then the problem (SD) is equivalent to its dual. If (SD) has a feasible solution $\tilde{u}$, then $\tilde{u}$ is also feasible in the dual problem, and the objective values sum to zero. Therefore, by the duality theory of linear programming, (SD) has an optimal solution and the optimal value is zero.

According to the methodology mentioned in [28], we can construct a homogeneous and self-dual artificial LP problem (HLP) related to (P) and (D) as below.
Given any $x^0 > 0$, $s^0 > 0$, and y^0 free,

$$\begin{array}{llrrrrl} & \min & & & & ((x^0)^T s^0 + 1)\theta & \\ (1) & \text{s.t.} & & Ax & -bt & +\overline{b}\theta & = 0 \\ (2) & & -A^T y & & +ct & -\overline{c}\theta & \geq 0 \\ (3) & & b^T y & -c^T x & & +\overline{z}\theta & \geq 0 \\ (4) & & -\overline{b}^T y & +\overline{c}^T x & -\overline{z}t & & = -((x^0)^T s^0 + 1) \\ (5) & & \text{y free}, & x \geq 0, & t \geq 0, & \theta \text{ free}, & \end{array}$$

where
$\overline{b} := b - Ax^0, \quad \overline{c} := c - A^T y^0 - s^0, \quad \overline{z} := c^T x^0 + 1 - b^T y^0.$

The relationships (1)-(3), with $t = 1$ and $\theta = 0$, represent primal and dual feasibility (with $x \geq 0$) and reversed weak duality, so they define primal and dual optimal solutions. To achieve feasibility for $x = x^0$ and $(y, s) = (y^0, s^0)$, the artificial variable θ is added with appropriate coefficients and the constraint (4) is added to achieve self duality.

Denote the slack vector for the inequality constraint (2) by s and by κ the slack scalar for the inequality constraint (3). We can see that (HLP) is homogeneous and self-dual.

The following are the properties of the (HLP) model [28].

- The Dual of (HLP), denoted by(HLD), has the same form as (HLP), i.e., (HLD) is simply (HLP) with (y, x, t, θ) being replaced by (y', x', t', θ'). Here y', x', t', θ' make up the dual multiplier vector for constraints (1), (2), (3), (4) respectively.
- (HLP) has a strictly feasible point for every choice of $x^0 > 0, s^0 > 0$:
  $y = y^0$, $x = x^0 > 0$, $t = 1$, $\theta = 1$, $s = s^0 > 0$, $\kappa = 1$.
- (HLP) has an optimal solution and its optimal solution set is bounded.
- The optimal value of (HLP) is zero, and for any feasible point, $(y, x, t, \theta, s, \kappa) \in F_h$, here F_h denotes the set of all points $(y, s, t, \theta, s, \kappa)$ that are feasible for (HLP):

$$((x^0)^T s^0 + 1)\theta = x^T s + t\kappa.$$

- There is an optimal solution $(y^\star, x^\star, t^\star, \theta^\star = 0, s^\star, \kappa^\star) \in F_h$, such that:

$$\begin{pmatrix} x^\star + s^\star \\ t^\star + \kappa^\star \end{pmatrix} > 0$$

  which we call a *strictly self-complementary solution.*

If we choose $y^0 := 0$, $x^0 := e$, and $s^0 := e$, then (HLP) becomes:

$$
\begin{array}{llrrrrl}
 & \min & & & & (n+1)\theta & \\
(1) & \text{s.t.} & & Ax & -bt & +\bar{b}\theta & = 0 \\
(2) & & -A^T y & & +ct & -\bar{c}\theta & \geq 0 \\
(3) & & b^T y & -c^T x & & +\bar{z}\theta & \geq 0 \\
(4) & & -\bar{b}^T y & +\bar{c}^T x & -\bar{z}t & & = -(n+1) \\
(5) & & x \geq 0, & t \geq 0, & & &
\end{array}
$$

where
$\bar{b} := b - Ae, \quad \bar{c} := c - e, \quad \bar{z} := c^T e + 1.$

If we look at the solution of (HLP), we can derive the optimal solution of (LP) by using the theorem below.

Theorem 2.2. [28] *Let* $(y^\star, x^\star, t^\star, \theta^\star = 0, s^\star, \kappa^\star)$ *be a strictly self-complementary solution for* (HLP)*. Then:*

- (P) *has a solution (neither infeasible nor unbounded) if and only if* $t^\star > 0$*. In this case,* $(x^\star/t^\star)$ *is an optimal solution for* (P) *and* $(y^\star/t^\star, s^\star/t^\star)$ *is an optimal solution for* (D)*;*

- *if* $t^\star = 0$*, then* $\kappa^\star > 0$*, which implies that* $c^T x^\star - b^T y^\star < 0$*, i.e., at least one of* $c^T x^\star$ *and* $-b^T y^\star$ *is strictly less than 0. If* $c^T x^\star < 0$ *then* (D) *is infeasible; if* $-b^T y^\star < 0$ *then* (P) *is infeasible; and if both* $c^T x^\star < 0$ *and* $-b^T y^\star < 0$ *then both* (P) *and* (D) *are infeasible.*

So homogeneous and self-dual model can guarantee that we have a strictly feasible solution when one interior point algorithm initializes.

2.5 Termination Technique for HSD Algorithms

For the Primal-Dual algorithms we will refer to in this thesis, we can always apply the termination technique described in [27]. Define φ be the index set $\{j : x_j^k \geq s_j^k, j = 1, 2, \ldots, n\}$,

denote by B those columns in A corresponding to φ and by N the rest of the columns in A. Then, we use a least-squares projection to create an optimal solution (y, x, t, κ) that is strictly self-complementary from an ϵ-solution $(y^k, x^k, s^k, \theta^k, t^k, \kappa^k)$ of (HLP).
Case 1. If $t^k \geq \kappa^k$, we solve for y, x_B, and t from

$$\begin{array}{lllll} \min & \|y^k - y\|^2 + & \|x_B^k - x_B\|^2 + & (t^k - t)^2 & \\ \text{s.t.} & & Bx_B & -bt & = 0 \\ & -B^T y & & +c_B t & = 0 \\ & b^T y & -c_B^T x_B & & = 0 \end{array}$$

otherwise,
Case 2. If $t^k < \kappa^k$, we solve for y, x_B, and κ from

$$\begin{array}{lllll} \min & \|y^k - y\|^2 + & \|x_B^k - x_B\|^2 + & (\kappa^k - \kappa)^2 & \\ \text{s.t.} & & Bx_B & & = 0 \\ & -B^T y & & & = 0 \\ & b^T y & -c_B^T x_B & -\kappa & = 0 \end{array}$$

This projection guarantees that the resulting $x_B^\star$ and $s_N^\star$ ($s_N^\star = c_N t^\star - N^T y^\star$ in Case 1 or $s_N^\star = -N^T y^\star$ in Case 2) are positive and $t^\star$ is positive in Case 1 and $\kappa^\star$ is positive in Case 2, as long as $(x^k)^T s^k + t^k \kappa^k$ is small enough in the algorithms. This is explained in detail in the paper of Ye [27] and Mehrotra [14]. It can be deduced from an interesting lemma as below.

Lemma 2.3. *[14] Given an interior solution x^k and s^k in the solution sequence generated by any primal-dual interior point algorithm, define $\varphi^k := \{j : x_j^k \geq s_j^k\}$. Then, we have:*
for some K sufficiently large but finite, $\varphi^k = \varphi$ for all $k \geq K$.

Remark: The statement "some K sufficiently large but finite" in the lemma above is equivalent to the statement that we have $(x^k)^T s^k + t^k \kappa^k$ less than some fixed small number which is independent of k [14].

In the paper of Ye [27], the performance of the algorithm after applying this termination technique is discussed. The termination technique is also seem to be very efficient using numerical tests [14].

Chapter 3

Some Fundamental Properties of the Entropic Search Directions

3.1 General Family of Entropic Directions

We have

$$w = -v + \delta v - V \ln\left(\frac{Vv}{\mu}\right)$$

in the v-space. This vector w completely determines the search direction for the algorithm in [25]. w as a vector in $\mathbb{R}^n$ is a non-negative linear combination of three vectors $-v$, v, $-V\ln(\frac{Vv}{\mu})$. One way to generalize the search direction w is to allow weights other than 1 for each of the three vectors above. Since $-v$ and v are clearly linearly dependent, we use two coefficients $\lambda_1 \in \mathbb{R}$, $\lambda_2 \in \mathbb{R}$ and consider

$$w(\lambda_1, \lambda_2) := -v + \lambda_1 v - \lambda_2 V \ln\frac{Vv}{\mu}.$$

Notice that any positive multiple of $w(\lambda_1, \lambda_2)$ yields the same primal-dual algorithm (multiplying $w(\lambda_1, \lambda_2)$ by k, divides step size α by k). Therefore, by a normalization of $w(\lambda_1, \lambda_2)$, we can reduce one of the parameters.
One obvious way to normalize is to enforce

$$\|w(\lambda_1, \lambda_2)\|_2 = 1.$$

Instead, we will normalize so that

$$x(\alpha)^T s(\alpha) = (1-\alpha)x^T s.$$

We can deduce the following relationship between λ_1 and λ_2 from this normalizing condition.

$$x^T d_s + s^T d_x = -n\mu = V^T w(\lambda_1, \lambda_2) \Leftrightarrow -n\mu = -n\mu + \lambda_1 n\mu - \lambda_2 n\mu\delta \Leftrightarrow \lambda_1 = \delta\lambda_2.$$

Therefore, we propose a class of search directions $w(\eta)$ based on the above statement. We use the following notations.
$w(\eta) := -v + \eta \left[\delta v - V \ln\left(\frac{Vv}{\mu}\right)\right]$, where $\eta \geq 0$.
$w(\eta)_p$ is the projection of $w(\eta)$ onto the null space of matrix AD.
$w(\eta)_q := w(\eta) - w(\eta)_p$, $d_x := Dw(\eta)_p$, $d_s := D^{-1}w(\eta)_q$.
Here d_x and d_s can also be achieved from the following system (HCPS):

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} d_x \\ d_y \\ d_s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ Vw(\eta) \end{bmatrix}.$$

Obviously the direction w is included in this family of search directions with $\eta = 1$. Moreover, the direction proposed in [30] is also included in this family with $\eta = \frac{1}{\sigma}$, i.e., $\max\left\{1, -\frac{1}{\ln(1-\beta)}\right\} < \eta \leq 2$.

3.2 Technical Lemmas for the Properties of Entropic Search Direction

To analyze the general family of entropic search directions, we need to understand the proximity measure δ well. We present some technical lemmas first which we will use in further analysis.

Lemma 3.1. *For every $\alpha \in \mathbb{R}$ such that $|\alpha| \leq 1$, we have :*

$$\alpha - \frac{\alpha^2}{2(1-|\alpha|)} \leq \ln(1+\alpha) \leq \alpha.$$

This lemma is well-known, one proof can be found in [24]. We will use it to estimate the value of δ in $\mathcal{N}_\infty(\beta)$ next.

Lemma 3.2. *Let $\beta \in [0,1)$ such that $(x,s) \in \mathcal{N}_\infty(\beta)$. Then,*

$$\frac{1-3\beta}{2(1-\beta)n}\left\|\frac{Xs}{\mu}-e\right\|_2^2 \le \delta \le \frac{1}{n}\left\|\frac{Xs}{\mu}-e\right\|_2^2.$$

Proof. The right-hand-side inequality was proved in [25]. We prove the left-hand-side inequality here. Let $\beta \in [0,1)$, such that $(x,s) \in \mathcal{N}_\infty(\beta)$. Then we have:

$$\begin{aligned}
\delta = \sum_{j=1}^{n} \frac{x_j s_j}{n\mu} \ln\left(\frac{x_j s_j}{\mu}\right) &\ge \sum_{j=1}^{n} \frac{x_j s_j}{n\mu}\left[\frac{x_j s_j}{\mu} - 1 - \frac{(\frac{x_j s_j}{\mu}-1)^2}{2(1-|\frac{x_j s_j}{\mu}-1|)}\right] \\
&\ge \frac{1}{n}\left\|\frac{Xs}{\mu}-e\right\|_2^2 - \max_{j=1,\ldots,n}\left(\frac{1}{2n(1-|\frac{x_j s_j}{\mu}-1|)}\right)\sum_{j=1}^{n}\frac{x_j s_j}{\mu}\left(\frac{x_j s_j}{\mu}-1\right)^2 \\
&\ge \frac{1}{n}\left\|\frac{Xs}{\mu}-e\right\|_2^2 - \frac{(1+\beta)}{2n(1-\beta)}\left\|\frac{Xs}{\mu}-e\right\|_2^2 \\
&= \frac{1-3\beta}{2(1-\beta)n}\left\|\frac{Xs}{\mu}-e\right\|_2^2.
\end{aligned}$$

In the above, the first inequality uses Lemma 3.1, the second inequality follows from the fact that $(x,s) \in \mathcal{N}_\infty(\beta)$. ☐

Corollary 3.3. *For every $(x,s) \in \mathcal{N}_\infty(\frac{1}{4})$, $\delta \ge \frac{1}{6n}\left\|\frac{Xs}{\mu}-e\right\|_2^2$. Moreover, for every $(x,s) \in \mathcal{N}_\infty(\frac{1}{10})$, $\delta \ge \frac{7}{18n}\left\|\frac{Xs}{\mu}-e\right\|_2^2$.*

We also observe that the lower bound of δ approaches $\frac{1}{2n}\left\|\frac{Xs}{\mu}-e\right\|_2^2$ as $\left\|\frac{Xs}{\mu}-e\right\|_\infty$ goes to 0.

After the analysis of the proximity measure δ, it is proper to review the search direction w in another aspect, i.e., we can decompose it into two orthogonal parts : affine-scaling and constant-gap centering.

Consider again the search direction $w = -v + \delta v - V\ln(\frac{Vv}{\mu})$. For $\delta < 1$, $-(1-\delta)v$ represents the affine-scaling component. It was stated in [25] that the other component of w behaves like the centering direction in terms of its first order behaviour. We already know the upper and lower bound of δ within $\mathcal{N}_\infty(\beta)$, so we can easily estimate the part of $-v + \delta v$. So we just need to estimate the part of $V \ln(\frac{Vv}{\mu})$ within $\mathcal{N}_\infty(\beta)$.

Lemma 3.4. *Let $\beta \in [0, \frac{1}{2})$. Then, for every $(x, s) \in \mathcal{N}_\infty(\beta)$, we have:*

$$\left(\delta - 2 - \frac{\beta^2}{4\beta^2 - 6\beta + 2}\right) v + \mu V^{-1} e \leq w \leq (\delta - 2)v + \mu V^{-1} e$$

Proof. If $\left\|\frac{Vv}{\mu} - e\right\|_\infty \leq \beta$, we know that $(1 - \beta)e \leq \frac{Vv}{\mu} \leq (1 + \beta)e$.

