2014

Curvature as a Complexity Bound in Interior-Point Methods

Murat Mut
Lehigh University

Curvature as a Complexity Bound in Interior-Point Methods

by

Murat Mut

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

in

Industrial Engineering

Lehigh University

August, 2014

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Date

Accepted Date

Dr. Tamás Terlaky, Dissertation Director

Committee Members:

Dr. Tamás Terlaky, Committee Chair

Dr. Frank E. Curtis

Dr. Katya Scheinberg

Dr. Eugenio Schuster

Dr. Yuriy Zinchenko

Acknowledgements

First and foremost, my gratitude goes to my Ph.D supervisor Prof. Tamás Terlaky. Without his guidance, encouragement and support, this thesis simply would not exist. His very broad knowledge of the subject, research insight, insistence for understanding and asking the fundamental questions has inspired me all along. Learning from him and watching the way he approaches a problem has been invaluable experience. It has been a pleasure and privilege to do research with him. I am grateful to the Lehigh ISE faculty and the department members for creating an ideal environment for both graduate study and social life. I consider myself lucky for the opportunity to learn both the theory and computational side of optimization. In particular, I would like to mention Prof. Ted Ralphs for his eye-opening courses addressing the computational methods in optimization. Generous funding of the department and the university has made it easy to attend conferences throughout my Lehigh years. Weekly seminar talks in the department have provided me with a healthy balance of both theory and applications in Operations Research. I thank my thesis committee members Dr. Frank E. Curtis and Prof. Katya Scheinberg for their constructive feedback. Their questions have enabled me to approach my thesis subject from different perspectives. I am lucky to have Dr. Yuriy Zinchenko in my committee who has done research closely related to mine. I thank Prof. Eugenio Schuster for willing to serve in my committee bringing an interdisciplinary viewpoint. My heart is full of gratitude for my parents and sister who always supported me in this long years of my graduate study. They have made sacrifices so that I could finish my Ph.D. Their accomplishments in life are not smaller than mine. I dedicate this thesis to my mother who passed away recently. I thank all my friends at University of Minnesota, Clemson and Lehigh University for their joyful company. I am grateful to them for making my life in U.S more enjoyable. Finally, I thank my wife, for her continuing support and companionship. I am indebted to her for tolerating my isolated mood at difficult phases of my Ph.D research.

Contents

List of Figures

Basic Notation

Abstract

In this thesis, we investigate the curvature of *interior paths* as a component of complexity bounds for interior-point methods (IPMs) in Linear Optimization (LO). LO is an optimization paradigm, where both the objective and the constraints of the model are represented by linear relationships of the decision variables. Among the class of algorithms for LO, our focus is on IPMs which have been an extremely active research area in the last three decades. IPMs in optimization are unique in the sense that they enjoy the best iteration-complexity bounds which are polynomial in the size of the LO problem. The main objects of our interest in this thesis are two distinct curvature measures in the literature, the *geometric* and the *Sonnevend curvature* of the central path. The central path is a fundamental tool for the design and the study of IPMs and we will see both that the geometric and Sonnevend's curvature of the central path are proven to be useful in approaching the iteration-complexity questions in IPMs. While the Sonnevend curvature of the central path has been rigorously shown to determine the iteration-complexity of certain IPMs, the role of the geometric curvature in the literature to explain the iteration-complexity is still not well-understood. The novel approach in this thesis is to explore whether or not there is a relationship between these two curvature concepts aiming to bring the geometric curvature into the picture. The structure of the thesis is as follows. In the first three chapters, we present the basic knowledge of path-following IPMs algorithms and review the literature on Sonnevend's curvature and

the geometric curvature of the central path. In Chapter 4, we analyze a certain class of LO problems and show that the geometric and Sonnevend's curvature for these problems display analogous behavior. In particular, the main result of this chapter states that in order to establish an upper bound for the total Sonnevend curvature of the central path, it is sufficient to consider only the case when the number of inequalities is twice as big as the dimension. In Chapter 5, we study the redundant Klee-Minty (KM) construction and prove that the classical polynomial upper bound for IPMs is essentially tight for the Mizuno-Todd-Ye predictor-corrector algorithm. This chapter also provides a negative answer to an open problem about the Sonnevend curvature posed by Stoer et al. in 1993. Chapter 6 investigates a condition number relevant to the Sonnevend curvature and yields a strongly polynomial bound for that curvature in some special cases. Chapter 7 deals with another self-concordant barrier function, the *volumetric* barrier, and the volumetric path. That chapter investigates some of the basic properties of the volumetric path and shows that certain fundamental properties of the central path fail to hold for the volumetric path. Chapter 8 concludes the thesis by providing some final remarks and pointing out future research directions.

Chapter 1

Introduction

1.1 Linear Optimization

A linear optimization problem can be expressed as follows: Let $n > m$ and A be an $m \times n$ matrix of full rank. For $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$, we consider the *primal* and *dual* linear optimization problems,

$$
\begin{array}{llll@{\qquad}llll}
 & (Primal) & & & & (Dual) & \\
\min & c^T x & & & \max & b^T y & \\
\text{s.t.} & Ax & = b & & \text{s.t.} & A^T y + s & = c \\
 & x & \geq 0, & & & s & \geq 0,
\end{array} \tag{1.1}
$$

where $x, s \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ are vectors of variables. We call the data (A, b, c) an LO problem instance.

In 1947, Dantzig developed the first efficient algorithm, the simplex method, to solve LO problems Dantzig (1965). The simplex method is still a widely used algorithm to solve LO problems. Although the simplex method is efficient in practice, it does not have the theoretical efficiency in the sense of polynomial iteration-complexity. In fact, a problem, referred to as (KM) designed by Klee and Minty (1972) showed that the simplex method with Dantzig's rule requires a number of arithmetical operations which grows exponentially with the number of variables of the problem.

The first polynomial algorithm for LO problems was developed by Khachiyan (1979, 1980). He applied the ellipsoidal method of Shor (1972), and Nemirovski and Yudin (1976) to LO problems expressed in integer data; and obtained a polynomial upper bound for the number of arithmetical operations to find an optimal solution Bland et al. (1981). This bound is $\mathcal{O}(n^4L)$, which depends on the problem dimension n and a number L, the length of the input, i.e., the total number of bits needed to describe the problem data. However Khachiyan's ellipsoid method was not practical for implementations.

In 1984, Karmarkar (1984) developed an algorithm, a so-called interior-point method, which had polynomial iteration-complexity bound of $\mathcal{O}(nL)$, with a total complexity of $\mathcal{O}(n^{3.5}L)$ arithmetic operations, a factor of $\sqrt{n}$ lower than Khachian's ellipsoid method. He also claimed that this algorithm was efficient in practice.

Soon after Karmarkar's work, Sonnevend (1985) introduced the concept of the "central path" and "central path-following" methods. In 1988, Renegar (1988) and Roos and Vial (1988), derived the first central path-following algorithm with an arithmetic operations complexity $\mathcal{O}(n^3L)$, which gave another $\sqrt{n}$ improvement over the Karmarkar's method. This complexity bound is still the best one as of today. The question whether there exists a strongly polynomial algorithm to solve LO problems (depending only the dimension of the problem) is still open.

A unifying theme in this thesis is the curvature of interior paths as a complexity bound in IPMs. The curvature of the central path will be our main focus for this purpose. As we will see later, curvature is a good measure of the number of iterations of path-following algorithms. This approach is multifaceted. By studying the curvature as a complexity bound, one might construct concrete examples of LO problems which give the worst case lower bound for the number of iterations in IPMs similar to what Klee and Minty have proved for the simplex method. On the other hand, the curvature of the central path is a more informative complexity measure than the classical bounds, and studying it more closely might help to understand why IPMs often perform much better than their worst case bound. Moreover, investigating the curvature might enable one to modify and improve existing algorithms possibly under special assumptions. In the rest of this introduction, we give the necessary background for IPMs and the central path.

1.2 IPMs and the central path

We refer to the feasible set $\mathcal{P} = \{x : Ax = b,\ x \geq 0\}$ of system (1.1) as the primal space and the set $\mathcal{D} = \{(y, s) : A^T y + s = c,\ s \geq 0\}$ as the dual space. We define the interior of $\mathcal{P}$ and $\mathcal{D}$ by $\mathcal{P}^+ = \{x > 0 : x \in \mathcal{P}\}$ and $\mathcal{D}^+ = \{(y, s) \in \mathcal{D} : s > 0\}$, respectively. For any $\mu > 0$, consider the pair of problems

$$\min\ \{c^T x + \mu F(x)\ :\ Ax = b, x > 0\}\ \text{ and }\ \min\ \{-b^T y + \mu F(s)\ :\ A^T y + s = c, s > 0\}, \tag{1.2}$$

where $F(\cdot)$ is a self-dual strictly convex barrier function meaning that $F(u) \to \infty$ as $u \to 0$ for $u \in \mathbb{R}^n$. Unless otherwise stated the barrier function will be the *logarithmic barrier* function $F(u) = -\sum_{i=1}^{n} \log(u_i)$. In Chapter 7, we will consider a different barrier called the *volumetric barrier* function.

For any $\mu > 0$, the following nonlinear problems, (see e.g. Roos et al. (2006)) with the primal and dual logarithmic barrier functions,

$$\begin{aligned} \min \quad & c^T x - \mu \sum_{i=1}^{n} \log x_i \\ \text{s.t.} \quad & Ax = b \end{aligned} \tag{1.3}$$

and

$$\begin{aligned} \min \quad & -b^T y - \mu \sum_{i=1}^{n} \log s_i \\ \text{s.t.} \quad & A^T y + s = c \end{aligned} \tag{1.4}$$

have the optimality conditions:

$$\begin{aligned} Ax &= b, \quad x \geq 0 \\ A^T y + s &= c, \quad s \geq 0 \\ xs &= \mu e, \end{aligned} \tag{1.5}$$

where uv denotes the Hadamard product $[u_1 v_1, \ldots, u_n v_n]^T$ for $u, v \in \mathbb{R}^n$ and $e = [1, \ldots, 1]^T$ is the all-one vector.

Notation: For a vector $u \in \mathbb{R}^n$, $U := diag(u)$ will be the diagonal matrix whose entries consist of those of u.

Theorem 1.2.1. Roos et al. (2006) *Suppose $\mathcal{P}$ and $\mathcal{D}$ satisfy the interior-point condition, i.e., there exist $x, s > 0$ satisfying (1.1). Then, for any $\mu > 0$, the system (1.5) has a unique solution $(x(\mu), y(\mu), s(\mu))$, where $x(\mu) > 0$ and $s(\mu) > 0$.*

Definition 1.2.2. *The projections $\{x(\mu) : \mu > 0\} \subset \mathcal{P}$ and $\{(y(\mu), s(\mu)) : \mu > 0\} \subset \mathcal{D}$ are called primal and dual central paths, respectively. We also denote the primal-dual central path by $(x(\mu), y(\mu), s(\mu))$ for $\mu > 0$.*

The relevance of (1.3) and (1.4) to solve (1.1) comes from the following fact:

Theorem 1.2.3. Roos et al. (2006) *Suppose that* (1.1) *satisfy the interior-point condition. Then,*

1. *$(x(\mu), y(\mu), s(\mu))$ is a smooth analytic curve.*
2. *The duality gap on the central path is $c^T x(\mu) - b^T y(\mu) = x(\mu)^T s(\mu) = n\mu$.*
3. *As $\mu \to 0$, the central path $(x(\mu), y(\mu), s(\mu))$ converges to an optimal solution (x^*, y^*, s^*) of (1.1). The duality gap is zero at optimality of (1.1), and the limit point is a strictly complementary optimal solution, i.e., $x^* s^* = 0$ and $x^* + s^* > 0$.*

The main idea of path-following algorithms is the following. Since as $\mu \to 0$ the optimal solutions of (1.3) and (1.4) lead to an optimal solution for (1.1), we can compute strictly feasible solutions close to $x(\mu)$, $(s(\mu)$, or both), and then reduce the barrier parameter μ, and repeat the procedure until μ is small enough. Computing approximate solutions for (1.3) and (1.4) is done by Newton's method. Such algorithms are called IPMs.

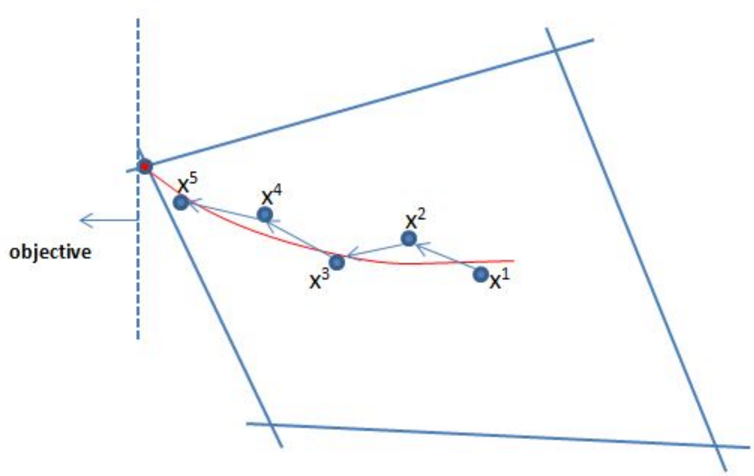

Figure 1.1: The points $x^1, \ldots, x^k, \ldots$ follow the central path and converge to an optimal solution.

In the next chapter, we review some properties of the central path and IPMs more closely.

Chapter 2

Path-following algorithms

Our goal in this chapter is to review in more detail the properties of primal-dual IPMs. We present two path-following algorithms for LO problems and review their basic properties. The first algorithm in Section 2.1 is a short-step primal-dual path-following algorithm Roos et al. (2006). The second algorithm in Section 2.2 is a predictor-corrector type primal-dual algorithm Mizuno et al. (1993). Both algorithms achieve the best iteration-complexity upper bound for IPMs as of today.

2.1 Short-step IPMs

We first review certain properties of a short-step primal-dual method. Given $(x,s) \in \mathcal{P}^+ \times \mathcal{D}^+$ and $\mu > 0$, we have the proximity measure $\delta(x,s,\mu) := \left\| \frac{xs}{\mu} - e \right\|_2$ measuring how close (x,s) is to the central path point $(x(\mu), y(\mu), s(\mu))$ that satisfies (1.5). Define the neighborhood $\mathcal{N}(\beta,\mu) = \{(x,s) \in \mathcal{P}^+ \times \mathcal{D}^+ : \delta(x,s,\mu) \leq \beta\}$. We define the β-neighborhood of the central path as $\mathcal{N}(\beta) := \bigcup_{\mu>0} \mathcal{N}(\beta,\mu)$. The following algorithm can be found in Roos et al. (2006).

A short-step path-following primal-dual algorithm:

Input:

An accuracy parameter $\epsilon > 0$;

a proximity parameter β, $0 \leq \beta < 1$;

a barrier update parameter θ, $0 < \theta < 1$;

an approximate solution $(x^0, s^0) \in \mathcal{P}^+ \times \mathcal{D}^+$ of (1.5) for some initial μ^0 such that $\delta(x^0, s^0, \mu^0) \leq \beta$ and $(x^0)^T s^0 = n\mu_0$;

begin $x := x^0$; $s := s^0$; $\mu := \mu^0$; **while** $n\mu \geq \epsilon$ **do**

begin

$x := x + \Delta x$;

$s := s + \Delta s$;

$\mu := (1 - \theta)\mu$;

end

end

For completeness, we include the linear system to solve in order to get the search directions $(\Delta x, \Delta s)$:

$$\begin{aligned} S\Delta x + X\Delta s &= \mu e - xs \\ A\Delta x &= 0 \\ A^T \Delta y + \Delta s &= 0 \end{aligned} \tag{2.1}$$

Next we define the input length L for an LO problem.

Definition 2.1.1. *Let the data A, b, c be integral. The input length L is defined as*

$$L = \sum_{i=1}^{m}(1+\lceil \log_2(1+|b_j|)\rceil) + \sum_{j=1}^{n}(1+\lceil \log_2(1+|c_j|)\rceil) + \sum_{\substack{i=1 \\ j=1}}^{m,n}(1+\lceil \log_2(1+|a_{ij}|)\rceil) \quad (2.2)$$

The complexity result for the short-step path-following algorithm is as follows:

Lemma 2.1.2. Roos et al. (2006)

1. *If* $\beta \leq \frac{1}{\sqrt{2}}$ *and* $\theta = \frac{1}{\sqrt{2n}}$, *the short-step path-following primal-dual algorithm requires at most* $\lceil \sqrt{2n} \log \frac{n\mu^0}{\epsilon} \rceil$ *iterations. The output is a primal dual pair* (x, s) *such that* $x^T s \leq \epsilon$.

2. *The accuracy* ϵ *needed to identify an exact optimal solution of the problems in* (1.1) *is* $\epsilon = \mathcal{O}(2^{-2L})$, *where* L *is the input length of the problem.*

2.2 Mizuno-Todd-Ye predictor-corrector algorithms

In this section, we review the properties of the Mizuno-Todd-Ye predictor-corrector (MTY predictor-corrector) algorithm Mizuno et al. (1993). Each iteration of the predictor-corrector algorithm consists of two steps, a predictor step and a corrector (or centrality) step. The search direction used by both steps at a given point in $u = (x, y, s) \in \mathcal{P}^+ \times \mathcal{D}^+$ is the unique solution of the following linear system of equations:

$$\begin{aligned} S\Delta x + X\Delta s &= \gamma\mu e - xs \\ A\Delta x &= 0 \\ A^T\Delta y + \Delta s &= 0. \end{aligned} \quad (2.3)$$

When $\gamma = 0$, the predictor search direction leads to the point for which the duality gap is zero. Hence if that point is feasible, it is optimal. However in general to maintain feasibility, a line search is performed. Another reason preventing a full Newton step is that the new point could be away from the central path with respect to a proximity measure. If (x, s) is on the central path, then the predictor search direction coincides with the tangent to the central path at that point. Hence intuitively if the curvature of the central path at that point is small, it should be possible to take a large step yielding a large reduction in the duality gap.

An iteration of the MTY predictor-corrector algorithm is as follows. Suppose that a constant $\beta = \frac{1}{4}$ is given. Given a point $u = (x, y, s)$ with normalized duality gap $\mu = \frac{x^T s}{n}$, suppose $\delta(x, s, \mu) \le \beta$. The algorithm generates

$$u^+ = (x^+, y^+, s^+) = (x, y, s) + \theta(\Delta x, \Delta y, \Delta s)$$

as follows. It first moves along the direction for which $\gamma = 0$, until it hits the boundary of the enlarged neighborhood $\delta(x^+, s^+, \mu^+) \le 2\beta$. In other words, compute the largest θ so that $\delta(x^+, s^+, \mu_g^+) \le 2\beta$ where the new duality gap $\mu_g^+ = \frac{{x^+}^T s^+}{n}$. Next, starting from the new point u^+, a new search direction with $\gamma = 1$ is computed. This search direction coincides with the search direction (2.1) corresponding to the point (x^+, y^+, s^+). The MTY predictor-corrector algorithm Mizuno et al. (1993) requires one corrector step with $\theta = 1$ to go back to the neighborhood where $\delta(x, s, \mu) \le \beta$.

Remark 2.2.1. *If the point (x, s) is on the central path with $xs = \mu e$, then the predictor step search direction (for $\gamma = 0$) $(\Delta x, \Delta y, \Delta s)$ in (2.3) are exactly the tangent directions $(\Delta x, \Delta s) = (-\dot{x}, -\dot{s})$. If $xs = w$ for $w \neq e$, then the search directions in (2.3) are the tangent directions for the equation system*

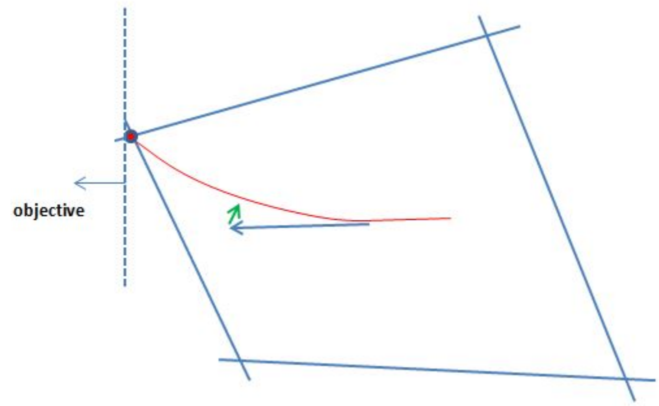

Figure 2.1: Illustration of a predictor-corrector algorithm.

$$\begin{aligned} Ax &= b, \quad x \geq 0 \\ A^T y + s &= c, \quad s \geq 0 \\ xs &= \mu w, \end{aligned} \tag{2.4}$$

which are the optimality conditions of the problems

$$\begin{array}{ll} \min & c^T x - \mu \sum_{i=1}^n w_i \log x_i \\ \text{s.t} & Ax = b \end{array} \tag{2.5}$$

and

$$\begin{array}{ll} \min & -b^T y - \mu \sum_{i=1}^n w_i \log s_i \\ \text{s.t} & A^T y + s = c. \end{array} \tag{2.6}$$

For any $w > 0$, $w \in \mathbb{R}^n$, *the path defined in* (2.4) *also converges to an optimal solution of* (1.1).

The following theorem presents the most important properties of the MTY predictor-corrector algorithm Mizuno et al. (1993); Monteiro and Tsuchiya (2005):

Theorem 2.2.2. (**Predictor step**)

Suppose that $(x, y, s) \in \mathcal{N}(\beta)$ *for some* $0 < \beta \leq \frac{1}{4}$ *with* $\frac{x^T s}{n} = \mu$. *For* $\gamma = 0$, *let the search directions be* $(\Delta x, \Delta y, \Delta s)$. *Let* θ *be the step length so that* $u^+ = (x^+, y^+, s^+) = (x, y, s) + \theta(\Delta x, \Delta y, \Delta s) \in \mathcal{N}(2\beta)$. *Then,*

1. $\mu^+ = (1-\theta)\mu$, *where* $\mu^+ = \dfrac{x^+ s^+}{n}$.

2. $$\left(\frac{x^+ s^+}{\mu^+} - e\right) = \left(\frac{xs}{\mu} - e\right) + \frac{\theta^2}{(1-\theta)} \frac{\Delta x \Delta s}{\mu} \tag{2.7}$$

3. *The step length* θ *satisfies* $\theta \geq \max\left\{\sqrt{\dfrac{\beta}{n}},\ \ 1 - \dfrac{\chi(\mu)}{\beta}\right\}$, *where* $\chi(\mu) := \dfrac{\|\Delta x \Delta s\|_2}{\mu}$. *Further, we have*

$$\theta \geq \frac{2}{1 + \sqrt{1 + \dfrac{4\chi(\mu)}{\beta}}}. \tag{2.8}$$

Theorem 2.2.3. (**Corrector step**)
Suppose that $(x, y, s) \in \mathcal{N}(2\beta)$. *For* $\gamma = 0$, *let the search directions be* $(\Delta x, \Delta y, \Delta s)$. *Then* $u^+ = (x^+, y^+, s^+) = (x, y, s) + (\Delta x, \Delta y, \Delta s) \in \mathcal{N}(\beta)$. *Moreover, the duality gap for* u^+ *is the same as* u.

The following theorem gives the complexity upper bound for the MTY predictor-corrector algorithm.

Theorem 2.2.4. *Let* $\beta = \frac{1}{4}$ *and* $(x_0, y_0, s_0) \in \mathcal{N}(\beta)$ *be given such that* $(x_0)^T s^0 = n\mu_0$ *for some* μ_0. *Then the MTY predictor-corrector algorithm will terminate in at most* $\mathcal{O}(\sqrt{n} \log \frac{n\mu_0}{\epsilon})$ *iterations with duality gap* $x^T s = c^T x - b^T y \leq \epsilon$.

Observe that if the error term χ in Theorem 2.2.2 is small, then we can choose a larger step length θ, hence get a larger reduction in the duality gap.