On the one hand, using Lemma 3.1, we have:

$$-V \ln\left(\frac{Vv}{\mu}\right) = V \ln(\mu V^{-2} e) = V \ln(e + \mu V^{-2} e - e) \leq V(\mu V^{-2} e - e) = \mu V^{-1} e - v,$$

on the other hand, using Lemma 3.1 again and the facts that $(x, s) \in \mathcal{N}_\infty(\beta)$, $\beta \in [0, \frac{1}{2})$, for every $i \in \{1, 2, \ldots, n\}$, we have:

$$\begin{aligned} v_i \ln \frac{\mu}{v_i^2} &\geq v_i \left[\frac{\mu}{v_i^2} - 1 - \frac{(\frac{\mu}{v_i^2} - 1)^2}{2(1 - |\frac{\mu}{v_i^2} - 1|)}\right] \\ &\geq v_i \left(\frac{\mu}{v_i^2} - 1\right) - \frac{v_i}{2(2 - \frac{1}{1-\beta})} \frac{\beta^2}{(1 - \beta)^2} \\ &\geq \mu v_i^{-1} - v_i - \frac{\beta^2 v_i}{4\beta^2 - 6\beta + 2}. \end{aligned}$$

Therefore, within $\mathcal{N}_\infty(\beta)$, for $\beta \in [0, \frac{1}{2})$, we can conclude this lemma holds.

□

Corollary 3.5. *For every $(x, s) \in \mathcal{N}_\infty(\frac{1}{4})$,*

$$(\delta - 2 - \frac{1}{12})v + \mu V^{-1} e \leq w \leq (\delta - 2)v + \mu V^{-1} e.$$

We are interested in the behavior of the search direction w which is proposed in [25] when it is applied to the predictor step of the P-C algorithm framework and the performance of w in the wide neighborhood. We found that the estimation of the following quantities plays an important role in the analysis. So we present the related results here.

We define $u_j := \frac{x_j s_j}{\mu}$, $\forall j \in \{1, 2, \dots, n\}$ and

$\Delta_{21}(u) := \sum_{j=1}^{n} u_j^2 \ln(u_j)$,

$\Delta_{12}(u) := \sum_{j=1}^{n} u_j \ln^2(u_j)$,

$\Delta_{22}(u) := \sum_{j=1}^{n} u_j^2 \ln^2(u_j)$.

We drop the argument u, (e.g. we write Δ_{ij} instead of $\Delta_{ij}(u)$) when u is clear from the context.

Lemma 3.6. *Let $\beta \in [0, \frac{1}{4}]$ and assume that $(x, s) \in \mathcal{N}_\infty(\beta)$. Then,*

$$\xi_{ij} n\delta \leq \Delta_{ij} \leq \zeta_{ij} n\delta, ij \in \{21, 22\}, \tag{3.1}$$

where

$\xi_{21} = 2(1-\beta)\ln(1-\beta) + 3(1-\beta)$,

$\zeta_{21} = 2(1+\beta)\ln(1+\beta) + 3(1+\beta)$,

$\xi_{22} = 6(1-\beta)\ln(1-\beta) + 2(1-\beta) + 6(1-\beta)\ln^2(1-\beta)$,

$\zeta_{22} = 6(1+\beta)\ln(1+\beta) + 2(1+\beta) + 6(1+\beta)\ln^2(1+\beta)$.

Let $\beta \in [0, \frac{1}{2}]$ and assume that $(x, s) \in \mathcal{N}_\infty^-(\beta)$. Then,

$$\xi_{12} n\delta \leq \Delta_{12} \leq \zeta_{12} n\delta,$$

where

$\xi_{12} = 0,\ \zeta_{12} = 2(\ln(n) + 1)$.

Proof. Let $f_{ij}(u) := \Delta_{ij} - \xi_{ij} n\delta$, $F_{ij}(u) := \zeta_{ij} n\delta - \Delta_{ij}$.

Define $\Omega := \{u \in \mathbb{R}^n : e^T u = n, \frac{3}{4}e \leq u \leq \frac{5}{4}e\}$, we define two classes of general constrained optimization problems:

$$\min_{u\in\Omega} F_{ij}(u) \qquad \text{and} \qquad \min_{u\in\Omega} f_{ij}(u).$$

$\nabla f_{21}(u) = (2U - \xi_{21} I)\ln(u) + u - \xi_{21} e$; $\nabla^2 f_{21}(u) = 2\text{Diag}(\ln(u)) - \xi_{21} U^{-1} + 3I$;
$\nabla F_{21}(u) = \zeta_{21} e + \zeta_{21} \ln u - 2U\ln(u) - u$; $\nabla^2 F_{21}(u) = \zeta_{21} U^{-1} - 3I - 2\text{Diag}(\ln(u))$;
$\nabla f_{22}(u) = -\xi_{22} e - \xi_{22} \ln u + 2U\ln^2(u) + 2U\ln u$;
$\nabla^2 f_{22}(u) = -\xi_{22} U^{-1} + 2I + 6\text{Diag}(\ln(u)) + 2\text{Diag}(\ln^2(u))$;
$\nabla F_{22}(u) = \zeta_{22} e + \zeta_{22} \ln u - 2U\ln^2(u) - 2U\ln u$;
$\nabla^2 F_{22}(u) = \zeta_{22} U^{-1} - 2I - 6\text{Diag}(\ln(u)) - 6\text{Diag}(\ln^2(u))$.

For $u > 0$, $2\ln(u_i) + 3 - \frac{\xi_{21}}{u_i}$ is an increasing function of u_i. Therefore, if $\nabla^2 f_{21}(u)$ is positive semidefinite at $u = (1-\beta)e$, then $\nabla^2 f_{21}(u)$ is positive semidefinite, $\forall u \in \Omega$. That is,
$2\ln(1-\beta) + 3 - \frac{\xi_{21}}{1-\beta} \geq 0 \Leftrightarrow \xi_{21} \leq 2(1-\beta)\ln(1-\beta) + 3(1-\beta)$.

For $u > 0$, $\frac{\zeta_{21}}{u_i} - 2\ln(u_i) - 3$ is a decreasing function of u_i. Whence, if $\nabla^2 F_{21}(u)$ is positive semidefinite at $u = (1+\beta)e$, then $\nabla^2 F_{21}(u)$ is positive semidefinite, $\forall u \in \Omega$. That is,
$-2\ln(1+\beta) - 3 + \frac{\zeta_{21}}{1+\beta} \geq 0 \Leftrightarrow \zeta_{21} \geq 2(1+\beta)\ln(1+\beta) + 3(1+\beta)$.

For $u > \mathbf{e}^{-3} e$, $\frac{\zeta_{22}}{u_i} - 6\ln(u_i) - 2 - 2\ln^2(u_i)$ is a decreasing function of u_i. Hence if $\nabla^2 F_{22}(u)$ is positive semidefinite at $u = (1+\beta)e$, then $\nabla^2 F_{22}(u)$ is positive semidefinite, $\forall u \in \Omega$. That is,
$-6\ln(1+\beta) - 2 - 2\ln^2(1+\beta) + \frac{\zeta_{22}}{1+\beta} \geq 0 \Leftrightarrow \zeta_{22} \geq 6(1+\beta)\ln(1+\beta) + 2(1+\beta) + 6(1+\beta)\ln^2(1+\beta)$.

Similarly, for $u > \mathbf{e}^{-3} e$, $-\frac{\xi_{22}}{u_i} + 6\ln(u_i) + 2 + 2\ln^2(u_i)$ is an increasing function of u_i. Therefore, if $\nabla^2 f_{22}(u)$ is positive semidefinite at $u = (1-\beta)e$, then $\nabla^2 f_{22}(u)$ is positive semidefinite, $\forall u \in \Omega$. That is,
$6\ln(1-\beta) + 2 + 2\ln^2(1-\beta) - \frac{\xi_{22}}{1-\beta} \geq 0 \Leftrightarrow \xi_{22} \leq 6(1-\beta)\ln(1-\beta) + 2(1-\beta) + 6(1-\beta)\ln^2(1-\beta)$.

If we consider the constrained optimization problems defined before and ξ_{21}, ζ_{21}, ξ_{22} and ζ_{22} bounded as above. The Lagrangeans have the forms below,
$\underline{\mathcal{L}}_{21}(u, \lambda) = f_{21}(u) - \lambda_1(n - e^T u) - \lambda_2^T(u - \frac{3}{4}e) - \lambda_3^T(\frac{5}{4}e - u)$,
$\overline{\mathcal{L}}_{21}(u, \lambda) = F_{21}(u) + \lambda_1(n - e^T u) + \lambda_2^T(u - \frac{3}{4}e) + \lambda_3^T(\frac{5}{4}e - u)$,
$\underline{\mathcal{L}}_{22}(u, \lambda) = f_{22}(u) - \lambda_1(n - e^T u) - \lambda_2^T(u - \frac{3}{4}e) - \lambda_3^T(\frac{5}{4}e - u)$,
$\overline{\mathcal{L}}_{22}(u, \lambda) = F_{22}(u) + \lambda_1(n - e^T u) + \lambda_2^T(u - \frac{3}{4}e) + \lambda_3^T(\frac{5}{4}e - u)$,
where $\lambda_1 \in \mathbb{R}$, $\lambda_2 \in \mathbb{R}^n_+$, $\lambda_3 \in \mathbb{R}^n_+$. We define $\lambda := [\lambda_1; \lambda_2; \lambda_3] \in \mathbb{R}^{2n+1}$.
We found that for $u^\star = e$, there are multiplier vectors $\underline{\lambda}_{21} = (\xi_{21} - 1)e_1$, $\overline{\lambda}_{21} = (\zeta_{21} - 1)e_1$, $\underline{\lambda}_{22} = \xi_{22} e_1$ and $\overline{\lambda}_{22} = \zeta_{22} e_1$ such the KKT conditions are satisfied respectively. Here e_i is the i^{th} unit vector whose dimension will be clear from the context (in this case $2n+1$). We also know $\nabla^2_{uu}\underline{\mathcal{L}}_{21}(u^\star, \underline{\lambda}_{21}) = \nabla^2 f_{21}(u)$, $\nabla^2_{uu}\overline{\mathcal{L}}_{21}(u^\star, \overline{\lambda_{21}}) = \nabla^2 F_{21}(u)$, $\nabla^2_{uu}\underline{\mathcal{L}}_{22}(u^\star, \underline{\lambda}_{22}) = \nabla^2 f_{22}(u)$, $\nabla^2_{uu}\overline{\mathcal{L}}_{22}(u^\star, \overline{\lambda_{22}}) = \nabla^2 F_{22}(u)$ are positive semidefinite respectively. Therefore, using the second order sufficient condition for a global minimizer, u^* is the global minimizer of these problems

with objective values all equal to 0. We can conclude $f_{21}(u^*) \geq 0$, $F_{21}(u^*) \geq 0$, $f_{22}(u^*) \geq 0$ and $F_{22}(u^*) \geq 0$ which imply the estimation of ξ_{21}, ζ_{21}, ξ_{21} and ζ_{22}.

The conclusion that $\xi_{12} = 0$ obviously holds due to the nonnegativity of the vectors x, s, u and $\text{Diag}(\ln(u))\ln(u)$.

We would like to discuss the estimation of ζ_{12} within $\mathcal{N}_\infty^-(\frac{1}{2})$ next.
For $\zeta_{12} = 2(\ln(n)+1)$, $\nabla F_{12}(u) = 2(\ln(n)+1)e + 2(\ln(n)+1)\ln(u) - \text{Diag}(\ln(u))\ln(u) - 2\ln(u)$, $\nabla^2 F_{12}(u) = 2\ln(n)U^{-1} - 2\text{Diag}(\ln(u))U^{-1}$.
If we consider the constrained optimization problem

$$\min_{u\in\mathbb{R}^n} F_{12}(u) \text{ subject to } e^T u - n = 0,\ u - \tfrac{1}{2}e \geq 0,$$

the Lagrangean has the form $\mathcal{L}_{12}^*(u,\lambda) = F_{12}(u) - \lambda_1(e^T u - n) - \lambda_2^T(u - \frac{1}{2}e)$.
It is obvious that $\nabla^2 F_{12}$ is positive definite if $u < ne$ while we know that $u \leq \frac{n+1}{2}e$ within $\mathcal{N}_\infty^-(\frac{1}{2})$. Hence we can conclude that F_{12} is strictly convex here. Moreover, we found for $u^\star = e$, there is a multiplier vector $\lambda_{12}^* = \zeta_{12}e_1$, such that the KKT conditions are satisfied. Therefore, $u^\star$ is the global minimizer of the optimization problem. We notice that $F_{12}(u^\star) = 0$ which implies the conclusion. □

Corollary 3.7. *For every $(x,s) \in \mathcal{N}_\infty(\frac{1}{4})$, $u = \frac{Xs}{\mu}$ as defined before, we have*

$$0 \leq 1.8n\mu^2\delta \leq \Delta_{21}(u) \leq \frac{9}{2}n\mu^2\delta,$$

$$\Delta_{22}(u)\mu^2 < 5n\delta.$$

Lemma 3.8. *Let $x > 0$, $s > 0$. Then $\Delta_{12} \geq n\delta^2$. Moreover, equality holds if and only if $Xs = \mu e$.*

Proof. We already know that $\Delta_{12} = \sum_{j=1}^n u_j \ln^2(u_j)$ and $\delta = \frac{1}{n}\sum_{j=1}^n u_j \ln(u_j)$, hence we only need to prove the following inequality in terms of u_j:

$$\sum_{j=1}^n u_j \ln^2(u_j) \geq \frac{1}{n}\left(\sum_{j=1}^n u_j \ln(u_j)\right)^2 \Leftrightarrow \left(\sum_{j=1}^n u_j\right)\sum_{j=1}^n u_j \ln^2(u_j) \geq \left(\sum_{j=1}^n u_j \ln(u_j)\right)^2.$$

Since $u_j \geq 0$, so $\sqrt{u_j} \geq 0$ and $\sqrt{u_j}|\ln(u_j)| \geq 0$, according to Cauchy-Schwartz inequality, we have

$$\left(\sum_{j=1}^{n} u_j\right) \sum_{j=1}^{n} u_j \ln^2(u_j) \geq \left(\sum_{j=1}^{n} u_j |\ln(u_j)|\right)^2 \geq \left(\sum_{j=1}^{n} u_j \ln(u_j)\right)^2$$

and the first equality holds if and only if $(U)^{\frac{1}{2}}e$ and $U^{\frac{1}{2}}|\ln(u)|$ are linearly dependent, that is, $|\ln(u_j)| = \mathit{constant}$ for every j. We know u_j can only be in the form of $\{c, \frac{1}{c}\}$ where $\frac{1}{n} \leq c \leq 1$. The second inequality holds if and only if $\sum_{j=1}^{n} u_j \ln(u_j) = \sum_{j=1}^{n} u_j |\ln(u_j)|$, i.e., $\sum_{u_j<1} u_j \ln(u_j) = 0$, this holds if and only if the set $\{j : u_j < 1\}$ is empty, which means $u = e$. The first equality will hold when we set $c = 1$ which also means $u \geq e$, but $\sum_{j=1}^{n} u_j = n$. Therefore, the conclusion holds if and only if $\frac{Xs}{\mu} = e$. □

3.3 Performance of Entropic Search Direction in Predictor-Corrector Algorithm

We are interested in the performance of the search direction w when it is applied in the predictor step of the Predictor-Corrector Algorithm. So we propose the following algorithm.