In the rest of this section, we present a variant of the MTY predictor-corrector algorithm developed by Stoer and Zhao (1993). Generally, both the algorithm of Stoer and Zhao (1993) and the MTY predictor-corrector algorithm use two nested neighborhoods $\mathcal{N}(\beta_0)$ and $\mathcal{N}(\beta_1)$ for $0 < \beta_0 < \beta_1 < 1$. The MTY predictor-corrector algorithm and the algorithm in Stoer and Zhao (1993) differ in the way the value of θ is determined. In the MTY predictor-corrector algorithm, we have $\beta_1 = 2\beta_0$ and θ is determined as being the largest number for which (x^+, s^+) stays within the enlarged neighborhood $\mathcal{N}(\beta_1)$. In the algorithm of Stoer and Zhao (1993), the value of θ is determined as the largest number for which $\left\| \frac{x^+ s^+}{\mu^+} - w \right\|_2 \leq \beta_1$, where $w = \frac{xs}{\mu}$. Then a constant number of pure centering steps are taken which will take the iterate back to the smaller neighborhood $\mathcal{N}(\beta_0)$ in such a way that the duality gap does not change. Both algorithms accelerate as they get close to optimality. For the MTY predictor-corrector algorithm, it is proved, see e.g. Potra (1994) that, μ^+ goes quadratically to zero. On the other hand, it is known that as $k \to \infty$, $\theta_k \to 1$ Stoer and Zhao (1993).

For the rest of the thesis, we will refer to the both algorithms as MTY predictor-corrector algorithm.

2.3 Polynomial iteration-complexity and local metric

In this section we highlight the role of the local norms in IPMs. In fact, it is possible to say that the polynomial iteration-complexity bound to solve LO problems is due to fact that Newton's method behaves nicely under the local Hessian norm. Our presentation here mostly follows that of Renegar (1987).

Consider the primal log-barrier problem in the form

$$\begin{array}{ll} \min & \frac{1}{\mu}c^T x - \sum_{i=1}^{n} \log x_i \\ \text{s.t} & Ax = b. \end{array} \tag{2.9}$$

Clearly the optimal solutions of (1.3) and (2.9) are the same. The reason for this rescaling is to make the Hessians free of the barrier parameter. First let's review Newton's method's quadratic convergence result for any general norm (see Renegar (1987) p:20).

Theorem 2.3.1. *Let $x(\mu)$ be the optimal solution of (2.9) and Δx be the Newton step to solve (2.9) so that $x^+ = x + \Delta x$. Then*

$$\|x^+ - x(\mu)\|_2 \le \|x - x(\mu)\|_2 \|H(x)^{-1}\|_2 \int_0^1 \|H(x + t(x(\mu) - x)) - H(x)\|_2 dt, \tag{2.10}$$

where $H(x)$ is the Hessian of the barrier function in (2.9).

Notice that if the norm used in (2.10) is Euclidean, then $\|H(x)^{-1}\|_2 = \|x\|_\infty^2$ and the convergence neighborhood parameter β will be affected by $\|x\|_\infty$. Under the local Hessian norm, Theorem 2.3.1 becomes the following.

Theorem 2.3.2. *Under the assumptions of Theorem 2.3.1, suppose that $\|x - x(\mu)\|_{H(x)} < 1$. Then*

$$\|x^+ - x(\mu)\|_{H(x)} \le \frac{\|x - x(\mu)\|_{H(x)}^2}{1 - \|x - x(\mu)\|_{H(x)}}. \tag{2.11}$$

Similarly we have the following:

Theorem 2.3.3. *If $\|\Delta x\|_{H(x)} \le \frac{1}{4}$, then*

$$\|x^+ - x(\mu)\|_{H(x)} \le \frac{3\|\Delta x\|_{H(x)}^2}{(1 - \|\Delta x\|_{H(x)})^3}. \tag{2.12}$$

Notice the simplicity of the bound in Theorem 2.3.3 compared to the generic bound (2.10). Now we are ready to introduce self-concordant barrier functions Nesterov and Nemirovskii (1994):

Definition 2.3.4. (Renegar (1987), p:23) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function with positive definite Hessian. Then f is said to be (strongly nondegenerate) self-concordant if*

- $\{z \in \mathbb{R}^n : \|z - x\|_{H(x)} \leq 1\} \subset \mathbb{R}^n_{++}$, *where $\mathbb{R}^n_{++}$ is the positive orthant.*
- *For all $x \in \mathbb{R}^n$, whenever $\|z - x\|_{H(x)} \leq 1$, one has*

$$1 - \|z - x\|_{H(x)} \leq \frac{\|v\|_{H(z)}}{\|v\|_{H(x)}} \leq \frac{1}{1 - \|z - x\|_{H(x)}}. \tag{2.13}$$

Note that logarithmic barrier function is self-concordant.

The following result is fundamental for showing the polynomial iteration-complexity of IPMs for LO problems. Suppose we have the system:

$$\begin{aligned} \min \quad & \tfrac{1}{\mu} c^T x + f(x) \\ \text{s.t.} \quad & Ax = b. \end{aligned} \tag{2.14}$$

where $f(x)$ is a (strongly nondegenerate) self-concordant function. Suppose $v_f := \sup_{x>0} \|\nabla f\|_{H(x)}$ is finite, where ∇f is the gradient with respect to the Hessian induced norm. Then,

Theorem 2.3.5. *The short-step IPMs algorithm (described as in* Renegar (1987), p:45*) has iteration-complexity*

$$\mathcal{O}(\sqrt{v_f} \log(v_f \mu_0 / \epsilon)),$$

where μ_0 is the initial barrier parameter and ϵ is the desired accuracy for the duality gap.

The algorithm of Renegar (1987), p:45 (see also Nesterov and Nemirovskii (1994)) is a primal short-step IPM for a general self-concordant function f with the corresponding v_f. It is well-known for example that v_f for the logarithmic barrier function is n, Nesterov and Nemirovskii (1994). Notice that for the logarithmic barrier function (1.3), the iteration-complexity upper bound in Theorem 2.3.5 matches the bound $\mathcal{O}(\sqrt{n}\log\frac{n\mu_0}{\epsilon})$ in Lemma 2.1.2 and Theorem 2.2.4.

Chapter 3

Curvature and IPMs

In Chapter 2, we indicated that IPMs follow the central path with Newton steps. Intuitively, this suggests that the sequence of points generated by the algorithm are somehow along the linear approximations of the central path and a central path with small curvature is easier to approximate with line segments yielding a lower number of Newton steps. Hence, if we can give an upper bound on the total curvature of the central path corresponding to an LO problem, this could also be used to bound the number of Newton iterations in IPMs. In addition, if we construct LO problems with large total curvature, such constructions would serve as examples of LO problems requiring a large number of iterations. In other words, it may be expected that

$$\#\text{ of iterations of Newton steps } = \Theta(\text{the curvature of the central path}).$$

In the following sections, we review several results that explore these ideas.

3.1 Geometric curvature of the central path

First we define the geometric curvature of a path (see Dedieu et al. (2005)). Intuitively, the curvature of a curve at a point is a measure of how far off it is from being a straight line around a neighborhood of that point. Let $h : [\alpha, \beta] \to \mathbb{R}^n$ be a C^2 map with non-zero derivative for any $t \in (\alpha, \beta)$. Denote the arc length by ℓ, where $\ell(t) = \int_\alpha^t \|\dot{h}(\tau)\|_2 d\tau$. The map ℓ establishes a one-to-one correspondence between the intervals $[\alpha, \beta]$ and $[\ell(\alpha), \ell(\beta)]$ and provides an arc length parametrization $h(\cdot)$. To $h(\cdot)$, there is an associated curve, called the Gauss curve, of unit length:

$$\text{For any } t \mapsto \ell, \text{ let } \gamma(t) = \frac{\dot{h}(t)}{\|\dot{h}(t)\|_2} = \frac{d}{d\ell}(h(\ell)).$$

The curvature at a point $h(\ell)$ is the second derivative with respect to the arc length parametrization, i.e.,

$$\Psi(l) = \frac{d}{d\ell}(\dot{h}(\ell)). \tag{3.1}$$

In terms of the original parameter t, it is written as

$$\Psi(t) = \frac{d}{dt}\left(\frac{\dot{h}(t)}{\|\dot{h}(t)\|_2}\right)\frac{1}{\|\dot{h}(t)\|_2}.$$

The total curvature K is the integral of the norm of the curvature vector, i.e.,

$$K = \int_0^{\ell(\beta)} \|\Psi(\ell)\|_2 d\ell = \int_\alpha^\beta \left\| \frac{d}{dt}\left(\frac{\dot{h}(t)}{\|\dot{h}(t)\|}\right)\right\|_2 dt. \tag{3.2}$$

The total curvature is independent of the initial parametrization t.

3.1.1 Upper bounds

Consider the dual problem and its feasible set,

$$\begin{aligned} \max \quad & b^T y & \\ \text{s.t.} \quad & A^T y + s & = c \\ & s & \geq 0. \end{aligned} \tag{3.3}$$

If we ignore the non negativity conditions $s_i \geq 0$ for each $i = 1, \ldots, n$, and allow either $s_i \geq 0$ or $s_i \leq 0$, we get 2^n possible polyhedra in the dual space corresponding to the sign configurations of s_i, $i = 1, \ldots, n$. Among such polyhedra, we consider only those that are bounded, i.e., only the polytopes. Given the matrix A and c, let $P(A, c)$ be the set of nonempty polytopes obtained this way. It is possible to show that the number of such polytopes is bounded above by $\binom{n-1}{m}$. This bound is achieved if the hyperplanes are in "generic" position: A hyperplane arrangement is called simple if any m hyperplanes intersect at a unique distinct point Deza and Xie (2007). If the hyperplane arrangement is simple, then $|P(A, c)| = \binom{n-1}{m}$.

Now fix an objective function $b^T y$ and for each bounded cell in the arrangement consider the central path corresponding to that bounded cell. For each bounded cell let $K(A, c; b)$ denote the total curvature of the corresponding central path. Using algebraic and integral geometry tools, Dedieu et al. Dedieu et al. (2005) proved that

$$\sum_{P \subset P(A,c)} K(A, c; b) \leq 2\pi(m-1)\binom{n-1}{m}. \tag{3.4}$$

Hence if the hyperplane arrangement is simple so that $|P(A, c)| = \binom{n-1}{m}$, we get an upper bound for the average total curvature as

$$\frac{\sum_{P \subset P(A,c)} K(A, c; b)}{|P(A, c)|} \leq 2\pi(m-1) = \mathcal{O}(m). \tag{3.5}$$

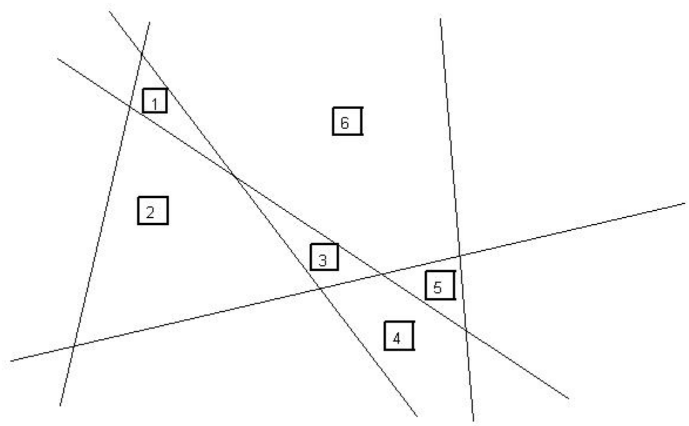

Figure 3.1: A simple arrangement of 5 hyperplanes in dimension 2.

The authors also conjectured that the worst case total curvature of a central path is $\mathcal{O}(m)$. The claim has been disproved by Deza et al. (2006).

In a recent paper De Loera et al. (2012), the authors consider the equation system

$$\begin{aligned} Ax &= b, \\ A^T y + s &= c, \\ xs &= \mu e \end{aligned} \tag{3.6}$$

without the nonnegativity conditions on x and s. Using algebraic geometry techniques, they prove a similar bound on the primal central path curve $x(\mu)$, which is the union of primal central paths resulting from system (3.6) (see Proposition 17 and Theorem 18 in De Loera et al. (2012)):

Theorem 3.1.1. *The total curvature K of the primal central path is bounded by*

$$K \leq 2\pi(n-m-1)\binom{n-1}{m-1}. \tag{3.7}$$

3.1.2 Lower bounds

Another approach to the curvature of the central path is to use it to construct worst case examples of LO problems.

Klee-Minty constructions:

- In Deza et al. (2006), the authors consider the following Klee-Minty cube variant where the m dimensional unit cube $[0,1]^m$ is tilted by a factor ρ. Following the dual problem formulation in (1.1), we have

$$\begin{array}{lrcll} \max & & -y_m & & \\ \text{s.t} & 0 \leq & y_1 & \leq 1 & \\ & \rho y_k \leq & y_k & \leq 1 - \rho y_{k-1} & \text{for } k = 2 \dots m. \end{array} \tag{3.8}$$

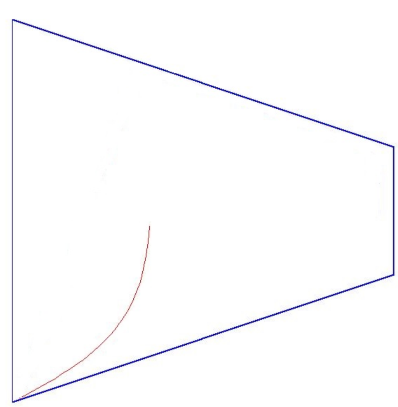

Figure 3.2: The central path in the non-redundant KM cube for $m = 2$.

Certain variants of the simplex method take $2^m - 1$ pivots to solve this problem. The simplex path for these variants starts from $(0, ..., 0, 1)^T$, it visits all the vertices ordered by the decreasing value of the last coordinate y_m until reaching the optimal point, which is the origin.

By adding redundant constraints at the same distance d parallel to the faces of the KM cube that includes the origin, they perturb the central path so that starting from the analytic center of the KM polytope towards the optimal point, it visits a small neighborhood of all the vertices of the cube.

$$
\begin{array}{llll}
\max & -y_m & & \\
\text{s.t} & 0 \leq \; y_1 & \leq 1 & \\
& \rho y_k \leq \; y_k & \leq 1 - \rho y_{k-1} & \text{for } k = 2 \ldots m. \\
& 0 & \leq d + y_1 & \text{repeated } h_1 \text{ times} \\
& \rho y_1 & \leq d + y_2 & \text{repeated } h_2 \text{ times} \\
& & \vdots & \\
& \rho y_{m-1} & \leq d + y_m & \text{repeated } h_m \text{ times.}
\end{array}
\tag{3.9}
$$

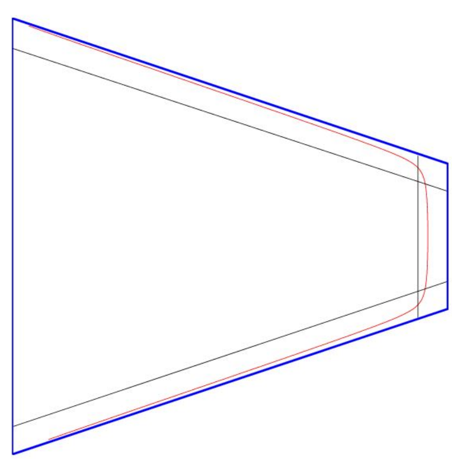

Figure 3.3: The central path after adding the redundant constraints to the KM cube for $m = 2$.

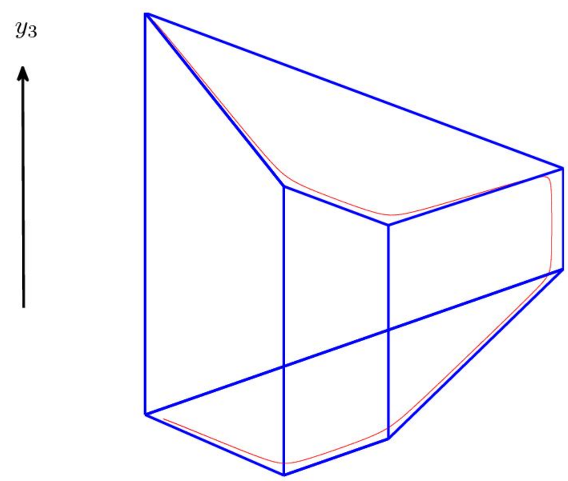

Figure 3.4: The central path after adding the redundant constraints to the KM cube for $m = 3$.

The number of redundant constraints needed to add in the construction is exponential in m, i.e., $h_1 + \ldots + h_m = \mathcal{O}(2^{6m}m^2)$ with an input length $L = \mathcal{O}(2^{6m}m^3)$. The distance of the redundant constraints to the KM hyperplanes is $d \geq m2^{m+1}$. Since the central path visits all the vertices, there must be $2^m - 2$ sharp turns, and a path-following IPMs needs at least $\Omega(2^m)$ number of iterations. In terms of the number of inequalities n, this gives a lower bound iteration-complexity of $\Omega((\frac{n}{\log^2 n})^{1/6})$.

- Note that in the above KM cube construction, if the central path is bent with a smaller number of redundant constraints, it will give a smaller n and hence will yield a relatively higher lower bound worst case iteration-complexity. In Deza et al. (2008a), the authors reduce the number of redundant constraints by allowing the distances d to vary. By placing the constraints at the distances $d = (m2^{m+4}, m2^{m+3}, \ldots, m2^5)$, they reduce the number of redundant constraints to $h_1 + \ldots + h_m = \mathcal{O}(2^{2m}m^3)$. This leads to the lower bound $\Omega((\frac{n}{\log^3 n})^{1/2})$ in the number iterations. By computing the iteration-complexity upper bound for this construction as $\mathcal{O}(\sqrt{n}\log n)$, the gap between upper and lower bound for complexities is reduced to a factor of $(\log n)^{5/2}$.

- Note that in the two previous KM constructions, the redundant constraints are placed parallel to the KM hyperplanes. It turns out Nematollahi and Terlaky (2008b) that placing the redundant constraints parallel to the coordinate axes is more efficient in the sense that less redundancy is required to bend the central path with $2^m - 2$ sharp turns. The distances of the constraints are still allowed to vary. More precisely, this formulation is

$$
\begin{array}{llll}
\max & -y_m & & \\
\text{s.t} & 0 \leq y_1 & \leq 1 & \\
& \rho y_k \leq y_k & \leq 1 - \rho y_{k-1} & \text{for } k = 2 \ldots m. \\
& 0 & \leq d_1 + y_1 & \text{repeated } h_1 \text{ times} \\
& 0 & \leq d_2 + y_2 & \text{repeated } h_2 \text{ times} \\
& & \vdots & \\
& 0 & \leq d_m + y_m & \text{repeated } h_m \text{ times .}
\end{array}
\tag{3.10}
$$

Here $d \cong (2^{m-1}, 2^{m-2}, \ldots, 2, 0)$. The number of added constraints is $h_1 + \ldots + h_m = \mathcal{O}(m2^{2m})$. This construction gives $\Omega(\frac{\sqrt{n}}{\sqrt{\log n}})$ while the iteration-complexity upper bound is $\mathcal{O}(\sqrt{n}\log n)$, and this last construction, to date, yields the smallest gap between the upper and lower bound complexities..

- Note that in all these KM constructions, highly redundant constraints are used. A natural question to ask is how curly the central path could be, if only those constraints are used which touch the feasible set. In Nematollahi and Terlaky (2008a) using a related KM construction, it has been shown that the same number $2^m - 2$

of sharp turns are obtained by making the distances set to zero. In other words, all of the redundant constraints touch the feasible region. For this construction, the number of constraints $h_1 + \ldots + h_m = \mathcal{O}(2^{m^2})$ which is exponentially higher than that of the previous KM example.

Remark 3.1.2. *From a practical point of view, it could be argued that many optimization solvers have pre-processing ability which would eliminate redundant constraints. However, this KM construction shows that in LO problems with less redundancy, the iteration-complexity could still be high.*

An example of a central path with curvature $\boldsymbol{\Omega}(\mathbf{n})$:

Another concrete case of an LO problem with a central path having a large curvature is developed in Deza et al. (2008b). Consider a polytope P in $\mathbb{R}^2$ defined by n inequalities as follows:

$$y_2 \leq 1,\ y_1 \leq \frac{y_2}{10} + \frac{1}{2}, \quad -y_1 \leq \frac{y_2}{3} + \frac{1}{3} \quad \text{and} \quad (-1)^i y_1 \leq \frac{10^{i-2} y_2}{11} + \frac{5}{11} - \frac{10^{-4}}{n}\frac{i}{n},\ i = 4, \ldots, n.$$

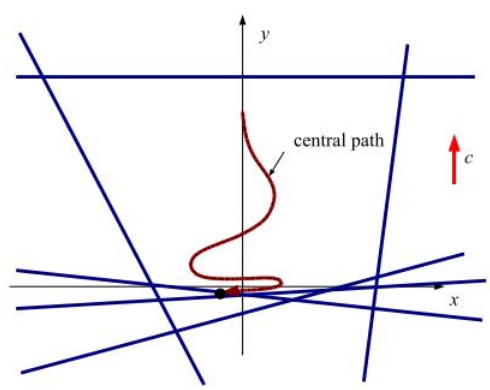

Figure 3.5: The polytope and its central path, picture by Deza et al. 2008.

The total curvature of the central path is asymptotically $\Omega(n)$. More precisely,

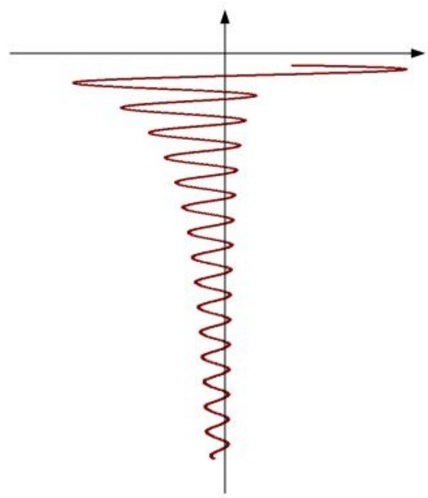

Figure 3.6: The central path with y_2 on logarithmic scale, picture by, Deza et al. 2008.

Theorem 3.1.3. Deza et al. (2008b) *The total curvature* K_n *of the central path of* $\min\{y_2 : (y_1, y_2) \in P\}$ *satisfies*

$$\liminf_{n\to\infty} \frac{K_n}{n} \geq \pi. \tag{3.11}$$

From Theorem 3.1.3, we see that the smallest possible upper bound for the curvature of the central path could be $\mathcal{O}(n)$. This is conjectured in Deza et al. (2008b). Recently, the conjecture has been disproved by Allamigeon et al. (2014). In fact, they showed that there exist LO problem instances (A, b, c) with total curvature $K = \Omega(2^n)$.

3.2 The Sonnevend curvature of the central path

Sonnevend's curvature is closely related to the iteration-complexity of the variant of MTY predictor-corrector algorithm which was introduced in Sonnevend et al. (1991). Let $\kappa(\mu) = \|\mu\dot{x}\dot{s}\|_2^{1/2}$. Stoer and Zhao (1993) proved that their algorithm has a iteration-complexity bound, which can be expressed in terms of $\kappa(\mu)$.

Theorem 3.2.1. *Let $\nu > 0$ be a constant with $\kappa(\mu) \geq \nu$ on $[\mu_1, \mu_1]$ and $\mathbf{N}$ be the number of iterations of Algorithm 2.1* Stoer and Zhao (1993) *to reduce the barrier parameter from μ_1 to μ_0. Then*

$$C_3 \int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu - 1 \leq \mathbf{N} \leq C_1 \int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu + C_2 \log\left(\frac{\mu_1}{\mu_0}\right) + 2 \tag{3.12}$$

for some "universal" constants C_1 and C_2 that depend only on the neighborhood of the central path. Moreover the constant C_3 depends on ν as well as the neighborhood of the central path.

The following proposition states the basic properties of Sonnevend's curvature.