Algorithm 3.1.
Input $(A, x^0, s^0, b, c, \epsilon)$, *where* (x^0, s^0) *is strictly feasible primal-dual initial point,*
ϵ *is the desired tolerance,* $\beta = \frac{1}{4}$.
define k:=0;
while $x^T s > \epsilon$

if k is even, then we execute the predictor step,

calculate $\delta, D, v, w, w_p, w_q$ *according to the notations stated in Section 2.3;*

solve the system (HCPS) to get the unique solution d_x *and* d_s*,*

let

$x(\alpha) := x^k + \alpha d_x > 0,$

$s(\alpha) := s^k + \alpha d_s > 0,$

$\alpha^\star := \max\{\alpha : (x(\alpha), s(\alpha)) \in \mathcal{N}_2(2\beta)\}.$

Let $x^{k+1} := x(\alpha^\star),\ s^{k+1} := s(\alpha^\star)$.

else

execute the corrector step as below,

solve the linear system (CPS) with $\gamma = 1$ *to get* d_x, d_s, *then*

let

$x^{k+1} := x^k + d_x,$

$s^{k+1} := s^k + d_s.$

end

$k := k + 1.$

repeat

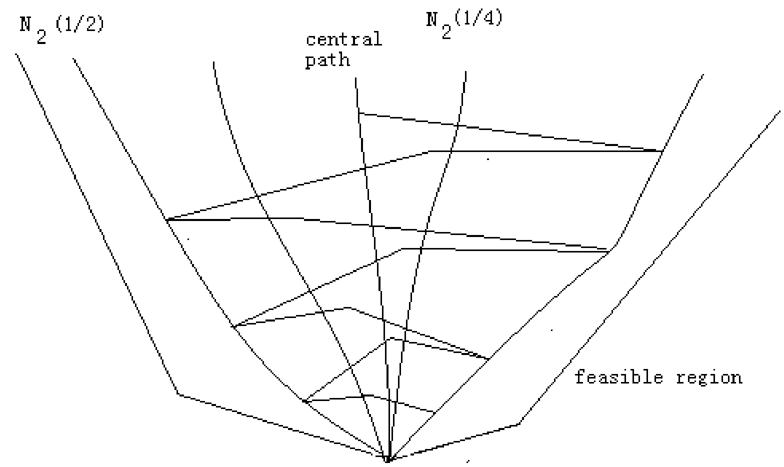


Figure 3.1: sketch of predictor-corrector algorithm

Lemma 3.9. *For the Primal-Dual path-following methods based on the search directions belong to the entropic family proposed in Section 3.1, we have the following relationship between step length and iteration complexity. That is, if the step length for each iteration is* $\Omega(\frac{1}{f(n)})$ *where* $f : \mathbb{N}_+ \to \mathbb{N}_+$, *then the algorithm needs at most* $O(f(n)\ln(\frac{1}{\epsilon}))$ *steps to achieve an* ϵ*-solution.*

Proof. We know that there is a good property for the entropic search direction family we proposed, i.e.,

$$x(\alpha)^T s(\alpha) = (1-\alpha)x^T s.$$

If the step length α_i for each iteration is $\Omega(\frac{1}{f(n)})$, then $\alpha_i > \frac{c}{f(n)}$ where c is a positive constant. From Lemma 3.1, we have,

$$(x^k)^T(s^k) = \prod_{i=1}^{k}(1-\alpha^i)(x^0)^T(s^0) \leq \mathbf{e}^{-\frac{kc}{f(n)}}(x^0)^T(s^0).$$

We need $\lceil \frac{f(n)}{c}\ln(\frac{1}{\epsilon})\rceil$ steps to decrease $(x^k)^T(s^k)$ from $(x^0)^T(s^0)$ to $\epsilon(x^0)^T(s^0)$, i.e., the iteration complexity is $O(f(n)\ln(\frac{1}{\epsilon}))$. □

For example: For the algorithm mention in [25], the step length for each iteration is at least $\Omega(\frac{1}{n})$, and the iteration complexity is $O(n\ln(\frac{1}{\epsilon}))$.

Theorem 3.10. *The iteration complexity of Algorithm 3.1 is* $O\left(\sqrt{n}\ln\frac{1}{\epsilon}\right)$.

Proof. If we apply the direction w to predictor step in P-C algorithm, then we need to find the lower bound for the step length to decide the iteration complexity.
For $(x,s)\in\mathcal{N}_2(\frac{1}{4})$, the following formula can guarantee that $(x(\alpha),s(\alpha))\in\mathcal{N}_2(\frac{1}{2})$.

$$\sum_{j=1}^{n}\left[\frac{x_j(\alpha)s_j(\alpha)}{(1-\alpha)\mu}-1\right]^2\leq\sum_{i=1}^{n}\left(\frac{x_js_j}{\mu}-1\right)^2+\frac{3}{16}. \tag{3.2}$$

The left-hand-side can be expanded as below:

$$\sum_{j=1}^{n}\left(\frac{x_j(\alpha)s_j(\alpha)}{(1-\alpha)\mu}-1\right)^2$$
$$=\sum_{j=1}^{n}\left(u_j-1+\frac{\alpha u}{(1-\alpha)}(\delta-\ln u_j)+\frac{\alpha^2}{(1-\alpha)\mu}(w_p)_j(w_q)_j\right)^2$$
$$=\sum_{j=1}^{n}\Bigg[(u_j-1)^2+\frac{\alpha^2(u_j)^2}{(1-\alpha)^2}(\delta^2+\ln^2u_j-2\delta\ln u_j)$$
$$+\tfrac{\alpha^4}{(1-\alpha)^2\mu^2}(w_p)_j^2(w_q)_j^2+2(u_j-1)(\tfrac{\alpha u_j}{(1-\alpha)}(\delta-\ln u_j)+\tfrac{\alpha^2}{(1-\alpha)\mu}(w_p)_j(w_q)_j)$$
$$+2\tfrac{\alpha^3\mu u_j}{(1-\alpha)^2\mu^2}(\delta-\ln u_j)(w_p)_j(w_q)_j\Bigg].$$

Further simplification results in the following formulation:

$$\tfrac{\alpha}{1-\alpha}\left(\delta^2\sum_{j=1}^{n}(\mu u_j)^2+\Delta_{22}-2\delta\Delta_{21}\right)+\tfrac{\alpha^3}{1-\alpha}\sum_{j=1}^{n}(w_p)_j^2(w_q)_j^2$$
$$+2\sum_{j=1}^{n}\left[u_j^2\mu^2\delta-\mu^2\delta u_j-u_j^2\mu^2\ln(u_j)+\mu^2u_j\ln u_j\right]$$
$$+2\alpha\sum_{j=1}^{n}\mu u_j(w_p)_j(w_q)_j+2\frac{\alpha^2}{1-\alpha}\sum_{j=1}^{n}\delta\mu u_j(w_p)_j(w_q)_j$$
$$-2\tfrac{\alpha^2}{1-\alpha}\sum_{j=1}^{n}\mu u_j(w_p)_j(w_q)_j\ln(u_j)\leq\frac{3(1-\alpha)\mu^2}{16\alpha}.$$

After expansion of the terms in the inequality (3.2) as shown above, we get an inequality providing a sufficient condition for the predictor step length α.
We have:
$d_4\alpha^4+d_3\alpha^3+d_2\alpha^2+d_1\alpha+d_0\leq 0,$
where
$B:=\sum_{j=1}^{n}\mu u_j\ln u_j(w_p)_j(w_q)_j,$

$C := \sum_{j=1}^{n} \mu u_j (w_p)_j (w_q)_j,$
$d_0 := -3\mu^2 \leq 0,$
$d_1 := 32\delta\mu^2\beta^2 + 32n\delta\mu^2 - 32\Delta_{21} + 6\mu^2 \leq 2\delta\mu^2 + 2\mu^2 + 6\mu^2 \leq 9\mu^2.$
The estimation for d_1 is based on Corollary 3.3.
$d_2 := 16(n\mu^2\delta^2 + \Delta_{22} + \delta^2\mu^2\beta^2 - 2\delta\Delta_{21} + 2C) - d_1 + 3\mu^2$,
$d_3 := 32(\delta - 1)C - 32B,$
$d_4 := 16\sum_{j=1}^{n} (w_p)_j^2 (w_q)_j^2.$
Now we should analyze the properties of these coefficients.
Since $v_j = x_j^{\frac{1}{2}} s_j^{\frac{1}{2}}$ and we know that within $\mathcal{N}_2(\frac{1}{4})$,
$\frac{3}{4} < |\frac{v_j^2}{\mu}| < \frac{5}{4}$, $0 \leq \delta \leq 1$ and more precisely, $\frac{1}{96n} \leq \frac{\beta^2}{6n} \leq \delta \leq \frac{\beta^2}{n} \leq \frac{1}{16n}$ on the boundary of $\mathcal{N}_2(\frac{1}{4})$.
According to the result in [15], we know that:
$\|W_p w_q\| \leq \frac{\sqrt{2}}{4}\|r\|^2$, i.e., $\sqrt{\sum_{j=1}^{n} (w_p)_j^2 (w_q)_j^2} \leq \frac{\sqrt{2}}{4}\sum_{j=1}^{n} r_j^2$
where $r_j = -v_j + \delta v_j - v_j \ln\left(\frac{v_j^2}{\mu}\right)$.
$|r_j^2| \leq v_j^2 \left|-1 + \delta - \ln\left(\frac{v_j^2}{\mu}\right)\right|^2 \leq \frac{25}{16} v_j^2.$

The second inequality of above is due to the facts that $\left|\ln\left(\frac{v_j^2}{\mu}\right)\right| \leq \frac{1}{4}$ and $0 \leq \delta \leq \frac{1}{16n}$ within $\mathcal{N}_2(\frac{1}{4})$.

$$d_4 = 16\sum_{j=1}^{n} (w_p)_j^2 (w_q)_j^2 \leq 2\left(\sum_{j=1}^{n} r_j^2\right)^2 \leq \frac{625}{128}\left(\sum_{j=1}^{n} \mu u_j\right)^2 \leq \frac{625}{128} n^2\mu^2 \leq 5n^2\mu^2.$$

Within $\mathcal{N}_2(\frac{1}{4})$, $\ln(u_j) \leq 1$, using Cauchy-Schwartz inequality we have:

$$\begin{aligned}
|B| &= \left|\sum_{j=1}^{n} \mu u_j \ln(u_j)(w_p)_j(w_q)_j\right| \leq \sum_{j=1}^{n} \mu u_j \left|\ln(u_j)(w_p)_j(w_q)_j\right| \\
&\leq \sum_{j=1}^{n} \mu u_j |(w_p)_j(w_q)_j| \leq \sqrt{\sum_{j=1}^{n} (w_p)_j^2 (w_q)_j^2}\sqrt{\sum_{j=1}^{n} \mu^2 u_j^2} \\
&\leq \frac{3n\mu}{4}\sqrt{(n+\beta^2)\mu^2} \leq \frac{3\sqrt{2}}{4} n^{\frac{3}{2}}\mu^2.
\end{aligned}$$

Similarly,

$$|C| \le \sum_{j=1}^{n} \mu u_j |(w_p)_j (w_q)_j| \le \frac{3\sqrt{2}}{4} n^{\frac{3}{2}} \mu^2;$$

Moreover,

$$C \le \sum_{\{i:(w_p)_j (w_q)_j \ge 0\}} \mu u_j (w_p)_j (w_q)_j \le \sum_{\{i:(w_p)_j (w_q)_j \ge 0\}} \frac{\mu^2 u_j^2}{4} \le \frac{1}{4}(n+\beta^2)\mu^2 \le \frac{n\mu^2}{2}.$$

Since $|\delta - 1| < 1$, we get:

$$d_3 = 32(\delta-1)C - 32B < 32|C| + 32|B| \le 96 n^{\frac{3}{2}} \mu^2.$$

Using Lemma 3.6 and Corollary 3.7, for every $(x,s) \in \mathcal{N}_2(\frac{1}{4})$,
we have: $\Delta_{22} \le \frac{5}{16}\mu^2$ and $\Delta_{21} \le \frac{9}{32}\mu^2$. So:

$$\begin{aligned}
d_2 &= 16(n\mu^2\delta^2 + \Delta_{22} + 2\Delta_{21} - 2n\mu^2\delta + \delta^2\mu^2\beta^2 - 2\delta\mu^2\beta^2 - 2\delta\Delta_{21} + 2C) - 3\mu^2 \\
&\le 16(n\mu^2\delta^2 + \frac{5}{16}\mu^2 + \frac{9}{16}\mu^2 + 0 + \delta^2\mu^2\beta^2 + 0 + 0 + 2C) - 3\mu^2 \\
&\le 16(n\beta\mu^2 + \mu^2 + n\mu^2) - 3\mu^2 \\
&\le 32n\mu^2.
\end{aligned}$$

After relaxation of the inequality, we get:

$$5n^2\alpha^4 + 96n^{\frac{3}{2}}\alpha^3 + 32n\alpha^2 + 9\alpha \le 3. \tag{3.3}$$

Suppose α has the form $\alpha = \frac{1}{E\sqrt{n}}$, we can see that if $E = 60$, then (3.3) is satisfied and so does (3.2).
We get the result that the step length α of predictor step is at least $\frac{1}{60\sqrt{n}}$. For the corrector step, the search direction is the constant-gap direction defined before. We know from the paper [28] that the complementary gap does not change in this step while the iterate will return to $\mathcal{N}_2(\frac{1}{4})$. Since we apply Algorithm 3.1 in the context of HSD model to get the starting point (x^0, s^0), we have the following inequality by using similar analysis in Lemma 3.9,

$$(x^k)^T(s^k) = \prod_{i=1, i \text{ is even}}^{k} (1-\alpha^i)(x^0)^T(s^0) \le \mathbf{e}^{-\frac{k}{120\sqrt{n}}}(x^0)^T(s^0).$$

We conclude that the iteration complexity for Algorithm 3.1 is $O\left(\sqrt{n}\ln\frac{1}{\epsilon}\right)$.

□

3.4 Performance of Entropic Search Direction in Monotone LCP Problem

The monotone linear complementarity problem [4], denoted by LCP, is to find $x \in \mathbb{R}^n$, such that $x \ge 0$, $Mx + q \ge 0$, $x^T(Mx+q) = 0$, where $M \in \mathbb{R}^{n\times n}$ is a positive semidefinite matrix, $q \in \mathbb{R}^n$ is a vector. Primal-Dual path-following IPM is effective in solving LCPs [26].

By using the technique of constructing an auxiliary sequence [31] and embedding the Monotone LCP problem into the HSD model, we find that we get some improvement on the iteration complexity of Primal-Dual path-following algorithm which uses the entropic search direction of [30] within $N_\infty^-(\beta)$. This search direction is derived by introducing algebraically equivalent central path with $\Psi(v) = \ln(v)$. It has the form:

$$-\sigma v + \delta v - V\ln(\tfrac{Vv}{\mu}), \text{ with } \sigma \in (0.5, 1), \text{ and } \sigma < \min\{1, \ln\tfrac{1}{1-\beta}\}.$$

The proof will use an auxiliary sequence (u^j, v^j) constructed for analysis purpose while it is not necessary to be calculated in practical algorithm.

Define: $(u^0, v^0) \in \{(x, y) \in \mathbb{R}^{2n} : Mx + h = y\}$, choose $(x^0, y^0) > 0$ and $(x^0, y^0) \ge (u^0, v^0)$. The auxiliary sequence (u^j, v^j) is defined as below:
$u^{k+1} := u^k + \alpha^k(d_{x^k} + x^k - u^k)$, $\quad v^{k+1} := v^k + \alpha^k(d_{y^k} + y^k - v^k)$.

In the paper [30], the authors constructed sequence (u^j, v^j) in a similar manner with that in [31]. We use different sequence construction within HSD model and get different result. We'd like to list some lemmas independent with the construction of (u^j, v^j) first.

The iteration complexity analysis is based on the potential reduction method and the potential function is :

$$\phi(x,s) := \frac{1}{\mu}\sum_{j=1}^{n} x_j s_j \ln(x_j s_j).$$

Lemma 3.11. *[31]*
$\|D^k d_{x^k}\|_2^2 + \|(D^k)^{-1} d_{s^k}\|_2^2 + 2(d_{x^k})^T d_{s^k} = \|(x^k s^k)^{\frac{1}{2}} \ln(x^k s^k/\mu_k \exp(\delta - \sigma))\|_2^2$,
where $D^k = (X^k)^{-\frac{1}{2}}(S^k)^{\frac{1}{2}}$.

The goals of this lemma are to estimate the value of $2(d_{x^k})^T d_{s^k}$ and to facilitate the definition of ς_k later.

Define:

$$\nu_k = \left(\frac{n}{1-\beta}\right)^{\frac{1}{2}} \frac{(x^k - u^k)^T s^k + (s^k - v^k)^T x^k}{(x^k)^T (s^k)},$$

$$\varsigma_k = \sigma^2 + \delta \ln n - \delta^2 + 2\beta_k \frac{(x^0 - u^0)(s^0 - v^0)}{(x^0)^T s^0},$$

$$\omega_k = \left(\nu_k + \sqrt{\nu_k^2 + \varsigma_k^2}\right)^2.$$

Lemma 3.12. *[31]* $\|D^k d_{x^k}\|_2^2 + \|(D^k)^{-1} d_{s^k}\|_2^2 \leq \omega_k (x^k)^T s^k$ *and* $\{\omega_k\}$ *is bounded.*

Since the sequence is proved to be global Q-linear convergent ([30], Lemma 7), i.e., the value for the potential function in iteration $k+1$ satisfies $\phi_{k+1} = \phi_0(1-\varrho)^k$. The above definitions will help us to identify the order of ω which is closely related to ϱ. It is essential to estimate the value of ϱ to determine the iteration complexity.

Lemma 3.13. *([30], Lemma 6)*
For the step length α*,*

$$\alpha_k \geq \min\left(\frac{\exp(-\sigma)}{1-\beta}, \frac{\min((1-\beta)\sigma, 2-2\sigma)}{n\omega}\right).$$

For the factor ϱ *which is related to the reduction of potential function:*

$$\varrho(\alpha_k) \geq \min\left(\frac{\exp(-\sigma)}{1-\beta}, \frac{\min((1-\beta)\sigma, 2-2\sigma)}{n\omega}\right)\left(\sigma - \frac{\min((1-\beta)\sigma, 2-2\sigma)}{n\omega}\right).$$

The above lemma gives a relationship between ω and ϱ.

Theorem 3.14. *If we apply search direction* $w_1 = -\sigma v + \delta v - V\ln(\frac{Vv}{\mu})$ *to Monotone LCP problem in the wide neighborhood* $\mathcal{N}_\infty^-(\beta)$ *where* $\sigma \in (0.5, 1)$, *and* $\sigma < \min\{1, \ln\frac{1}{1-\beta}\}$, *We can achieve iteration complexity of* $O(n\ln^4 n\ln(\frac{1}{\epsilon}))$.

Proof. If we embed this problem in a HSD model and assume $(u^0, v^0) = e$, $(x^0, y^0) = e$ where e is the all one vector of dimension $2n$, we get a strictly feasible initial point at the beginning. Then we know $\nu_k = 0$, $\varsigma_k = \sigma^2 + \delta\ln n - \delta^2$, where $0 \le \delta < \frac{n}{2}\ln\frac{n}{2}$. Since $0 < \delta < \ln n$, in the worst case, $\varsigma^k \le \ln^2 n + 1$, $\omega_k \le (\ln^2 n + 1)^2$, so $\varrho_k = O(\frac{1}{n\ln^4 n})$, the iteration complexity is at most $O(n\ln^4 n\ln(\frac{1}{\epsilon}))$. □

In [30], the author proposed an infeasible-start algorithm based on one entropic search direction and proved the iteration complexity of $O(n^2\ln(\frac{1}{\epsilon}))$. We used the same technique and proved the iteration complexity of $O(n\ln^4 n\ln(\frac{1}{\epsilon}))$ for this search direction because that we employ the properties of homogeneous and self-dual model to get a feasible starting point. In the next chapter, we use some other technique to analyze the iteration complexity of this entropic search direction for LP problems in HSD model. The result is $O(n\ln n\ln(\frac{1}{\epsilon}))$.

Chapter 4

Analysis on the General Family of Search Directions Based on Primal-Dual Entropy

We proposed a general family of search directions based on Primal-Dual Entropy in the last chapter. We'd like to analyze the iteration complexity of Primal-Dual path-following algorithm based on this family here.

Lemma 4.1. *Let $x > 0$, $s > 0$. For $\eta \geq 0$, $\|w(\eta)\|_2^2 = n\mu[1 - \eta^2(\delta^2 - \frac{\Delta_{12}}{n})]$.*

Proof. Let $x > 0$, $s > 0$. We have,

$$\begin{aligned}
\|w(\eta)\|_2^2 &= \sum_{j=1}^{n} v_j^2 \left(\delta\eta - 1 - \eta \ln \frac{x_j s_j}{\mu}\right)^2 \\
&= \sum_{j=1}^{n} x_j s_j \left(\delta^2\eta^2 + 1 + \eta^2 \ln^2\left(\frac{x_j s_j}{\mu}\right) + 2\eta \ln\left(\frac{x_j s_j}{\mu}\right) - 2\delta\eta - 2\delta\eta^2 \ln\left(\frac{x_j s_j}{\mu}\right)\right) \\
&= n\mu\delta^2\eta^2 + n\mu + \eta^2\Delta_{12} + 2n\eta\mu\delta - 2\delta\eta n\mu - 2n\mu\delta\eta^2\delta \\
&= n\mu + \eta^2\Delta_{12} - n\mu\eta^2\delta^2 \\
&= n\mu\left[1 - \eta^2\left(\delta^2 - \frac{\Delta_{12}}{n}\right)\right].
\end{aligned}$$

The third equality is due to the definition of δ and Δ_{12}, the other equalities are trivial. □

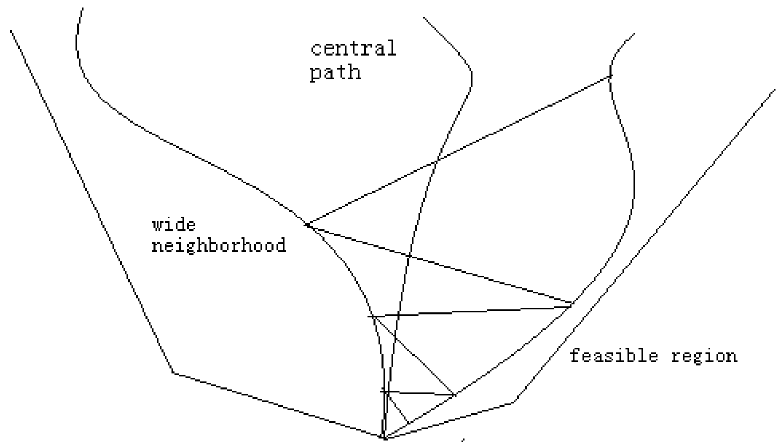


Figure 4.1: Sketch for Primal-Dual path-following method

Algorithm 4.1.

Input $(A, x^0, s^0, b, c, \epsilon, \eta)$, *where* A, x^0, s^0, b, c *are defined in the* (HLP) *formulation in Chapter 2,* ϵ *is the desired tolerance,* $1 \leq \eta = O(1)$, $\beta := \frac{1}{2}$.

while $x^T s > \epsilon$

calculate $\delta, D, v, w(\eta), w(\eta)_p, w(\eta)_q$ *according to the notations above;*

solve the system (HCPS) to get the unique solution d_x *and* d_s,

let

$x(\alpha) := x + \alpha d_x > 0$,

$s(\alpha) := s + \alpha d_s > 0$,

$\alpha^\star := \max\{\alpha : (x(\alpha), s(\alpha)) \in \mathcal{N}_\infty^-(\beta)\}$.

Let $x := x(\alpha^\star);\ s := s(\alpha^\star)$.

repeat

Theorem 4.2. *When we apply the Algorithm 4.1 to a LP problem, we can deduce that the algorithm will converge to an ϵ-solution in at most $O(n\ln(n)\ln\frac{1}{\epsilon})$ steps.*

Proof. For each $j \in \{1,\dots,n\}$, let $u_j := \frac{x_j s_j}{\mu}$ as before.
Let us consider the search direction w in the wide neighborhood of $\mathcal{N}_\infty^-(\frac{1}{2})$. For the next iteration to stay in the same neighborhood, the step length α should satisfy the following condition. For every $j \in \{1,2,\dots,n\}$,

$$u_j + \frac{\alpha}{1-\alpha}u_j(\delta\eta - \eta\ln(u_j)) + \frac{\alpha^2}{1-\alpha}\frac{(w_p)_j(w_q)_j}{\mu} \geq 1/2. \tag{4.1}$$

We know from [15] that $|(w_p)_i(w_q)_i| \leq \|w(\eta)\|^2/4$. Since all of our discussion is within $\mathcal{N}_\infty^-(\frac{1}{2})$, we know that: $u_j \geq \frac{1}{2}$, $\forall j$. We also deduce that $u_j \leq \frac{n+1}{2}$, $\forall j$ from the fact $\sum_{j=1}^{n} u_j = n\mu$. So we know:

$$0 \leq \Delta_{12} \leq \max(\ln^2 u_j)\sum_{j=1}^{n} u_j = n\ln^2\left(\frac{n+1}{2}\right),\ 0 \leq \delta \leq \ln\tfrac{n+1}{2}.$$

After relaxation, (4.1) can be transformed to the following formulation:

$$u_j + \frac{\alpha}{1-\alpha}u_j(\delta\eta - \eta\ln(u_j)) - \frac{\alpha^2}{1-\alpha}\left(\frac{n\mu + \eta^2\mu\Delta_{12} - n\mu\delta^2\eta^2}{4\mu}\right) \geq 1/2,$$

i.e.,

$$4\left(u_j - \frac{1}{2}\right) + 4\alpha\left(u_j\delta\eta - u_j\ln(u_j)\eta - u_j + \frac{1}{2}\right) - \alpha^2(n + \eta^2\Delta_{12} - n\delta^2\eta^2) \geq 0.$$

We will discuss 3 cases by the range of u_j.
If $u_j \geq 1$, further relaxation can achieve the following:

$$-\alpha^2\left(n + n\ln^2\frac{n+1}{2}\eta^2\right) + 1 + 4\alpha\left(u_j\delta\eta - \eta u_j\ln(u_j) - u_j + \frac{1}{2}\right) + 2\left(u_j - \frac{1}{2}\right) \geq 0,$$

i.e.,

$$-\alpha^2\left(n+n\ln^2\frac{n+1}{2}\eta^2\right)+1+2\left(u_j-\frac{1}{2}\right)\left[2\frac{u_j}{u_j-\frac{1}{2}}\alpha\left(\delta\eta-\eta\ln(u_j)-1\right)+1\right]+2\alpha\geq 0.$$

Let $\alpha \leq \frac{1}{4\sqrt{n}\ln(n)\eta}$, then the above inequality holds.

If $1 > u_j \geq \frac{9}{16}$, relaxation can result in the following formulation:

$$-\alpha^2\left(n+n\ln^2\frac{n+1}{2}\eta^2\right)-2\alpha+\frac{1}{4}\geq 0.$$

Let $\alpha \leq \frac{1}{16\sqrt{n}\ln(n)\eta}$, then the above inequality holds.

If $\frac{1}{2} \leq u_j \leq \frac{9}{16}$, relaxation can result in the following formulation:

$$-\alpha^2(n+\Delta_{12}\eta^2)+4\alpha u_j\delta\eta+4\alpha\left(\frac{1}{2}\ln\frac{16}{9}\eta-\frac{1}{16}\right)\geq 0,$$

for $\eta \geq 1$, further relaxation can achieve the following inequality,

$$-\alpha^2(n+\Delta_{12}\eta^2)+2\alpha\delta\eta+\frac{3}{4}\alpha\geq 0.$$

Using Lemma 3.6, we know that if $\alpha \leq \frac{1}{10n\ln(n)\eta}$ with $\eta \geq 1$, then the above inequality holds.

The step length should have the value of at least $\min\left\{\frac{1}{10n\ln(n)\eta}, \frac{1}{16\sqrt{n}\ln(n)\eta}\right\}$.

For the case $\eta = 1$, we can see the $\alpha = \frac{1}{10n\ln(n)}$ will be a proper step length for each step. Hence the iteration complexity for Algorithm 4.1 is $O(n\ln(n)\ln\frac{1}{\epsilon})$ according to Lemma 3.9. □

We note that we need $\|w(\eta)\|_2^2$ to be small to achieve better iteration complexity, so we consider another choice of η as below.

Algorithm 4.2.