Proposition 3.2.2. Sonnevend et al. (1991); Zhao (2010) *The following holds for the central path.*

1. *We have $\kappa(\mu) = \left\| \frac{\mu \dot{s}(\mu)}{s(\mu)} - \left(\frac{\mu \dot{s}(\mu)}{s(\mu)}\right)^2 \right\|_2^{\frac{1}{2}}$.*
2. *We have $\frac{\mu \dot{s}(\mu)}{s(\mu)} = Me$, where $M(\mu) = S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1}$ is the projection matrix. For a bounded dual feasible set, we have $\frac{\mu \dot{s}(\mu)}{s(\mu)} \to 0$ as $\mu \to \infty$.*
3. *We have $\left\| \frac{\mu \dot{s}(\mu)}{s(\mu)} \right\|_2 \leq \sqrt{n}$ and $\kappa(\mu) \leq \sqrt{n}$ implying that*

$$\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu = \mathcal{O}\left(\sqrt{n} \log\left(\frac{\mu_1}{\mu_0}\right)\right).$$

Next we show that the term $\frac{\kappa(\mu)}{\mu}$ can be expressed as a local curvature in a specific sense. Recall that given the Euclidean inner product $\langle \cdot \rangle$, any positive definite matrix $H \in \mathbb{R}^{n \times n}$ induces a new inner product $\langle \cdot \rangle_H$ as follows.

$$\langle u, v \rangle_H = \langle u, Hv \rangle = u^T H v \text{ for } u, v \in \mathbb{R}^n$$

Let the Hessian of the primal and dual logarithmic barrier functions $-\sum_{i=1}^{n}\log(x_i)$ and $-\sum_{i=1}^{n}\log(s_i)$ be $H(x)$ and $H(s)$, respectively.

Theorem 3.2.3. Ohara and Tsuchiya (2007) *We have*

$$\left(\frac{\kappa(\mu)}{\mu}\right)^2 = \frac{1}{2}\sqrt{\|\ddot{x}\|^2_{H(x)} + \|\ddot{s}\|^2_{H(s)}} \tag{3.13}$$

so that

$$\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu = \frac{1}{\sqrt{2}} \int_{\mu_0}^{\mu_1} (\|\ddot{x}\|^2_{H(x)} + \|\ddot{s}\|^2_{H(s)})^{1/4} d\mu \tag{3.14}$$

Theorem 3.2.3 shows that the iteration-complexity of the MTY predictor-corrector algorithm can indeed be interpreted as a curvature with respect to the Hessian norm induced by the logarithmic barrier function.

Monteiro and Tsuchiya (2008) proved that, as $\mu_0 \to 0$ and $\mu_1 \to \infty$, $\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu$ admits an upper bound expression which involves a condition number depending only on A. This condition number is defined as

$$\overline{\chi}_A := \sup_D \{\|A^T(ADA^T)^{-1}AD\|_2\}, \tag{3.15}$$

where D ranges over the set of positive diagonal matrices. It is known that (Vavasis and Ye (1996), Lemma 24), $\log(\overline{\chi}_A) = \mathcal{O}(L_A)$, where L_A is the input bit length of A when the matrix has all integer entries. Then we have the following bound for Sonnevend's curvature.

Theorem 3.2.4. Monteiro and Tsuchiya (2008) *We have*

$$\int_0^{\infty} \frac{\kappa(\mu)}{\mu} d\mu = \mathcal{O}(n^{3.5}\log(n + \overline{\chi}_A)).$$

Note that if we rescale our primal and dual LO problems with a positive diagonal matrix as

$$\begin{aligned} \min \quad & Dc^T x \\ \text{s.t.} \quad & ADx = b \\ & x \geq 0, \end{aligned} \tag{3.16}$$

$$\begin{aligned} \max \quad & b^T y \\ \text{s.t.} \quad & DA^T y + s = Dc \\ & s \geq 0, \end{aligned} \tag{3.17}$$

then the rescaled central path becomes $\overline{x}(\mu), \overline{y}(\mu), \overline{s}(\mu) = (D^{-1}x(\mu), y(\mu), Ds(\mu))$. It is easy to see that $\kappa(\mu)$ is scaling independent, while $\overline{\chi}_A$ is not. This gives a possibility of reducing the bound $\mathcal{O}(n^{3.5}\log(n + \overline{\chi}_A)$. Let $\overline{\chi}_A^* := \inf\{\overline{\chi}_{AD})\}$ where D ranges over all positive diagonal matrices. Then the bound in Theorem 3.2.4 becomes

$$\int_0^\infty \frac{\kappa(\mu)}{\mu} d\mu = \mathcal{O}(n^{3.5}\log(n + \overline{\chi}_A^*)). \tag{3.18}$$

In Monteiro and Tsuchiya (2005), the authors also show that $\overline{\chi}_A^*$ and $\overline{\chi}_A$ may have arbitrarily different orders of magnitude.

Chapter 4

A Klee-Walkup type result for Sonnevend's curvature

4.1 Overview

In this chapter, we prove that in order to establish a polynomial upper bound for the total Sonnevend curvature of the central path, it is sufficient to consider the case when $n = 2m$. This also implies that analyzing the worst-case behavior for any size of LO problem can be done simply by considering the case of $n = 2m$. As a by-product, our construction yields an asymptotically $\Omega(n)$ worst-case lower bound for Sonnevend's curvature. Our research is motivated by the work of Deza et al.(2008) for the geometric curvature of the central path, which is analogous to the Klee-Walkup result for the diameter of a polytope.

The idea of using a sequence of polytopes whose size and dimension increase by one was first used by Klee and Walkup (1967) in the context of the diameter of a polytope. The diameter of a polytope is the maximum of the shortest edge path's over all pairs of vertices. A lower bound in the worst-case for the diameter of a polytope implies the same lower bound for the iteration-complexity of any simplex type algorithm. In Klee and Walkup (1967), it is shown that proving an upper bound for the diameter of a polytope for general (m, n) reduces to the case of $(m, 2m)$. From an optimization perspective, it is interesting to note the analogies between the diameter of a polytope, the geometric, and the Sonnevend curvature of the central path. Moreover, this similarity suggests that the most "difficult" LO problems also occur when $n = 2m$.

The main result we obtain in this chapter for the Sonnevend curvature $\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu$ can be described as follows. Starting with an LO problem of size (m, n) with a bounded dual feasible set, we give a new LO problem whose size is $(m + 1, n + 1)$. The Sonnevend curvature for the latter is greater than that of the former by a constant independent of the problem data. Starting with a LO problem of size (m, n), and by continuing this process, we get an LO problem with size $(\overline{m}, 2\overline{m})$ whose curvature is greater than that of the original problem. This implies that in order to prove an upper bound for the Sonnevend curvature of the central path, it is sufficient to consider only the case when $n = 2m$.

Our work is motivated by the paper of Deza et al. (2009). In that paper, the authors construct a sequence of polytopes whose central path approximates that of the previous one. Furthermore, it is shown that total geometric curvature of the central path increases by a constant. In this thesis, we use the very same construction for the case of $n > 2m$. Hence, for the aforementioned construction, it can be concluded that Sonnevend's curvature and the geometric curvature of the central path have similar behavior. In Sonnevend et al. (1991), the authors use a different construction, which gives rise to the lower bound of $\Omega(n)$ for Sonnevend's curvature asymptotically. Our main result implies a bound which also achieves this worst-case lower bound.

Theorem 3.2.4 shows that the Sonnevend curvature admits an upper bound independent of both $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. In light of this fact, we make the following definition.

Definition 4.1.1. *Given $A \in \mathbb{R}^{m\times n}$, define*

$$\Lambda(m,n,A) = \sup\left\{ \int_0^\infty \frac{\kappa(\mu)}{\mu} d\mu : b \in \mathbb{R}^m,\ c \in \mathbb{R}^n \right\}.$$

4.2 Main construction

4.2.1 Embedding the central path

In this section, we introduce the construction used in Deza et al. (2009). First assume that $n > 2m$. We will later reduce the case $m < n < 2m$ to this case. Consider the LO problem

$$\max\{b^T y : y \in P\}, \text{ where } P = \{y \in \mathbb{R}^m : A^T y \le c\} \text{ is a polytope.} \tag{4.1}$$

Without loss of generality, we may assume that:

(A1) The analytic center y^* of P is the origin, and

(A2) $c = e$ where e is the all-one vector.

First, we can always shift a general polytope P so that assumption (A1) is satisfied. Since $\kappa(\mu)$ only depends on μ and the derivatives $\dot{x}$ and $\dot{s}$, this transformation would not change the Sonnevend curvature. Note that $y^* = 0$ being an interior point in P implies that $c > 0$. Second, if we rescale our LO problem with a positive diagonal matrix as given in (3.16) and (3.17), then the rescaled central path becomes $(\overline{x}(\mu), \overline{y}(\mu), \overline{s}(\mu)) = \big(D^{-1}x(\mu), y(\mu), Ds(\mu)\big)$ implying that $\kappa(\mu)$ does not change. Since $c > 0$ by assumption, by choosing D with $De = c^{-1}$, we can make $c = e$.

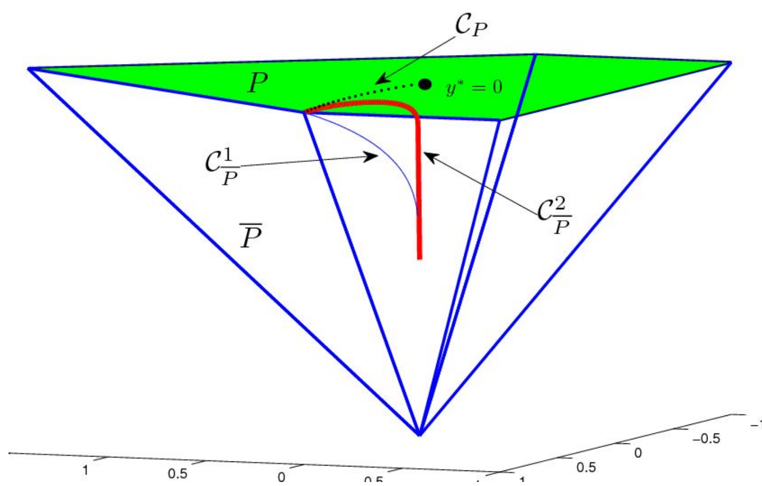

Figure 4.1: The dotted path $\mathcal{C}_P$ is the central path of the original polytope P. The figure shows how the central path $\mathcal{C}_{\overline{P}}$ is changing with θ. A smaller θ_1 leads to the path $\mathcal{C}_{\overline{P}}^1$, while $\mathcal{C}_{\overline{P}}^2$ results from $\theta_2 >> \theta_1$.

We now associate problem (4.1) with a sequence of LO problems parameterized by $\theta > 0$ as follows:

$$\max \quad b^T y + \theta z$$

$$\begin{bmatrix} A^T & -e_{n\times 1} \\ 0_{1\times m} & 1 \end{bmatrix} \begin{bmatrix} \overline{y} \\ z \end{bmatrix} + \begin{bmatrix} \overline{s} \\ \overline{s}_{n+1} \end{bmatrix} = \begin{bmatrix} 0_{n\times 1} \\ 1 \end{bmatrix} \qquad (4.2)$$

$$\overline{s}, \overline{s}_{n+1} \geq 0.$$

The feasible set for the problem (4.2) can be written as $\overline{P} = \{A^T\overline{y} \leq ze,\ z \leq 1\}$. Let $\overline{A} = \begin{bmatrix} A & 0_{m\times 1} \\ -e_{1\times n} & 1 \end{bmatrix}$. The associated central path equations for (4.2) are

$$A^T \frac{\overline{y}(\mu)}{z(\mu)} + \frac{\overline{s}(\mu)}{z(\mu)} = e, \quad A\overline{s}(\mu)^{-1} = \frac{b}{\mu}, \qquad (4.3)$$

$$\frac{1}{\overline{s}_{n+1}(\mu)} = \frac{1}{1-z(\mu)} = \sum_{i=1}^{n} \frac{1}{\overline{s}_i(\mu)} + \frac{\theta}{\mu}. \tag{4.4}$$

Note that $\overline{y}$, $\overline{s}$ and $\overline{z}$ in (4.3) and (4.4) are functions of both μ and θ. We will sometimes drop θ or μ, when no confusion arises. We denote the central path of P and $\overline{P}$ by $\mathcal{C}_P$ and $\mathcal{C}_{\overline{P}}$, respectively.