Input $(A, x^0, s^0, b, c, \beta, \epsilon)$, *where* A, x^0, s^0, b, c *are defined in the* (HLP) *formulation in Chapter 2,* ϵ *is the desired tolerance,* β *determines the wide neighborhood we set.*

while $x^T s > \epsilon$

if $\left\|\frac{x_j s_j}{\mu} - e\right\|_2^2 \leq \frac{1}{4n}$,

calculate the affine-scaling direction (d_x, d_s) *as stated in Chapter 2,*

define

$x(\alpha) := x + \alpha d_x > 0,$

$s(\alpha) := s + \alpha d_s > 0,$

$\alpha^\star := \max\{\alpha : (x(\alpha), s(\alpha)) \in \mathcal{N}_\infty^-(\beta)\}.$

Let $x := x(\alpha^\star),\ s := s(\alpha^\star)$;

else

define $\eta := \frac{1}{\sqrt{\frac{\Delta_{12}}{n} - \delta^2}}$,

calculate $\Delta_{12}, \delta, D, v, w(\eta), w(\eta)_p, w(\eta)_q$ *as defined before,*

solve the system (HCPS) to get the unique solution d_x *and* d_s,

define

$x(\alpha) := x + \alpha d_x > 0,$

$s(\alpha) := s + \alpha d_s > 0,$

$\alpha^\star := \max\{\alpha : (x(\alpha), s(\alpha)) \in \mathcal{N}_\infty^-(\beta)\}.$

Let $x := x(\alpha^\star),\ s := s(\alpha^\star)$.

end

repeat

Lemma 4.3. *Let $x > 0$, $s > 0$ be such that $\Delta_{12} > n\delta^2$. Then for $\eta := \frac{1}{\sqrt{\frac{\Delta_{12}}{n} - \delta^2}}$, we have $\|w(\eta)\|_2^2 = 2n\mu$.*

Proof. Since $\Delta_{12} > n\delta^2$, η is well defined, then this lemma is trivial according to Lemma 4.1. □

Lemma 4.4. *Let $(x, s) \in \partial\mathcal{N}_\infty^-(\frac{1}{2})$. Then $\eta = \frac{1}{\sqrt{\frac{\Delta_{12}}{n} - \delta^2}} \leq 3\sqrt{n}$.*

Proof. $\eta = \frac{1}{\sqrt{\frac{\Delta_{12}}{n} - \delta^2}} = \frac{\sqrt{n}}{\sqrt{\Delta_{12} - n\delta^2}}$. Hence we need to estimate the value of $(\Delta_{12} - n\delta^2)$ at the boundary of $\mathcal{N}_\infty^-(\frac{1}{2})$, i.e., we have at least one entry of u to be $\frac{1}{2}$, without loss of generality, we set $u_n = \frac{1}{2}$. Therefore, we have:

$$
\begin{aligned}
\Delta_{12} - n\delta^2 &= \frac{1}{2}\ln^2(2) + \sum_{j=1}^{n-1} u_j \ln^2(u_j) - \frac{1}{n}\left(\sum_{j=1}^{n-1} u_j \ln(u_j) - \frac{\ln 2}{2}\right)^2 \\
&= \frac{1}{2}\ln^2(2) + \frac{\ln 2}{n}\left(\sum_{j=1}^{n-1} u_j \ln(u_j) - \frac{\ln 2}{4}\right) + \sum_{j=1}^{n-1} u_j \ln^2(u_j) - \frac{1}{n}\left(\sum_{j=1}^{n-1} u_j \ln(u_j)\right)^2 \\
&= \frac{1}{2}\ln^2(2) + \frac{\ln 2}{n}\left(n\delta + \frac{\ln 2}{4}\right) + \left[\sum_{j=1}^{n-1} u_j \ln^2(u_j) - \frac{1}{n}\left(\sum_{j=1}^{n-1} u_j \ln(u_j)\right)^2\right] \\
&\geq \frac{1}{2}\ln^2(2) + \left[\sum_{j=1}^{n-1} u_j \ln^2(u_j) - \frac{1}{n - \frac{1}{2}}\left(\sum_{j=1}^{n-1} u_j \ln(u_j)\right)^2\right] \\
&\geq \frac{1}{2}\ln^2(2).
\end{aligned}
$$

The first inequality is due to the nonnegativity of δ and $\frac{1}{n} < \frac{1}{n-\frac{1}{2}}$. The second inequality can be deduced by Cauchy-Schwartz inequality using similar method as in Lemma 3.5. Therefore $\eta \leq \frac{\sqrt{2n}}{\ln 2} \leq 3\sqrt{n}$. □

Theorem 4.5. *If we apply the Algorithm 4.2 to $\mathcal{N}_\infty^-(\frac{1}{2})$, we can deduce that the algorithm converges to an ϵ-solution in at most $O(n \ln(n) \ln \frac{1}{\epsilon})$ steps.*

Proof. For each $j \in \{1, \ldots, n\}$, let $u_j := \frac{x_j s_j}{\mu}$ as before. We use w to stand for $w(\eta)$ in this proof. Let us consider the search direction w in the wide neighborhood of $\mathcal{N}_\infty^-(\frac{1}{2})$. For the next iteration to stay in the same neighborhood, the step length α should satisfy the following condition for every $j \in \{1, 2, \dots, n\}$,

$$u_j + \frac{\alpha}{1-\alpha} u_j(\delta\eta - \eta \ln(u_j)) + \frac{\alpha^2}{1-\alpha} \frac{(w_p)_j (w_q)_j}{\mu} \geq 1/2. \tag{4.2}$$

We know from [15] and Lemma 4.1, Lemma 4.2 that $|(w_p)_j(w_q)_j| \leq \|w(\eta)\|^2/4 = \frac{n\mu}{2}$.
After relaxation, (4.2) can be transformed to the following formulation:

$$u_j + \frac{\alpha}{1-\alpha} u_j(\delta\eta - \eta \ln(u_j)) - \frac{\alpha^2}{1-\alpha} \frac{n}{2} \geq 1/2,$$

i.e.,

$$2\left(u_j - \frac{1}{2}\right) + 2\alpha\left(u_j\delta\eta - u_j \ln(u_j)\eta - u_j + \frac{1}{2}\right) - n\alpha^2 \geq 0.$$

We discuss 3 cases by the range of u_j.
If $u_j \geq 1$, further relaxation can achieve the following:

$$-2n\alpha^2 + 1 + 4\alpha\left(u_j\delta\eta - \eta u_j \ln(u_j) - u_j + \frac{1}{2}\right) + 2\left(u_j - \frac{1}{2}\right) \geq 0.$$

i.e.,

$$-2n\alpha^2 + 1 + 2\left(u_j - \frac{1}{2}\right)\left[2\frac{u_j}{u_j - \frac{1}{2}}\alpha\left(\delta\eta - \eta\ln(u_j) - 1\right) + 1\right] + 2\alpha \geq 0.$$

We know the estimation of η from Lemma 4.4 and we can deduce that the above inequality holds for $\alpha \leq \frac{1}{6\sqrt{n}\ln n}$.
If $\frac{9}{16} \leq u_j < 1$, relaxation can result in the following formulation:

$$-2n\alpha^2 - 2\alpha + \frac{1}{4} \geq 0.$$

Let $\alpha \leq \frac{1}{16\sqrt{n}}$, then the above inequality holds.

If $\frac{1}{2} \le u_j \le \frac{9}{16}$, relaxation can result in the following formulation:

$$-2n\alpha^2 + 2\alpha u_j \delta\eta + 2\alpha\left(\frac{1}{2}\ln\frac{16}{9}\eta\right) + (1-\alpha)\left(u_j - \frac{1}{2}\right) \ge 0,$$

further relaxation can get the following formulation,

$$-2n\alpha^2 + \frac{\alpha\eta}{4} \ge 0.$$

Since $\Delta_{12} < n\ln^2 n$ and $\delta \ge 0$, we know $\eta = \frac{1}{\sqrt{\frac{\Delta_{12}}{n} - \delta^2}} > \frac{1}{\ln(n)}$, we can deduce that the above inequality holds for $\alpha \le \frac{1}{8n\ln(n)}$.

The step length α should have the value of at least $\min\{\frac{1}{6\sqrt{n}\ln(n)}, \frac{1}{16\sqrt{n}}, \frac{1}{8n\ln(n)}\}$. We can conclude the $\alpha = \frac{1}{8n\ln(n)}$ is a proper step length for each step. From Lemma 3.9 we can conclude that the iteration complexity is $O(n\ln(n)\ln\frac{1}{\epsilon})$. □

In order to study the relationship between the parameter η and step length α, we did some analysis and computational experiments to explore the rule for "best" η to achieve maximal step length. (Because, maximizing the step length is equivalent to minimizing the complementary gap along the given direction.)

In order to find the "best" η such that the step length α can achieve maximum within $\mathcal{N}_\infty^-(\frac{1}{2})$ at every iteration, we need to solve a 2-dimensional constrained programming problem in every iteration.

Define $u_j = \frac{x_j s_j}{\mu}$, $v = X^{\frac{1}{2}}S^{\frac{1}{2}}e$, $w(\eta) = -v + \eta\left[\delta v - V\ln(\frac{Vv}{\mu})\right]$ as before, $w_p := P_{AD}w(\eta)$, P_{AD} is the projection operator onto the null space of AD, $w_q := w(\eta) - w_p$.

$$\begin{array}{llll} & \text{maximize} & \alpha & \text{(CS)} \\ (1) & \text{subject to} & 0 < \alpha < 1, & \\ (2) & & \eta \ge 0, & \\ (3) & \forall j \in \{1, 2\ldots, n\}, & \frac{(w_p)_j (w_q)_j}{\mu}\alpha^2 + \alpha\left(u_j\delta\eta - u_j\ln(u_j)\eta - u_j + \frac{1}{2}\right) + \left(u_j - \frac{1}{2}\right) \ge 0. & \end{array}$$

Define:
$t_p := P_{AD}(\delta v - V\ln(\frac{Vv}{\mu}))$, $t_q := \delta v - V\ln(\frac{Vv}{\mu}) - t_p$, $v_p := P_{AD}(-v)$, $v_q := -v - v_p$.

In the plane search for the optimal solution η^* and α^*, we want η^* more accurately than α^*, e.g., we need η^* to have the accuracy of 10^{-10} while we may only need α^* to have the accuracy of 10^{-4} since only η^* will be involved in the computation of the search directions d_x and d_s.

Note that the quadratic terms in α and η for the constraint in (3) have different properties as below.
For η, it has the form $\alpha^2\frac{(t_p)_i(t_q)_i}{\mu}\eta^2$. The sign of the coefficient of η^2 does not change as α changes.
For α, coefficient of α^2 is $\frac{(v_p+\eta t_p)_i(v_q+\eta t_q)_i}{\mu}$. It may change sign as η changes. So the algorithm below will decrease α in some strategy and try to find a feasible η for a fixed α.

Algorithm 4.3.
Input $(A, x^0, s^0, b, c, \epsilon)$, *where* A, x^0, s^0, b, c *are defined in the* (HLP) *formulation in Chapter 2,* ϵ *is the desired tolerance.*
while $x^T s > \epsilon$
    $\alpha := 1$;
    while $\alpha > 0$
        calculate $\delta, u, v_p, v_q, t_p, t_q$ *as defined before, and transform*
        (CS) into quadratic system of η*. Then solve the system (CS)*
        to verify if there exists positive η *satisfying the constraints (3) of (CS) for current* α*.*
        if such η *exists,* $\eta^\star := \eta$, $\alpha_\eta := \alpha$, *exit;*
        if such η *does not exists, decrease* α *as below:*
        if $\alpha > \frac{1}{10}$, *let* $\alpha := \alpha - 0.05$; *else* $\alpha := 0.95\alpha$.

    repeat
    calculate $\delta, D, v, w(\eta^*), w(\eta^*)_p, w(\eta^*)_q$ *as defined before,*
    solve the system (HCPS) to get the unique solution d_x *and* d_s*.*
    Let $x := x(\alpha_{\eta^*})$; $s := s(\alpha_{\eta^*})$.

repeat

For the procedure to verify if there exists positive η which satisfies the n inequalities in the constraints (3) of (CS), each constraint is a quadratic problem for η and can induce a feasible interval for η. If the intersection of all the intervals corresponding to the n inequality constraints is not empty, we then find the η corresponding to a step length α. We use the following procedure to determine the feasible interval of η for a given step length α.

Algorithm 4.4.
Input $\delta, u, v_p, v_q, t_p, t_q, \alpha$, *where* $\delta, u, v_p, v_q, t_p, t_q$ *as defined before.*
α *is the current step length for testing.*
For $j = 1$ *to* n

$a_0(j) := \alpha^2 \frac{(t_p)_j (t_q)_j}{\mu}$,

$b_0(j) := \alpha u_j \delta - u_j \ln u_j \alpha - \alpha^2((v_p)_j (t_q)_j + (v_q)_j (t_p)_j)$,

$c_0(j) := \alpha^2 (v_p)_j (v_q)_j + (1-\alpha)(u_j - \frac{1}{2})$.

Solve the inequality of η, *i.e.,* $a_0(j)\eta^2 + b_0(j)\eta + c_0(j) \geq 0$.

If $a_0(j) = 0$, $b_0(j) = 0$ *and* $c_0(j) \geq 0$,

$r1(j) := 0$, $r2(j) := 0$, $r3(j) := -\infty$, $r4(j) := +\infty$;

if $a_0(j) = 0$, $b_0(j) = 0$ *and* $c_0(j) < 0$,

$r1(j) := -\infty$, $r2(j) := +\infty$, $r3(j) := -\infty$, $r4(j) := -\infty$;

If $a_0(j) = 0$ *and* $b_0(j) > 0$,

$r1(j) := 0$, $r2(j) := 0$, $r3(j) := -\frac{c_0(j)}{b_0(j)}$, $r4(j) := +\infty$;

if $a_0(j) = 0$ *and* $b_0(j) < 0$,

$r1(j) := 0$, $r2(j) := 0$, $r3(j) := -\infty$, $r4(j) := -\frac{c_0(j)}{b_0(j)}$;

If $a_0(j) > 0$ *and* $b_0(j)^2 - 4a_0(j)c_0(j) < 0$,

$r1(j) := 0$, $r2(j) := 0$, $r3(j) := -\infty$, $r4(j) := +\infty$;

if $a_0(j) < 0$ *and* $b_0(j)^2 - 4a_0(j)c_0(j) < 0$,

$r1(j) := -\infty$, $r2(j) := +\infty$, $r3(j) := -\infty$, $r4(j) := -\infty$;

If $a_0(j) > 0$ *and* $b_0(j)^2 - 4a_0(j)c_0(j) \geq 0$,

$r1(j) := \frac{-b_0(j) - \sqrt{b_0(j)^2 - 4a_0(j)c_0(j)}}{2a_0(j)}$, $r2(j) := \frac{-b_0(j) + \sqrt{b_0(j)^2 - 4a_0(j)c_0(j)}}{2a_0(j)}$,

$r3(j) := -\infty$, $r4(j) := +\infty$;

if $a_0(j) < 0$ *and* $b_0(j)^2 - 4a_0(j)c_0(j) \geq 0$,

$r1(j) := 0$, $r2(j) := 0$,

$r3(j) := \frac{-b_0(j) + \sqrt{b_0(j)^2 - 4a_0(j)c_0(j)}}{2a_0(j)}$, $r4(j) := \frac{-b_0(j) - \sqrt{b_0(j)^2 - 4a_0(j)c_0(j)}}{2a_0(j)}$;

end

repeat

Define $[t1, t2] := [\max_j(r3(j)), \min_j(r4(j))]$.

The feasible region is the intersection of $(-\infty, r2(j)] \cup [r1(j), +\infty)$, $j \in \{1, \ldots, n\}$ *with* $[t1, t2]$.

While $\exists j$ *such that* $(r2(j)-t1)(t2-r1(j))<0$

$(r2(j)-t1)(t2-r1(j))<0$ *means that* $(-\infty, r2(j)] \cup [r1(j), +\infty)$ *intersects with* $[t1, t2]$ *at only one end.*

if $(r2(j)-t1)(t2-r1(j)) \leq 0$ *and* $r2(j) \geq t1$,

$t2 := r2(j)$;

if $(r2(j)-t1)(t2-r1(j)) \leq 0$ *and* $r1(j) \leq t2$,

$t1 := r1(j)$;

if $(r2(j)-t1)(t2-r1(j)) \geq 0$ *and* $r1(j) > t2$ *or* $r2(j) < t1$,

stop searching η *for current* α.

end

repeat

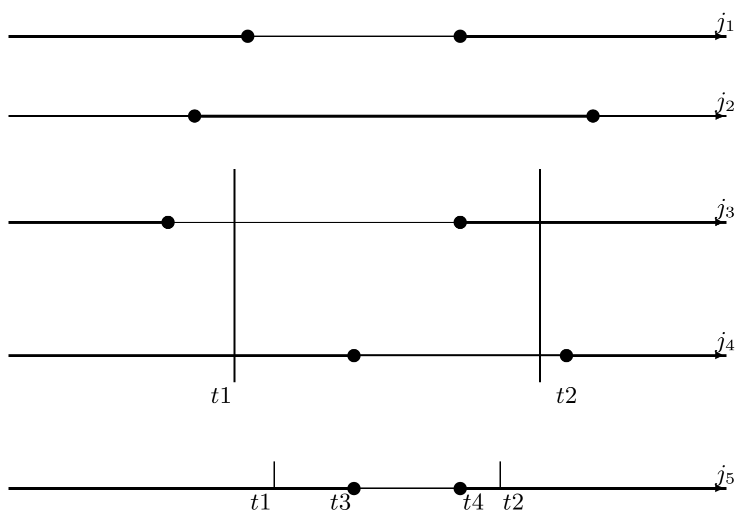
Define $M_0 := \{j : t1 \leq r2(j) \leq t2, t1 \leq r1(j) \leq t2\}$.

If $t1 \leq t2$,

the union of intervals $[t1, \min_{j \in M_0}(r1(j))]$ *and* $[\max_{j \in M_0}(r2(j)), t2]$ *are the feasible region for* η *if one of them exists. We choose the maximal* η *for convenience.*

end

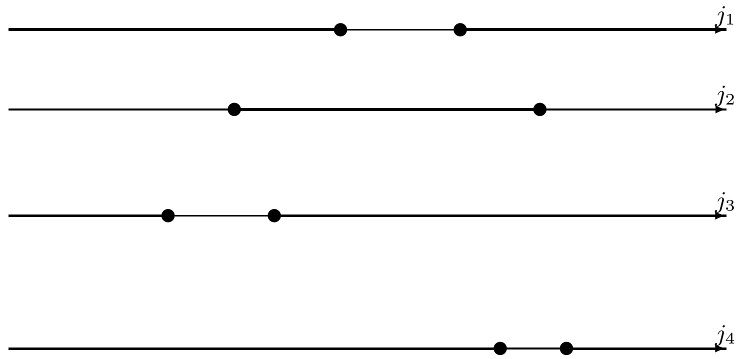
The sketch for the η verifying procedure.



Lemma 4.6. *The Algorithm 4.4 defined above is valid to determine if the feasible η exists and can find the "best" η if it exists.*

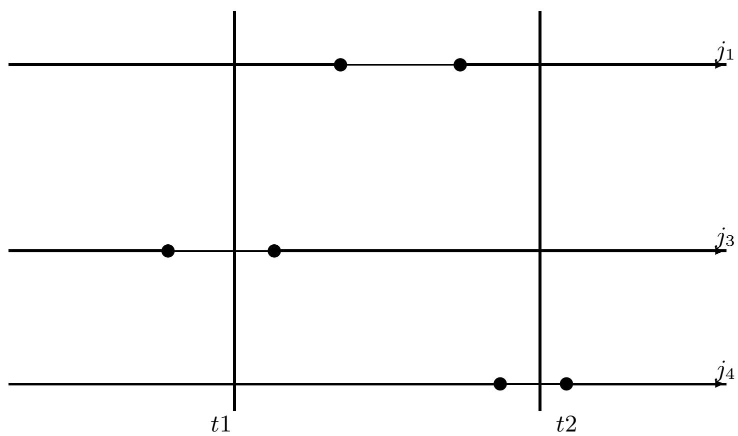
Proof. From the Algorithm 4.4 defined above, we can see that there are two kinds of intervals for the constraints in (3). One form is the union of two open intervals ,i.e., $(-\infty, r1(j)]$ and $[r2(j), \infty)$, denote the constraints in this class as K_1. Another is the convex interval $[r3(j), r4(j)]$ and denoted by K_2 the constraints in this class. These two cases are shown in the j_1 and the j_2 axis above. It is easy to find the intersection of the convex intervals which is $[t1, t2] := [\max_{j \in K_2}(r3(j)), \min_{j \in K_2}(r4(j))]$. While it is hard to have the intersection of $[t1, t2]$ with the intervals in class K_1. In the above algorithm, we decrease the closed interval $[t1, t2]$ by intersecting it with the intervals which intersect with the convex interval at only one end, i.e., to eliminate the cases belong to the situations of the j_3 and the j_4 axis. If the convex interval $[t1, t2]$ exists, we can assume that all the remaining intervals in K_1 intersect the convex interval at both ends. Denote these remaining intervals by M', then we find two feasible intervals for η as shown in the j_5 axis, i.e., $[t1, t3 := \min_{j \in M'}(r1(j))]$ or $[t4 := \max_{j \in M'}(r2(j)), t2]$. $\square$

We give out an example as below.



One instance for the η verifying procedure.(step 1)

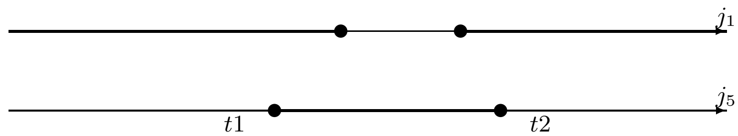
We will try to find the intersection of intervals on the above 4 axis by using Algorithm 4.4



One instance for the η verifying procedure.(step 2) j_2 processed.

We will try to find the intersection of convex interval $[t_1, t_2]$ with the non-convex intervals.

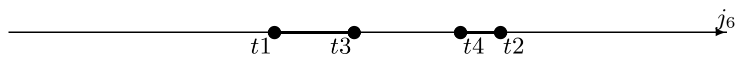
One instance for the η verifying procedure (step 3). j_3, j_4 processed.



t_1, t_2 are adjusted according to the intervals on j_3, j_4 which intersect with $[t1, t2]$ only at one end.

One instance for the η verifying procedure (step 4). j_1 processed.

$[t_1, t3]$, $[t_4, t_2]$ is the feasible region of η.



With Algorithm 4.3, we improve the iteration complexity dramatically in practice as shown in the next chapter. We analyze its iteration complexity as below.

Theorem 4.7. *When we apply the Algorithm 4.3 to solve LP problems, we can deduce that the algorithm will converges to an ϵ-solution in at most $O(n\ln(n)\ln\frac{1}{\epsilon})$ steps.*

Proof. From the analysis of Algorithm 4.1, we can see the step length is at least $\frac{1}{10n\ln n}$ for each iteration step if $\eta = O(1) \geq 1$. From the definition of Algorithm 4.3, we know the step length derived for each iteration should be at least $\frac{1}{10n\ln n}$, i.e., longer than Algorithm 4.1. So from Lemma 3.9, we know the iteration complexity for Algorithm 4.3 is at most $O(n\ln(n)\ln\frac{1}{\epsilon})$. ☐

The result above is not satisfying. It is no better than the algorithms using some fixed η, we may need some other theoretical "tools" to explore if better iteration complexity can be achieved in the future.

Chapter 5

Computational Results on the Entropic Direction Family

We did some experiments on a PC (HP P-M 1.4G 256M RAM) using the software MATLAB version 6.5. The LP problems for test are those in the problem set of NETLIB [33]. They represented most of the cases of interesting LP problems.

We implemented the Algorithms 4.1 -4.3 in MATLAB code for the artificial HSD LP problems as shown in Chapter 2. In the coding we reused some input and output subroutines from the LIPSOL software [32] developed by Zhang. The initial feasible solution is $y^0 := 0$, $x^0 := e$, $s^0 := e$, $\theta = 1$, $t = 1$ and $\kappa = 1$. we solve the quadratic inequality system $\frac{X(\alpha)s(\alpha)}{(1-\alpha)\mu} \geq \frac{1}{2}e$ to get the maximal step length α such that $(x(\alpha), s(\alpha)) \in \mathcal{N}_\infty^-(\frac{1}{2})$ in each iteration. This method is of better efficiency than backtracking method. Define $\epsilon := 10^{-6}$, we set the termination criteria as following:

$$(x/t)^T s/t < \epsilon,\ \|Ax/t - b\|_\infty < \epsilon \text{ and } \|A^T y/t + s/t - c\|_\infty < \epsilon.$$

If the criteria is satisfied, we will get one ϵ-solution of the original problem in the form $(\frac{x}{t}, \frac{y}{t}, \frac{s}{t})$ from a strictly self-complementary solution $(y, x, \theta = 0, t, \kappa, s)$ of the HSD problem [28].

In order to avoid the accumulated error in the equality constraints, after each iteration we let:
$\overline{b}$=$\overline{b} - \chi_1/\theta$,
$\overline{c}$=$\overline{c} + \chi_2/\theta$,

$\overline{z}$=$\overline{z} - \chi_3/\theta$,
where χ_i is the error, i.e.,
$\chi_1 = Ax - bt + \overline{b}\theta$,
$\chi_2 = A^T y + ct - \overline{c}\theta - s$,
$\chi_3 = b^T y - c^T x + \overline{z}\theta - \kappa$.
By using this method, we can dump errors into the column that is guaranteed to be a nonbasic column. This technique is inspired by the content in paper [27].

The table below includes the majority of NETLIB problems and compares the performance of the directions proposed in this thesis within the wide neighborhood $\mathcal{N}_\infty^-(1/2)$. All the search directions belong to the entropic family we proposed in Chapter 3 and have the form $w(\eta) = -v + \eta[\delta v - V \ln(\frac{Vv}{\mu})]$.

The parameter η is the key factor for the performance of the algorithms. The columns $\eta = 1, \ldots, 4$ denote the iteration numbers we used to achieve ϵ-solution using Algorithm 4.1 with different choice of $\eta = O(1)$. the column η^0 denotes the iteration numbers we used to achieve ϵ-solution with Algorithm 4.2. The last column η^* stands for the iteration numbers for Algorithm 4.3 to achieve ϵ-solution.

From the table below, we find that Algorithm 4.2 has the worst performance and Algorithm 4.3 needs the least number of iterations to achieve convergence. For Algorithm 4.1, different values of η result in different performance even for the same LP problem while it seems $\eta = 2$ wins most of the problems in NETLIB. We think it is interesting to have some statistical results for the "best" η in Algorithm 4.3 so that we may find some rules inside.

We select the problems "bandm", "scsd6", and "lotfi" as examples. The following figures give out some impression for the value of η^*, α^*, δ and Δ_{12}. We also give out the figures for the value of $\frac{\Delta_{12}}{(n\delta)}$ and $\frac{\Delta_{12}}{(n\delta^2)}$. When the iterate is on the central path, we define $\frac{\Delta_{12}}{(n\delta)} := 0$ and $\frac{\Delta_{12}}{(n\delta^2)} := 0$.

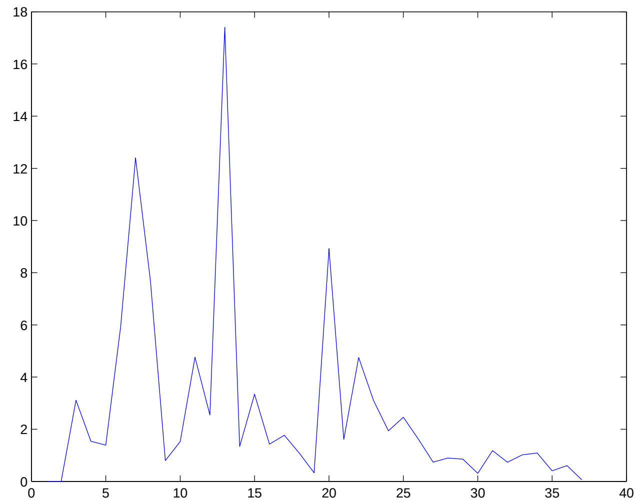


Figure 5.1: The "best" η for each iteration for "bandm"

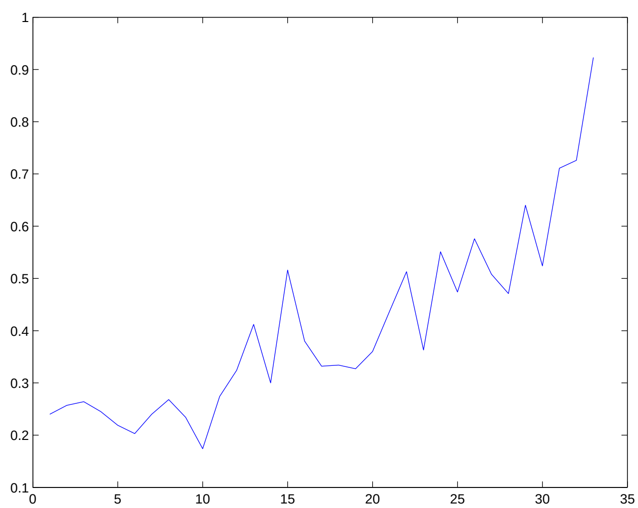


Figure 5.2: Step length for each iteration for "bandm"

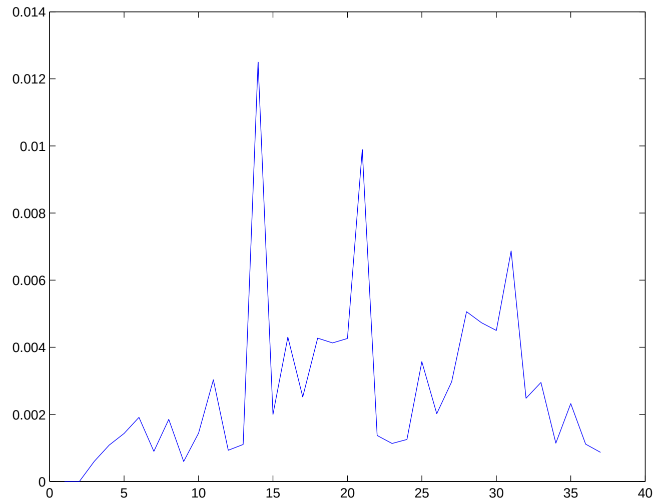


Figure 5.3: The value of δ for each iteration for "bandm"