Intuitively a large θ should force $z \cong 1$ in such a way that, the central path $\mathcal{C}_{\overline{P}}$ first follows an almost straight line from the analytic center to the face $P \times \{1\}$ and then stays close to the central path $\mathcal{C}_P$. The following proposition, first proved in Deza et al. (2009), shows that this is indeed the case.

Proposition 4.2.1. *Let $[\mu_0, \mu_1]$ be a fixed interval. Then, as $\theta \to \infty$, on $[\mu_0, \mu_1]$ we have,*

1. *$z(\mu) \to 1$ and $\overline{y}(\mu) \to y(\mu)$ uniformly;*

2. *$\overline{s}_{n+1}(\mu) \to 0$ and $\overline{s}(\mu) \to s(\mu)$ uniformly.*

Proof. Claim *1.* is the same as Proposition 2.1 in Deza et al. (2009) (see also the remark following it). Claim *2.* follows from the first part since $\overline{s}_{n+1}(\mu) = 1 - z(\mu)$ and $\overline{s}(\mu) = z(\mu) - A^T\overline{y}(\mu)$. □

The following proposition shows that if $z(\mu)$ in (4.3) and (4.4) is known, then $\overline{y}(\mu)$ is completely determined by the central path $\mathcal{C}_P$.

Proposition 4.2.2. *Let $z(\mu)$ satisfy the central path equations (4.3) and (4.4). Then*

$$\overline{y}(\mu) = z(\mu) y\left(\frac{\mu}{z(\mu)}\right) \quad \text{and} \quad \overline{s}(\mu) = z(\mu) s\left(\frac{\mu}{z(\mu)}\right).$$

Proof. Direct substitution into (4.3) shows that $\overline{y}(\mu) = z(\mu) y\left(\frac{\mu}{z(\mu)}\right)$ and $\overline{s}(\mu) = z(\mu) s\left(\frac{\mu}{z(\mu)}\right)$ satisfy the equations in (4.3), which are the central path equations for (4.1) with the choice of $\mu' = \frac{\mu}{z(\mu)}$. Since the solution is unique, the claim follows. □

Note that Proposition 4.2.1 and Proposition 4.2.2 show that for a fixed interval $[\mu_0, \mu_1]$, the parameter θ can be chosen large enough so that the central paths $\mathcal{C}_P$ and $\mathcal{C}_{\overline{P}}$ become close to each other on that interval. Hence, it is natural to expect that Sonnevend's curvature for $\mathcal{C}_P$ and $\mathcal{C}_{\overline{P}}$ on the same interval should have similar order of magnitudes.

Proposition 4.2.3. *Let $\overline{\kappa}(\mu)$ correspond to the central path $\mathcal{C}_{\overline{P}}$. Then, on the fixed interval $[\mu_0, \mu_1]$, we have $\dot{\overline{s}}(\mu) \to \begin{bmatrix} \dot{s}(\mu) \\ 0 \end{bmatrix}$ uniformly as $\theta \to \infty$. Consequently, as $\theta \to \infty$, $\overline{\kappa}(\mu) \to \kappa(\mu)$ on $[\mu_0, \mu_1]$ uniformly as well.*

Proof. It is well-known, see Roos et al. (2006) e.g., that for system (1.5), we have $\dot{s} = \frac{1}{\mu} A^T (AS^{-2}A^T)^{-1} AS^{-1}e$. Now we calculate

$$U := \overline{A} \begin{bmatrix} \overline{S}^{-1} & 0 \\ 0 & s_{n+1}^{-1} \end{bmatrix} = \begin{bmatrix} A\overline{S}^{-1} & 0 \\ -\overline{s}^{-1} & (s_{n+1})^{-1} \end{bmatrix},$$

which gives

$$UU^T = \begin{bmatrix} A\overline{S}^{-2}A^T & -A\overline{s}^{-1} \\ (-A\overline{s}^{-1})^T & e^T\overline{s}^{-2} + \frac{1}{\overline{s}_{n+1}^2} \end{bmatrix}. \tag{4.5}$$

From the formula for the inverse of a block diagonal matrix, we obtain

$$(UU^T)^{-1} = \begin{bmatrix} (A\overline{S}^{-2}A^T)^{-1} + \frac{W_1}{r} & \frac{W_2}{r} \\ (\frac{W_2}{r})^T & \frac{1}{r} \end{bmatrix}, \tag{4.6}$$

where $r = e^T\overline{s}^{-2} + \frac{1}{\overline{s}_{n+1}^2} - (A\overline{s}^{-1})^T (A\overline{S}^{-2}A^T)^{-1} A\overline{s}^{-1}$, $W_2 = (A\overline{S}^{-2}A^T)^{-1}A\overline{s}^{-1}$, and $W_1 = W_2W_2^T$. Then, since $\overline{s} \to s$ as $\theta \to \infty$, it follows that the terms W_1 and W_2 converge to finite limits that are only determined by (4.1). Then, in terms of $\overline{s}_{n+1}$, we get $\frac{1}{r} = \mathcal{O}(\overline{s}_{n+1}^2)$. Thus, we conclude that

$$(UU^T)^{-1} = \begin{bmatrix} (A\overline{S}^{-2}A^T)^{-1} + \mathcal{O}(\overline{s}_{n+1}^2) & \mathcal{O}(\overline{s}_{n+1}^2) \\ \mathcal{O}(\overline{s}_{n+1}^2) & \mathcal{O}(\overline{s}_{n+1}^2) \end{bmatrix},$$

where $\mathcal{O}(.)$ should be understood to apply to the entries of a matrix, vector, or to a scalar depending on the context. Calculate

$$\begin{aligned} (UU^T)^{-1}\overline{A} \begin{bmatrix} \overline{s}^{-1} \\ \overline{s}_{n+1}^{-1} \end{bmatrix} &= (UU^T)^{-1} \begin{bmatrix} A\overline{s}^{-1} \\ -e^T\overline{s}^{-1} + \overline{s}_{n+1}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} (A\overline{S}^{-2}A^T)^{-1}A\overline{s}^{-1} + \mathcal{O}(\overline{s}_{n+1}) \\ \mathcal{O}(\overline{s}_{n+1}) \end{bmatrix}. \end{aligned} \tag{4.7}$$

Finally, from (4.7), we obtain

$$\overline{A}^T \ (UU^T)^{-1}\overline{A} \begin{bmatrix} \overline{s}^{-1} \\ \overline{s}_{n+1}^{-1} \end{bmatrix} = \begin{bmatrix} A^T(A\overline{S}^{-2}A^T)^{-1}A\overline{s}^{-1} + \mathcal{O}(\overline{s}_{n+1}) \\ \mathcal{O}(\overline{s}_{n+1}) \end{bmatrix}.$$

Taking the limit in θ, we get

$$\dot{\overline{s}}(\mu) = \frac{1}{\mu}\overline{A}^T(UU^T)^{-1}\overline{A} \begin{bmatrix} \overline{s}^{-1} \\ \overline{s}_{n+1}^{-1} \end{bmatrix} \to \begin{bmatrix} \frac{1}{\mu}A^T(AS^{-2}A^T)^{-1}As^{-1} \\ 0 \end{bmatrix} = \begin{bmatrix} \dot{s} \\ 0 \end{bmatrix}. \tag{4.8}$$

Since from Proposition 3.2.2, all the terms in $\overline{\kappa}(\mu)$ converge uniformly, we conclude that $\overline{\kappa}(\mu) \to \kappa(\mu)$ uniformly on $[\mu_0, \mu_1]$ as $\theta \to \infty$. □

Corollary 4.2.4. *On the fixed interval $[\mu_0, \mu_1]$, consider the Sonnevend curvature $\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu$ for the central path $\mathcal{C}_{\overline{P}}$. Then, for any $\epsilon > 0$, there is an LO problem of* *size* $(m+1, n+1)$ *with the Sonnevend curvature*

$$\int_{\mu_0}^{\mu_1} \frac{\overline{\kappa}(\mu)}{\mu} d\mu \geq \int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu - \epsilon.$$

Proof. By Proposition 4.2.3, we can choose a θ large enough so that $\kappa(\mu)$ and $\overline{\kappa}(\mu)$ is arbitrarily close to each other on $[\mu_0, \mu_1]$. Hence the claim follows. □

4.2.2 Constant increase of Sonnevend's curvature

We proved that on a fixed interval $[\mu_0, \mu_1]$, one can always make the Sonnevend curvature of $\mathcal{C}_{\overline{P}}$ and $\mathcal{C}_P$ arbitrarily close to each other. In the sequel, we will further show that there exists an interval $[\mu_1, \mu_2]$ such that while Sonnevend's curvature of $\mathcal{C}_P$ stays small on $[\mu_1, \mu_2]$, it can be made as large as a constant for $\mathcal{C}_{\overline{P}}$ on the same interval by increasing θ. To this end, the following proposition provides important tools. First, we need some special notation.

Notation: Let $\Delta : \mathbb{R}^2 \to \mathbb{R}$ be a function such that $\Delta(\alpha_1, \alpha_2)$ converges uniformly in α_2 to 0 as $\alpha_1 \to \infty$. Then we will write $\Delta(\alpha_1, \alpha_2) = o(1)$ as $\alpha_1 \to \infty$, and write *the bound is uniform* in α_2.

To display the dependence on θ, in the sequel we write the relevant quantities as functions of both μ and θ.

Proposition 4.2.5. *As* $\mu \to \infty$ *one has,*

1. $\overline{s}_i(\mu, \theta) - z(\mu, \theta) = o(1)$ *for* $i = 1, \ldots, n$,
2. $z(\mu, \theta) > \frac{1}{2}$, *and*
3. $\frac{\mu\dot{\overline{s_i}}(\mu, \theta)}{\overline{s_i}(\mu, \theta)} - \frac{\mu\dot{z}(\mu, \theta)}{z(\mu, \theta)} = o(1)$ *for* $i = 1, \ldots, n$.

Moreover, in statements 1. and 3., the bound is uniform in θ.

Proof.

1. From Proposition 4.2.2, we have $\overline{s}(\mu,\theta) = z(\mu,\theta)\left(e - A^T y(\frac{\mu}{z(\mu,\theta)})\right)$. Since by assumption, the analytic center of P is $y^* = 0$, we have $y(\mu) \to 0$ as $\mu \to \infty$. This proves the claim.

2. Since the analytic center of P is $y^* = 0$, we conclude $s_i(\mu,\theta) \le n$ for large μ with $i = 1,\dots,n$. From (4.4) and Proposition 4.2.2, we have

$$\frac{1}{1 - z(\mu,\theta)} - \frac{1}{z(\mu,\theta)}\left(\sum_{i=1}^{n} \frac{1}{s_i(\frac{\mu}{z(\mu,\theta)})}\right) = \frac{\theta}{\mu} > 0,$$

which implies

$$z(\mu,\theta) > \frac{\displaystyle\sum_{i=1}^{n} \frac{1}{s_i(\frac{\mu}{z(\mu,\theta)})}}{1 + \displaystyle\sum_{i=1}^{n} \frac{1}{s_i(\frac{\mu}{z(\mu,\theta)})}}. \tag{4.9}$$

Since for large μ, $s_i(\mu,\theta) \le n,\ i = 1,\dots,n$, the inequality (4.9) yields $\sum_{i=1}^{n} \frac{1}{s_i(\mu,\theta)} \ge 1$ for large μ. Then from (4.9), and using the fact that $z(\mu,\theta) \le 1$, for large μ we obtain $z(\mu,\theta) > \frac{1}{2}$.