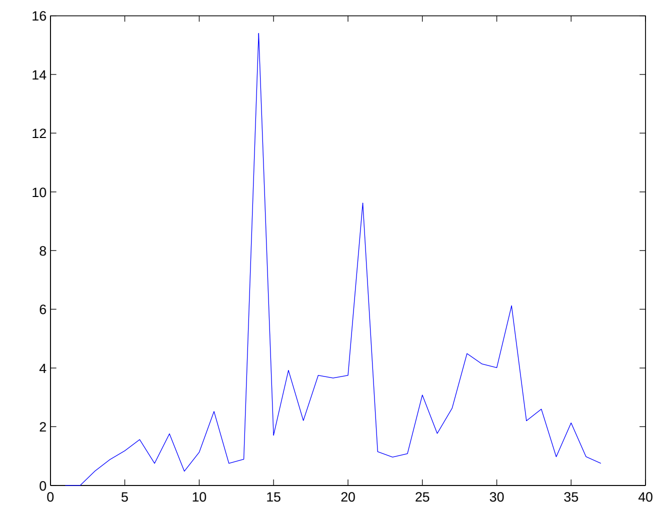


Figure 5.4: The value of Δ_{12} for each iteration for "bandm"

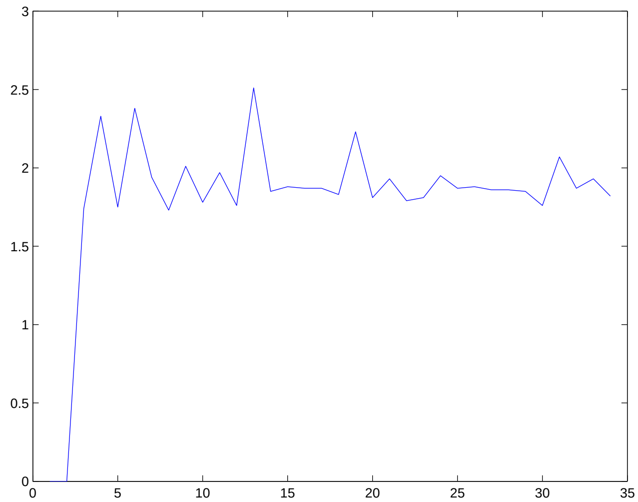


Figure 5.5: The value of $\frac{\Delta_{12}}{(n\delta)}$ for each iteration for "bandm"

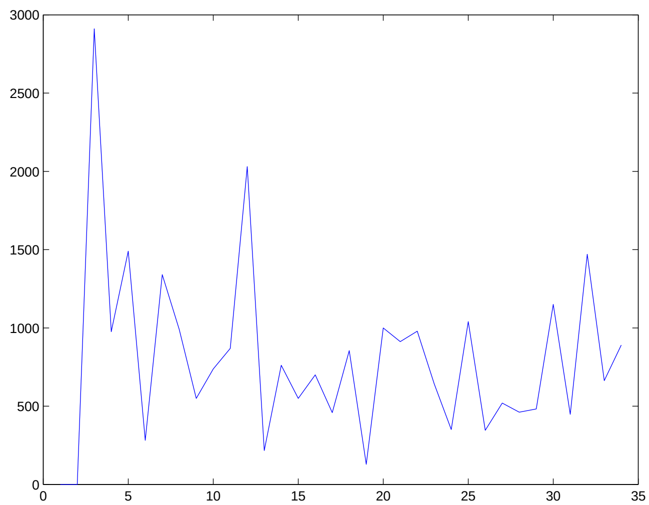


Figure 5.6: The value of $\frac{\Delta_{12}}{(n\delta^2)}$ for each iteration for "bandm"

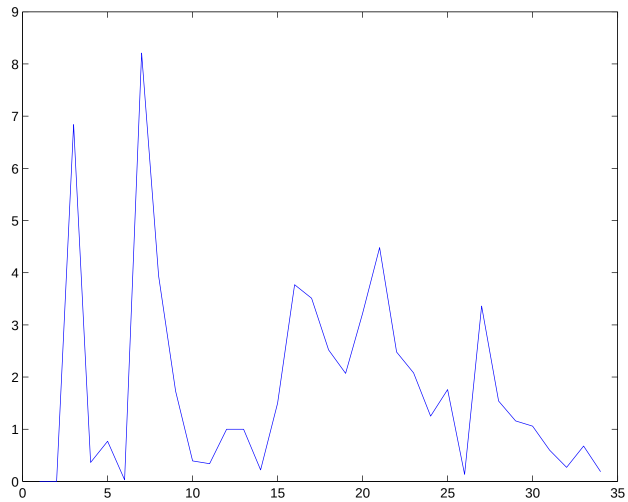


Figure 5.7: The "best" η for each iteration for "lotfi"

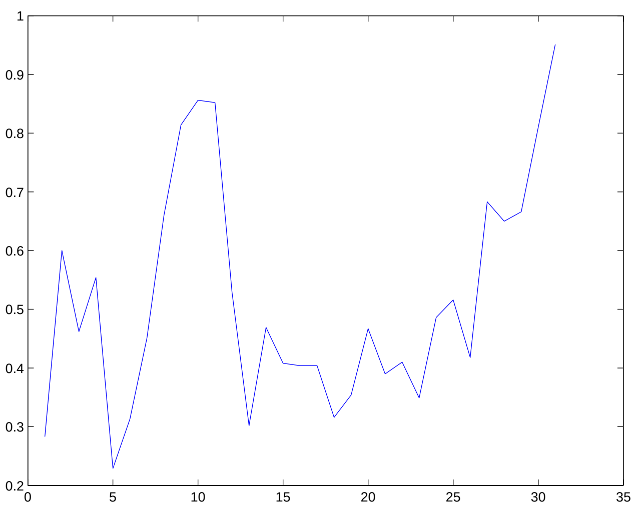


Figure 5.8: Step length for each iteration for "lotfi"

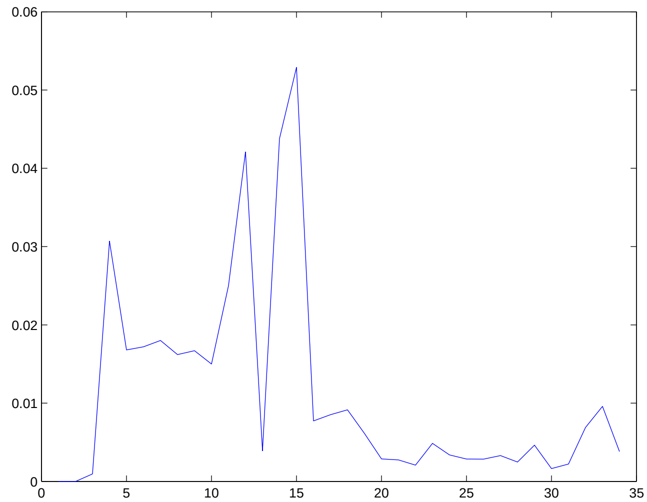


Figure 5.9: The value of δ for each iteration for "lotfi"

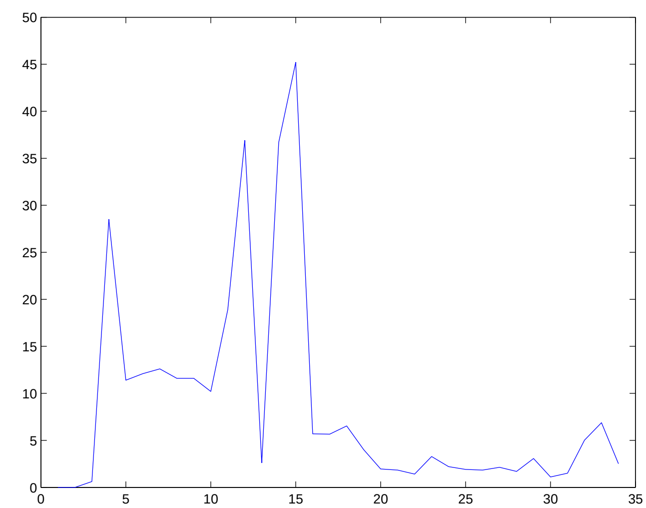


Figure 5.10: The value of Δ_{12} for each iteration for "lotfi"

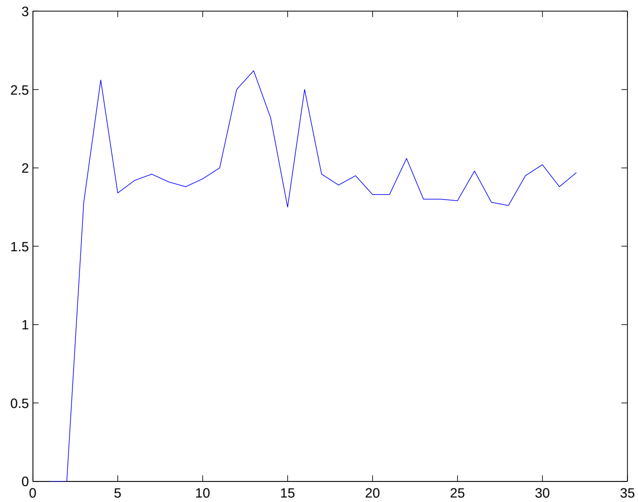


Figure 5.11: The value of $\frac{\Delta_{12}}{(n\delta)}$ for each iteration for "lotfi"

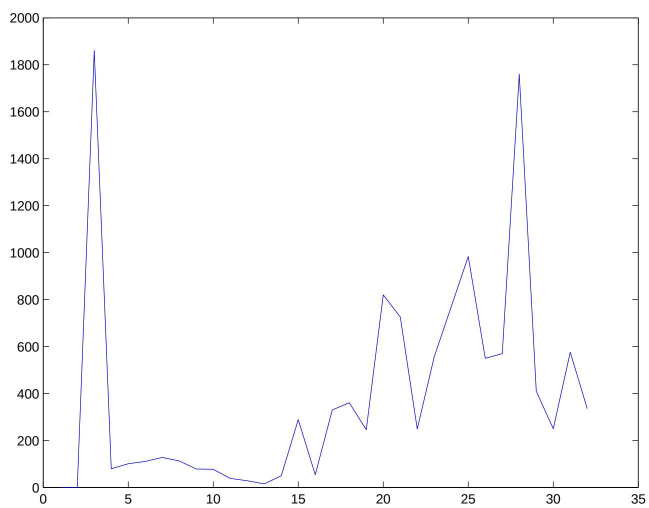


Figure 5.12: The value of $\frac{\Delta_{12}}{(n\delta^2)}$ for each iteration for "lotfi"

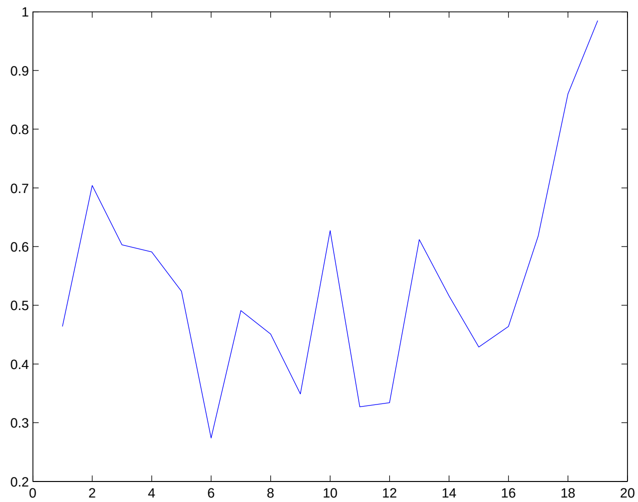


Figure 5.13: The value of α for each iteration for "scsd6"

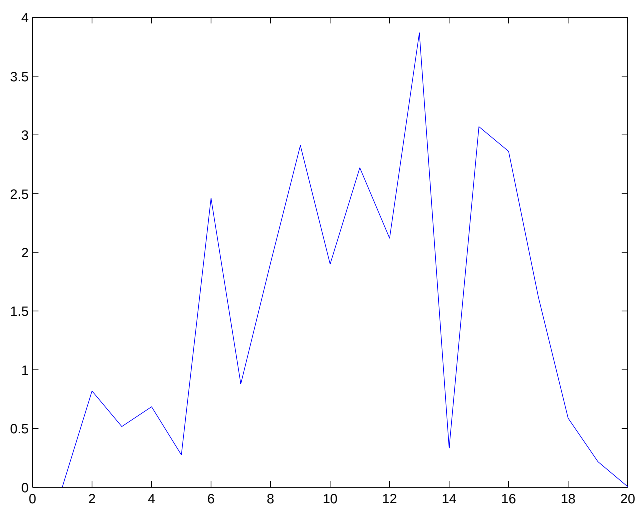


Figure 5.14: The value of η for each iteration for "scsd6"

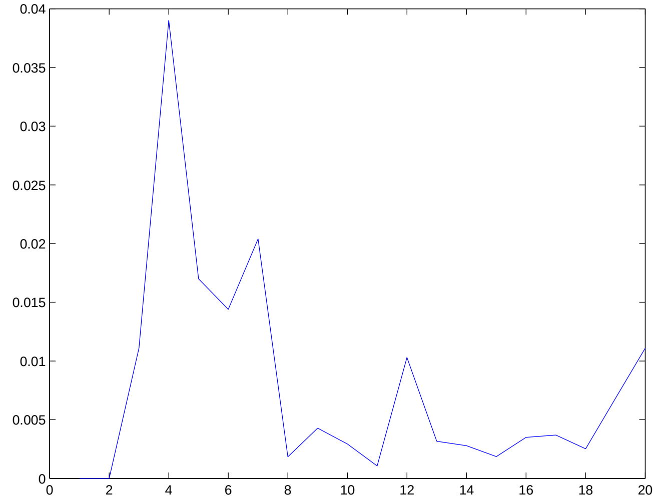


Figure 5.15: The value of δ for each iteration for "scsd6"

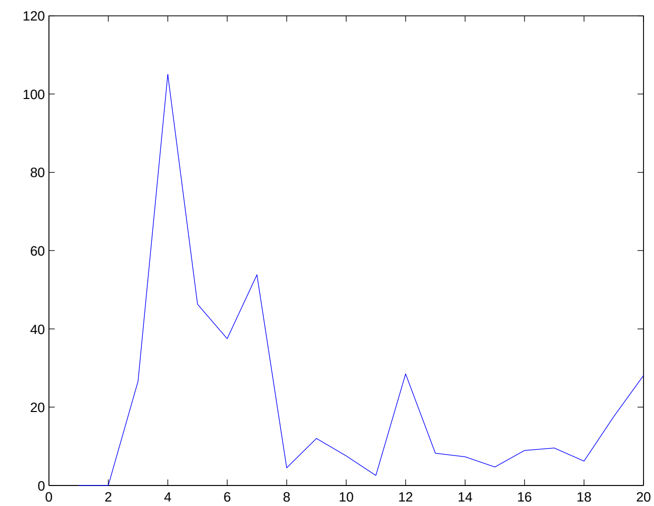


Figure 5.16: The value of Δ_{12} for each iteration for "scsd6"

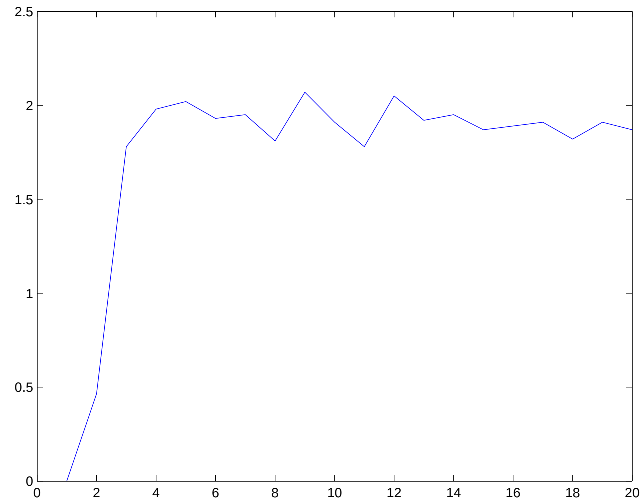


Figure 5.17: The value of $\frac{\Delta_{12}}{(n\delta)}$ for each iteration for "scsd6"

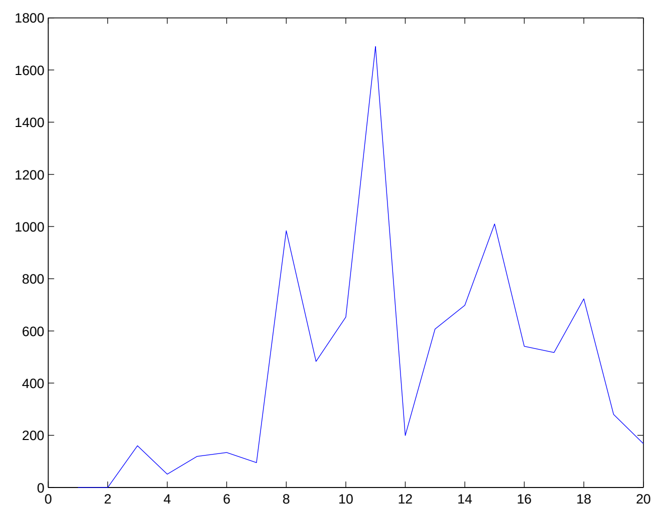


Figure 5.18: The value of $\frac{\Delta_{12}}{(n\delta^2)}$ for each iteration for "scsd6"

Remark: From the figures above, we observe that $\frac{\Delta_{12}}{n\delta}$ has a upper bound 3 in all the 3 cases. If this is true in general cases, this will help us to improve the iteration complexity of Algorithm 4.2 to be $O(n\sqrt{\ln n}\ln(\frac{1}{\epsilon}))$. We hope that the value of $\frac{\Delta_{12}}{n\delta^2}$ may also have some upper bound, then we can deduce that the Algorithm 4.2 may have better iteration complexity than we proved in Chapter 4. This is an open problem we may explore in the future. We also tested the Algorithm 4.3 without the restriction that η is positive. We observe that although in some steps, negative η is feasible, we always can find some positive η which is also feasible for the given α. It is an interesting problem to explore whether there is always a positive η for the maximal possible step length.

Direction			$\eta = 1$	$\eta = 2$	$\eta = 3$	$\eta = 4$	η^0	η^*
NETLIBName	Dimension	Nonzero						
$sc105$	$106 * 103$	281	**29**	38	49	68	52	20
$sc205$	$206 * 203$	552	**30**	36	49	69	761	22
$scagr7$	$130 * 140$	553	**35**	39	46	57	56	26
$scagr25$	$472 * 500$	2029	**39**	44	54	67	86	35
$gfrd - pnc$	$617 * 1092$	3467	**44**	51	55	65	111	32
$afiro$	$28 * 32$	88	**28**	31	37	44	64	17
$scsd1$	$78 * 760$	3148	**27**	31	37	42	71	18
$scsd8$	$398 * 2750$	11334	**32**	33	39	47	92	23
$grow22$	$441 * 946$	8318	**57**	70	79	111	156	35
$grow7$	$141 * 301$	2633	**44**	58	67	78	144	34
$blend$	$75 * 83$	522	**31**	32	33	41	41	20
$sc50b$	$51 * 48$	119	**27**	34	40	46	41	16
$sc50a$	$51 * 48$	131	**33**	34	38	41	44	18
$grow15$	$301 * 645$	5665	**39**	47	78	103	144	36
$vtp.base$	$199 * 203$	914	**50**	**50**	58	62	122	42
$scsd6$	$148 * 1350$	5666	**33**	**33**	38	42	78	22
$share2b$	$97 * 79$	730	**40**	**40**	41	47	66	22
$e226$	$224 * 282$	2767	61	**53**	57	76	68	38
$tuff$	$334 * 587$	4523	123	**115**	119	123	98	40
$Israel$	$175 * 142$	2358	57	**44**	46	63	80	44
$lotfi$	$154 * 308$	1086	48	**47**	54	60	72	35
$capri$	$272 * 353$	1786	55	**54**	65	63	81	36
$fit1p$	$628 * 1677$	10894	120	**84**	98	133	166	34
$beaconfd$	$174 * 262$	3476	46	**41**	47	56	57	21
$modszk1$	$686 * 1622$	3170	111	**86**	89	112	167	56

Direction			$\eta = 1$	$\eta = 2$	$\eta = 3$	$\eta = 4$	η^0	η^*
degen3	$1504 * 1818$	26230	40	**38**	59	60	69	34
recipe	$92 * 180$	752	40	**38**	44	48	99	19
sctap3	$1481 * 2480$	10734	49	**44**	48	55	113	32
kb2	$44 * 41$	291	41	**40**	45	54	72	30
ganges	$1310 * 1681$	7021	55	**53**	59	70	165	42
fit1d	$25 * 1026$	14430	52	**47**	51	58	165	35
degen2	$445 * 534$	4449	39	**37**	40	48	69	33
ship08s	$779 * 2387$	9501	62	**55**	63	67	127	34
ship04l	$403 * 2118$	8450	62	**58**	64	80	161	33
ship04s	$403 * 1458$	5810	60	**56**	59	81	214	34
agg	$489 * 163$	2541	281	**267**	336	> 400	357	49
agg2	$517 * 302$	4515	65	**59**	63	78	140	42
agg3	$517 * 302$	4531	58	**54**	61	67	129	42
boeing1	$351 * 384$	3865	66	**57**	59	74	162	47
boeing2	$167 * 143$	1339	66	**52**	55	70	113	38
stocfor2	$2158 * 2031$	9492	94	**75**	77	89	157	58
d6cube	$404 * 6184$	37704	77	**74**	80	89	81	37
wood1p	$245 * 2594$	70216	73	**50**	55	89	> 140	39
adlittle	$57 * 97$	465	64	**52**	53	61	77	27
stair	$357 * 467$	3857	86	**37**	56	70	87	31
scfxm3	$991 * 1371$	7846	216	**108**	**108**	167	145	71
Nesm	$663 * 2923$	13988	99	**79**	**79**	89	218	68
brandy	$221 * 249$	2150	70	**50**	**50**	52	79	36
maros	$847 * 1443$	10006	103	77	**72**	83	212	37
ship12s	$1152 * 2763$	10941	147	98	**85**	86	146	51
ship12l	$1152 * 5437$	21597	129	125	**105**	108	187	48

Direction			$\eta = 1$	$\eta = 2$	$\eta = 3$	$\eta = 4$	η^0	η^*
$sctap1$	$301 * 480$	2052	105	78	**73**	79	91	34
$standata$	$360 * 1075$	3038	124	87	**80**	87	129	28
$fffff800$	$525 * 854$	6235	101	95	**87**	92	167	61
$stair$	$357 * 467$	3857	47	42	**37**	46	86	32
$stocfor1$	$118 * 111$	474	82	60	**57**	64	88	24
$bnl1$	$644 * 1175$	6129	91	81	**80**	83	120	69
$sctap2$	$1091 * 1880$	8124	53	46	**45**	56	110	34
$pilot4$	$411 * 1000$	5145	109	110	**95**	97	> 160	86
$bore3d$	$234 * 315$	1525	80	60	**52**	63	97	38
$bandm$	$305 * 472$	2659	61	51	**48**	51	61	36
$forplan$	$162 * 421$	4916	93	73	**72**	101	138	56
$scrs8$	$491 * 1169$	4029	136	98	**82**	83	127	50
$sctap2$	$1091 * 1880$	8124	78	66	60	**57**	102	36
$scfxm1$	$331 * 457$	2612	123	162	141	**127**	121	46
$scfxm2$	$661 * 914$	5229	147	164	144	**159**	140	51
$seba$	$516 * 1028$	4874	130	119	102	**100**	115	56
$standmps$	$468 * 1075$	3686	124	81	69	**65**	167	38
$share1b$	$118 * 225$	1182	239	199	175	**161**	129	53
$standgub$	$361 * 1366$	3281	188	150	117	**99**	329	31
$czprob$	$930 * 3523$	14173	237	156	130	**116**	> 350	67
$25fv47$	$822 * 1571$	11127	158	82	92	**80**	116	57
$bnl2$	$2325 * 3489$	16124	146	111	109	**101**	173	65

Bibliography

[1] E. D. Anderson, J. Gondzio, Cs. M*é*sz*á*ros, and X. Xu, "Implementation of interior-point methods for large scale linear programs, in Interior Point Method of Mathematical Programming," T.Terlaky, ed., Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 189-252, 1996.

[2] Y. Q. Bai, M. El Ghami and C. Roos, "A comparative study of kernel function for primal-dual interior-point algorithms in linear optimization," SIAM Journal of Optimization, vol. 15,pp. 101-128, 2004.

[3] A. Ben-Israel, "Newton's Method with Modified Functions," Contemporary Mathematics, vol. 204, pp39-50, 1995.

[4] R. W. Cottle, J. S. Pang and R.E. Stone, "The linear complementarity Problem," Boston Academic, U.S.A., 1992.

[5] A. Decarreau, D. Hilhorst, C. Lemaréchal, and J. Navaza. "Dual methods in entropy maximization. Application to some problems in crystallography," SIAM J. Optimization, Vol.2(2) pp.173-197, 1992.

[6] N. J. Dusaussoy and I. E. Abdou. "The extended MENT algorithm: a maximum entropy type algorithm using prior knowledge for computerized tomography," IEEE Transactions on Signal Processing, Vol. 39(5), pp. 1164-1180, 1991.

[7] S. Erlander, "Accessibility, Entropy and the Distribution and Assignment of Traffic, " Transportation Research, vol. 11, pp. 149-153, 1977.

[8] S. Erlander, "Entropy in Linear Programs," Mathematical Programming, vol. 21, pp. 137-151, 1981.

[9] S. C. Fang, J. R. Rajasekera and H. S. J. Tsao, "Entropy Optimization and Mathematical Programming," Kluwer Academic Publishers, Boston/London/Dordrecht, (ISBN 0-7923-9939-0), 1997.

[10] H. J. Greenberg, W. P. Pierskalla, " Surrogate Mathematical Programming," Operations Research, vol. 19, pp. 924-939, 1970.

[11] B. Jansen, C. Roos, and T. Terlaky, "A polynomial primal-dual Dikin-type algorithm for linear programming," Mathematics of Operations Research, Vol.21, pp. 341-353, 1996.

[12] N. K. Karmarkar, "A new polynomial-time algorithm for linear programming," Combinatorica, vol.4, pp. 373-395, 1984.

[13] X. Li, "Entropy and Optimization," Ph.D. Dissertation, University of Liverpool, U.K., 1987.

[14] S. Mehrotra, Y.Ye, "Finding an interior point in the optimal face of linear programs," Mathematical Programming, vol 62, pp. 497-515, 1993.

[15] S. Mizuno, "A new polynominal time method for a linear complementarity problem" Mathematical Programming, vol. 56, pp. 31-47, 1993.

[16] S. Mizuno, M. J. Todd and Y. Ye, "On adaptive-step primal-dual interior-point algorithms for linear programming," Mathematics of Operations Research, vol. 19, pp. 964-981, 1993.

[17] C. Nadeu and M. Bertran. ”A flatness based generalized optimization approach to spectral estimation,” Signal Processing, pp. 311-320, 1990.

[18] J. L. Nazareth, ”A reformulation of the central path equations and its algorithmic implications,” Technical Report 94-1, Department of Pure and Applied Mathematics, Washington State University, Pullman, WA, 1994.

[19] S. Pan, X. Li, and S. He, ”An infeasible primal-dual interior point algorithm for linear programs based on logarithmic equivalent transformation,” Journal of Mathematical Analysis and Applications, vol. 314, pp.644-660, 2006.

[20] J. Peng, C. Roos, and T. Terlaky, ”Self-regular functions and new search direcion for linear and semidefinite optimization,” Mathmatical Programming, vol. 93, pp. 129-171, 2002.

[21] J. Peng, C. Roos, and T. Terlaky, ”Self-Regularity: A new paradigm for Primal-Dual interior-point algorithms,” Princeton University Press, Princeton, NJ, 2002.

[22] C. E. Shannon, ”The Mathematical Theory of Communication,” Bell System Technical Journal, vol. 27, pp. 279-428 and 623-656, 1948.

[23] L. Tunçel, ”Semidefinite Optimization” lecture notes, 2005.

[24] L. Tunçel, ”Convex optimization: barrier function and interior point methods,” lecture notes, March 1998.

[25] L. Tunçel, M.J.Todd, ”On the interplay among Entropy, Variable Metrics and Potential Functions in Interior-Point Algorithms,” Computational Optimization and Applications, vol. 8, pp. 5-19, 1997.

[26] S. Wright, ”Primal-Dual Interior-Point Methods,” SIAM, Philadephis Hall, U.S.A., 1997.

[27] Y. Ye, "On the finite convergence of interior-point algorithms for linear programming," Mathmatical Programming, vol. 57, pp. 325-335, 1992.

[28] Y. Ye, M. J. Todd and S. Mizuno, "An $O(\sqrt{n}L)$ iteration homogeneous and self-dual linear programming algorithm," Mathematics of Operations Research, vol. 19, pp. 53-67, 1994.

[29] Y. Ye, " Interior Point Algorithms, Theory and Analysis," John Wiley & Sons, Chichester, UK, 1997.

[30] P. Zhang, X. Li, "An infeasible-start path-following method for monotone LCPs," Mathematical and Computer Modelling, vol 39, pp. 23-31, 2003.

[31] Y. Zhang, "On the convergence of a class of infeasible interior point methods for the horizontal linear complementarity problems," SIAM J. Optim Vol 4, pp208-227, 1994.

[32] Y. Zhang, LIPSOL code using MATLAB, www.caam.rice.edu/ zhang/lipsol/, 1992.

[33] The NETLIB lp/data, http://www.netlib.org/lp/data/ .