3. Differentiating the equation $\overline{s_i}(\mu,\theta) = z(\mu,\theta)s_i\left(\frac{\mu}{z(\mu,\theta)}\right)$, we can derive from Proposition 4.2.2 that,

$$\dot{\overline{s_i}}(\mu,\theta) = \dot{z}(\mu,\theta)s\left(\frac{\mu}{z(\mu,\theta)}\right) + z(\mu,\theta)\dot{s_i}\left(\frac{\mu}{z(\mu,\theta)}\right)\left(\frac{1}{z(\mu,\theta)} - \frac{\mu\dot{z}(\mu,\theta)}{z(\mu,\theta)^2}\right). \tag{4.10}$$

Using (4.10), we get

$$\frac{\mu\dot{\overline{s_i}}(\mu,\theta)}{\overline{s_i}(\mu,\theta)} = \frac{\mu\dot{z}(\mu,\theta)}{z(\mu,\theta)} + \frac{\mu\dot{s}_i\left(\frac{\mu}{z(\mu,\theta)}\right)}{s_i\left(\frac{\mu}{z(\mu,\theta)}\right)}\left(\frac{1}{z(\mu,\theta)} - \frac{\mu\dot{z}(\mu,\theta)}{z(\mu,\theta)^2}\right). \tag{4.11}$$

Proposition 3.2.2 part *2.* implies that $\frac{\mu\dot{s}_i\left(\frac{\mu}{z(\mu,\theta)}\right)}{s_i\left(\frac{\mu}{z(\mu,\theta)}\right)} \to 0$ as $\mu \to \infty$. Further, Proposition 3.2.2 part *3.* implies that

$$\left|\frac{\mu\dot{z}(\mu,\theta)}{1-z(\mu,\theta)}\right| = \left|\frac{\mu\dot{\overline{s}}_{n+1}(\mu,\theta)}{\overline{s}_{n+1}(\mu,\theta)}\right| \leq \sqrt{n},$$

which further implies that

$$\left|\frac{\mu\dot{z}(\mu,\theta)}{z(\mu,\theta)^2}\right| \leq \sqrt{n}\frac{(1-z(\mu,\theta))}{z(\mu,\theta)^2}.$$

The bound $\frac{1}{2} < z(\mu,\theta) \leq 1$ from part *2.* implies that

$$\left|\frac{\mu\dot{z}(\mu,\theta)}{z(\mu,\theta)^2}\right| \leq \sqrt{n}\frac{(1-z(\mu,\theta))}{z(\mu,\theta)^2} \leq 2\sqrt{n},$$

which yields

$$\left|\left(\frac{1}{z(\mu,\theta)} - \frac{\mu\dot{z}(\mu,\theta)}{z(\mu,\theta)^2}\right)\right| \leq 2 + 2\sqrt{n}.$$

Hence we conclude from (4.11) that, $\frac{\mu\dot{\overline{s_i}}(\mu,\theta)}{\overline{s_i}(\mu,\theta)} \to \frac{\mu\dot{z}(\mu,\theta)}{z(\mu,\theta)}$ as $\mu \to \infty$. Note also that all the bounds come from problem (4.1), and therefore independent of θ. This proves that the bounds in statements *1.* and *3.* are uniform in θ. □

Now we are ready to present our main tool which leads to a constant increase in Sonnevend's curvature of $\mathcal{C}_{\overline{P}}$.

Lemma 4.2.6. *As $\mu \to \infty$, we have*

$$\frac{\mu\dot{\overline{s}}_{n+1}(\mu,\theta)}{\overline{s}_{n+1}(\mu,\theta)} = \frac{\frac{\theta}{\mu}}{\frac{\theta}{\mu} + \frac{n}{z(\mu,\theta)^2}} + o(1).$$

Moreover the bound is uniform in θ.

Proof. From (4.4), we have $\overline{s}_{n+1}(\mu,\theta) = \dfrac{1}{\frac{\theta}{\mu} + \sum_{i=1}^{n} \frac{1}{\overline{s}_i(\mu,\theta)}}$. Then one has

$$\log(\overline{s}_{n+1}(\mu,\theta)) = -\log\left(\frac{\theta}{\mu} + \sum_{i=1}^{n} \frac{1}{\overline{s}_i(\mu,\theta)}\right). \tag{4.12}$$

By differentiating (4.12) and multiplying by μ, we get

$$\frac{\mu\dot{\overline{s}}_{n+1}(\mu,\theta)}{\overline{s}_{n+1}(\mu,\theta)} = \frac{\frac{\theta}{\mu} + \sum_{i=1}^{n} \frac{\mu\dot{\overline{s}}_i(\mu,\theta)}{\overline{s}_i(\mu,\theta)}^2}{\frac{\theta}{\mu} + \sum_{i=1}^{n} \frac{1}{\overline{s}_i(\mu,\theta)}}. \tag{4.13}$$

Substituting $\overline{s}_{n+1}(\mu,\theta) = 1 - z(\mu,\theta)$ in (4.13) and using parts *1.* and *3.* of Proposition 4.2.5, as $\mu \to \infty$, we can write;

$$-\frac{\mu\dot{z}(\mu,\theta)}{1 - z(\mu,\theta)} = \frac{\frac{\theta}{\mu} + \frac{n\mu\dot{z}(\mu,\theta) + o(1)}{z(\mu,\theta)^2 + o(1)}}{\frac{\theta}{\mu} + \frac{n}{z(\mu,\theta)} + o(1)}. \tag{4.14}$$

Rearranging the terms in (4.14), we have

$$-\mu\dot{z}(\mu,\theta) = \frac{\frac{(1 - z(\mu,\theta))\theta}{\mu} + (1 - z(\mu,\theta))\left(\frac{n\mu\dot{z}(\mu,\theta) + o(1)}{z(\mu,\theta)^2 + o(1)}\right)}{\frac{\theta}{\mu} + \frac{n}{z(\mu,\theta)} + o(1)}. \tag{4.15}$$

To solve (4.15) for $\mu\dot{z}(\mu,\theta)$ explicitly, we first get

$$-\mu\dot{z}(\mu,\theta)\left(\frac{\theta}{\mu}+\frac{n}{z(\mu,\theta)}+\left(\frac{(1-z(\mu,\theta))n}{z(\mu,\theta)^2+o(1)}\right)+o(1)\right)$$

$$=\frac{(1-z(\mu,\theta))\theta}{\mu}+(1-z(\mu,\theta))o(1).$$

Finally we obtain,

$$-\frac{\mu\dot{z}(\mu,\theta)}{(1-z(\mu,\theta))} = \frac{\frac{\theta}{\mu}+o(1)}{\frac{\theta}{\mu}+\frac{n}{z(\mu,\theta)}+\left(\frac{(1-z(\mu,\theta))n}{z(\mu,\theta)^2+o(1)}\right)+o(1)}$$

$$=\frac{\frac{\theta}{\mu}+o(1)}{\frac{\theta}{\mu}+\frac{n}{z(\mu,\theta)^2}+o(1)}=\frac{\frac{\theta}{\mu}}{\frac{\theta}{\mu}+\frac{n}{z(\mu,\theta)^2}}+o(1),$$

which proves the claim. Moreover, since all the bounds come from Proposition 4.2.5, the bound is uniform in θ. This concludes the proof. $\square$

Corollary 4.2.7. *There exists a* $\tau \geq \frac{\sqrt{19}}{40}\log 2$ *such that*

$$\int_0^\infty \frac{\overline{\kappa}(\mu)}{\mu}d\mu \geq \int_0^\infty \frac{\kappa(\mu)}{\mu}d\mu+\tau.$$

Proof. Let $\epsilon>0$. Since $\int_0^\infty \frac{\kappa(\mu)}{\mu}d\mu$ is finite by Theorem 3.2.4, one can find a μ_0 and μ_1 such that $\int_0^{\mu_0} \frac{\kappa(\mu)}{\mu}d\mu \leq \epsilon$ and $\int_{\mu_1}^\infty \frac{\kappa(\mu)}{\mu}d\mu \leq \epsilon$. Note that from Lemma 4.2.6, we can also choose a μ_1 such that

$$\left| \frac{\mu \dot{\overline{s}}_{n+1}(\mu,\theta)}{\overline{s}_{n+1}(\mu,\theta)} - \frac{\frac{\theta}{\mu}}{\frac{\theta}{\mu} + \frac{n}{z(\mu,\theta)^2}} \right| \le \frac{1}{30} \tag{4.16}$$

for $\mu \ge \mu_1$ and for any $\theta > 0$.

Let $v = \int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu$. Having μ_1 chosen, we need to choose a θ' large enough so that both $\frac{\theta'}{\mu_1} > n$, and $\int_{\mu_0}^{\mu_1} \frac{\overline{\kappa}(\mu)}{\mu} d\mu \ge v - \epsilon$ are satisfied. Note that Corollary 4.2.4 implies that such a θ' exists.

Since by Proposition 4.2.5, part *2.*, we have $\frac{1}{2} \le z(\mu,\theta') \le 1$, it follows that $n \le \frac{n}{z(\mu,\theta')^2} \le 4n$ for $n \ge 2$. Since by assumption $\frac{\theta'}{\mu_1} > n$, there exist $\mu_2 > \mu_1$ such that $\frac{\theta'}{\mu_2} = n$. Then on $\mu \in [\mu_2, 2\mu_2]$, we have $\frac{n}{2} \le \frac{\theta'}{\mu} \le n$ and $n \le \frac{n}{z(\mu,\theta')^2} \le 4n$, which together implies that

$$\frac{1}{10} \le \frac{\frac{\theta'}{\mu}}{\frac{\theta'}{\mu} + \frac{n}{z(\mu,\theta')^2}} \le \frac{2}{3}. \tag{4.17}$$

Then for $\mu \in [\mu_2, 2\mu_2]$, (4.16) and (4.17) together imply $\frac{1}{20} \le \frac{\mu \dot{\overline{s}}_{n+1}(\mu,\theta')}{\overline{s}_{n+1}(\mu,\theta')} \le \frac{7}{10}$. Thus for $\mu \in [\mu_2, 2\mu_2]$, we obtain

$$\left| \left(\frac{\mu \dot{\overline{s}}_{n+1}(\mu,\theta')}{\overline{s}_{n+1}(\mu,\theta')} \right)^2 - \left(\frac{\mu \dot{\overline{s}}_{n+1}(\mu,\theta')}{\overline{s}_{n+1}(\mu,\theta')} \right) \right|^{\frac{1}{2}} \ge \frac{\sqrt{19}}{20}. \tag{4.18}$$

Hence, from (4.18) and Proposition 3.2.2, part *1.*, we obtain

$$\int_{\mu_2}^{2\mu_2} \frac{\overline{\kappa}(\mu)}{\mu} d\mu \ge \frac{\sqrt{19}}{20} \log 2.$$

Finally, we have

$$
\begin{aligned}
\int_0^\infty \frac{\overline{\kappa}(\mu)}{\mu} d\mu \geq \int_{\mu_0}^\infty \frac{\overline{\kappa}(\mu)}{\mu} d\mu \geq \int_{\mu_0}^{2\mu_2} \frac{\overline{\kappa}(\mu)}{\mu} d\mu \geq & \int_{\mu_0}^{\mu_1} \frac{\overline{\kappa}(\mu)}{\mu} d\mu + \int_{\mu_2}^{2\mu_2} \frac{\overline{\kappa}(\mu)}{\mu} d\mu \\
\geq & \ (v - \epsilon) + \frac{\sqrt{19}}{20} \log 2 \\
\geq & \int_{\mu_0}^\infty \frac{\kappa(\mu)}{\mu} d\mu - 2\epsilon + \frac{\sqrt{19}}{20} \log 2 \\
\geq & \int_0^\infty \frac{\kappa(\mu)}{\mu} d\mu - 3\epsilon + \frac{\sqrt{19}}{20} \log 2.
\end{aligned}
$$

The claim follows, since ϵ can be chosen arbitrarily small. □

Finally we deal with the case when $m < n < 2m$. In this case let $\hat{A} = [A \quad A]$, $\hat{b} = 2b$, and $\hat{c}^T = [c^T \quad c^T]$ so that $\hat{n} = 2n > 2m$. Then the central path is given as $\hat{x}(\mu)^T = [x(\mu)^T \quad [x(\mu)^T]$, $\hat{y}(\mu) = y(\mu)$ and $\hat{s}(\mu)^T = [s(\mu)^T \quad [s(\mu)^T]$. From these formulas that, one can easily deduce that, $\hat{\kappa}(\mu) = 2^{\frac{1}{4}}\kappa(\mu)$. Thus, since we have $\hat{n} > 2m$, our previous results apply.

The following corollary summarizes our findings in terms of $\Lambda(m, n, A)$, see Definition 4.1.1.

Corollary 4.2.8. *Let $A \in \mathbb{R}^{m \times n}$. Then there exists an $\overline{m}$, a matrix $\overline{A} \in \mathbb{R}^{\overline{m} \times 2\overline{m}}$, and a constant τ independent of problem data such that,*

- *If $n > 2m$, then $\Lambda(m, n, A) + (n - 2m)\tau \leq \Lambda(\overline{m}, 2\overline{m}, \overline{A})$, where $\overline{m} = n - m$.*
- *If $m < n < 2m$, then $\Lambda(m, n, A) + 2(n - m)\tau \leq 2^{\frac{1}{4}}\Lambda(\overline{m}, 2\overline{m}, \overline{A})$, where $\overline{m} = 2n - m$.*

Hence, in either case, we conclude that there is an $\overline{m} < 2n$ such that

$$\Lambda(m,n,A) \leq 2^{\frac{1}{4}}\Lambda(\overline{m},2\overline{m},\overline{A}).$$

Proof. We give the proof only for the case $n > 2m$. The proof for the case $m < n < 2m$ is analogous.

Let $\epsilon > 0$ and $A \in \mathbb{R}^{m\times n}$ be given. From Definition 4.1.1, one can find $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ such that $\Lambda(m,n,A) \leq \int_0^\infty \frac{\kappa(\mu)}{\mu}d\mu + \epsilon$. From Corollary 4.2.7, increasing the size of the problem $n-2m$ times, we obtain a new problem data $\overline{A} \in \mathbb{R}^{\overline{m}\times 2\overline{m}}$, $b \in \mathbb{R}^{\overline{m}}$ and $c \in \mathbb{R}^{2\overline{m}}$ such that

$$\int_0^\infty \frac{\kappa(\mu)}{\mu}d\mu \leq \int_0^\infty \frac{\overline{\kappa}(\mu)}{\mu}d\mu - (n-2m)\tau,$$

where $\int_0^\infty \frac{\overline{\kappa}(\mu)}{\mu}d\mu$ is the Sonnevend curvature of the new central path and τ is the constant derived in the proof of Corollary 4.2.7. Using Definition 4.1.1 once again, it follows that

$$\begin{aligned}\Lambda(m,n,A) \leq & \int_0^\infty \frac{\overline{\kappa}(\mu)}{\mu}d\mu - (n-2m)\tau + \epsilon \\ \leq & \Lambda(\overline{m},\overline{n},\overline{A}) - (n-2m)\tau + \epsilon.\end{aligned}$$

Since ϵ is arbitrarily small, the result follows. □

In the end, several observations are in order. First, even though we presented construction (4.2) for $n > 2m$, the same construction is valid for any m and n, and the increase in the Sonnevend curvature is still at least a constant. Second, repeating (4.2) leads to an $\Omega(n)$ worst-case lower bound for the Sonnevend curvature for a problem data $\overline{A}, \overline{b}, \overline{c}$, where the increase occurs for $\mu_i << \mu_{i+1}$. Since the constant increase occurs around a point on the central path close to the analytic center, each μ_i will be large. However, in the final LO problem, as Proposition A.3 shows, by doing the scaling $\hat{b} := \frac{\overline{b}}{\eta}$ by a large η, the same $\Omega(n)$ worst-case iteration-complexity can occur on any interval $[\mu^{'}, \mu^{''}]$.

Chapter 5

The iteration-complexity upper bound for IPMs is tight

It is an open question whether there is a IPM algorithm for the class of LO problems with $\mathcal{O}(n^{\alpha}\log(\frac{\mu_1}{\mu_0}))$ upper bound iteration-complexity for $\alpha < \frac{1}{2}$ to reduce the barrier parameter from μ_1 to μ_0. Sonnevend et al. Sonnevend et al. (1991) showed that for two distinct special classes of LO problems, we have the upper bounds $\mathcal{O}(n^{\frac{1}{4}}\log(\frac{\mu_1}{\mu_0}))$ and $\mathcal{O}(n^{\frac{3}{8}}\log(\frac{\mu_1}{\mu_0}))$. Another direction of research regarding the iteration-complexity of IPMs is to construct worst-case examples. Sonnevend et al. (1991) showed that a variant of MTY predictor-corrector algorithm requires $\Omega(n^{\frac{1}{3}})$ iterations to reduce the duality gap by $\log n$ for certain LO problems. A similar result has been obtained by Todd (1993) for the primal-dual affine scaling algorithm and has been later extended by Todd and Ye (1996) for long step primal-dual algorithm variants; they showed that these algorithms take $\Omega(n^{\frac{1}{3}})$ iterations to reduce the duality gap by a constant.

In this regard, a related open question raised by Stoer and Zhao (1993), was whether there is an $\alpha < \frac{1}{2}$ with $\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu \leq n^{\alpha} \log\left(\frac{\mu_1}{\mu_0}\right)$ for all LO problems. This chapter provides a negative answer to the latter question. In fact, we show that for any $\epsilon > 0$, there is a redundant KM cube as constructed by Nematollahi and Terlaky (2008b) for which $\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu = \Omega\left(n^{\left(\frac{1}{2}-\epsilon\right)} \log\left(\frac{\mu_1}{\mu_0}\right)\right)$.

5.1 KM cube construction

First we recall the KM construction in Nematollahi and Terlaky (2008b) and review its fundamental properties.

$$\begin{array}{llll} \max & -y_m & & \\ \text{s.t.} & 0 \leq y_1 & \leq 1 & \\ & \rho y_{k-1} \leq y_k & \leq 1 - \rho y_{k-1} & \text{for } k = 2 \ldots m. \\ & 0 & \leq d_1 + y_1 & \text{repeated } h_1 \text{ times} \\ & 0 & \leq d_2 + y_2 & \text{repeated } h_2 \text{ times} \\ & & \vdots & \\ & 0 & \leq d_m + y_m & \text{repeated } h_m \text{ times .} \end{array} \tag{5.1}$$

As in Nematollahi and Terlaky (2008b), we fix $\rho(m) = \frac{m}{2(m+1)}$ and choose $d = \left(\frac{1}{\sqrt{\rho^{m-1}}}, \frac{1}{\sqrt{\rho^{m-2}}}, \ldots, \frac{1}{\sqrt{\rho}}, 0\right)$. We denote the m-dimensional KM cube by $\mathcal{KM}(m, \rho(m))$.

Let us denote the slack variables by $\bar{s}_k = 1 - \rho y_{k-1}$ and $s_k = y_k - \rho y_{k-1}$ for $k = 2, \ldots, n$ with the convention $\bar{s}_1 = 1 - y_1$ and $s_1 = y_1$. There is a one-to-one correspondence between the vertices of $\mathcal{KM}(m, \rho(m))$ with the m-tuples $v^i \in \{0,1\}^m$, $i = 1, \ldots, 2^m$ as follows. Each vertex of $\mathcal{KM}(m, \rho(m))$ is determined by whether exactly one of $s_i = 0$ or $\bar{s}_i = 0$ for each $i = 1, \ldots, m$ in (5.1). If $s_i = 0$, then the i-th coordinate of the corresponding m-tuple in $\{0,1\}^m$ is 0; if $\bar{s}_i = 1$ it is 1. For our purpose, we describe the relevant terms of $\mathcal{KM}(m, \rho(m))$ inductively as follows:

First we describe the order of the vertices $\mathcal{V}(m)$ of $\mathcal{KM}(m, \rho(m))$ as the simplex path visits them. In this encoding, the i-th coordinate of a point in $\mathcal{V}(m)$ is set to 1 when its actual coordinate is larger than $\frac{1}{2}$; and to 0, when its actual coordinate is smaller than $\frac{1}{2}$. Note that $\mathcal{V}(m)$ is an encoding of the vertices of $\mathcal{KM}(m, \rho(m))$, they are not the actual vertex points in $\mathbb{R}^m$. For $m = 2$, let

$$\mathcal{V}(2) = \{v^1, v^2, v^3, v^4\} = \{(0,1), (1,1), (1,0), (0,0)\}. \tag{5.2}$$

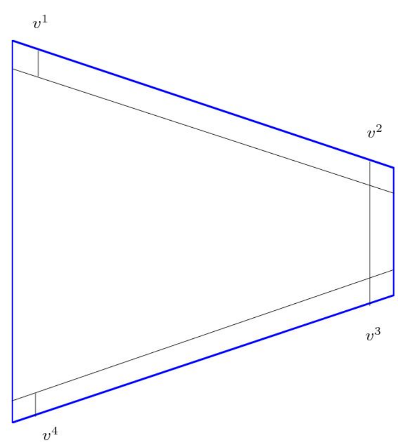

Figure 5.1: $\mathcal{V}(2) = \{v^1, v^2, v^3, v^4\} = \{(0,1), (1,1), (1,0), (0,0)\}$ gives an encoding of the vertices in the order they are visited by the central path.

Figure 5.1 shows the vertices of the $\mathcal{KM}(m, \rho(m))$. Then let

$$\mathcal{V}(m+1) = \{(v^{2^m}, 1), (v^{2^m-1}, 1), \ldots, (v^1, 1), (v^1, 0), (v^2, 0), \ldots, (v^{2^m}, 0)\}. \tag{5.3}$$

It was shown by Nematollahi and Terlaky (2008b) that there exists a redundant KM cube $\mathcal{KM}(m, \rho(m))$ whose central path, denoted by $\mathcal{CP}(m)$, visits the vertices in the order given in the set $\mathcal{V}(m)$. Figures 5.2 and 5.3 show the central path for $m = 2$ and $m = 3$.

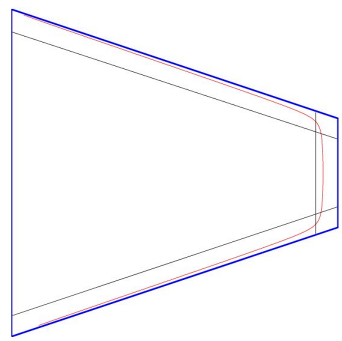

Figure 5.2: The central path visits the vertices $\mathcal{V}(2) = \{v^1, v^2, v^3, v^4\}$ of the $\mathcal{KM}(m, \rho(m))$ cube for $m = 2$ in the given order as μ decreases.

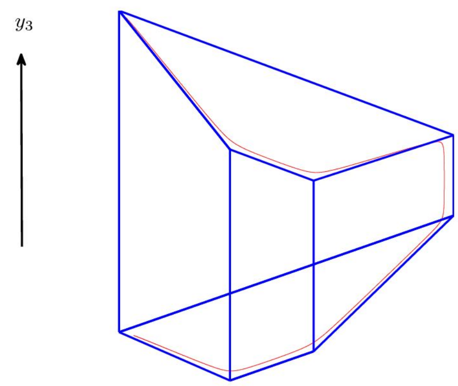

Figure 5.3: The central path in the redundant cube $\mathcal{KM}(m, \rho(m))$ cube for $m = 3$.

Next we define inductively a tube along the edges of the simplex path in $\mathcal{KM}(m, \rho(m))$ as follows. Let $\delta \leq \frac{1}{4(m+1)}$. Let $\mathcal{T}_\delta^U(2) = \{y : \mathbb{R}^2 : \bar{s}_2 \leq \delta\}$, $\mathcal{T}_\delta^L(2) = \{y : \mathbb{R}^2 : s_2 \leq \delta\}$ and $\mathcal{C}_\delta(m) = \{y : \mathbb{R}^2 : \bar{s}_m \geq \delta,\ s_m \geq \delta\}$ for $m \geq 2$. Note that $\mathcal{T}_\delta^U(2)$ and $\mathcal{T}_\delta^L(2)$ corresponds to a tube for the upper and lower facets of $\mathcal{KM}(2, \rho(2))$, respectively, while $\mathcal{C}_\delta(2)$ corresponds to the central part of $\mathcal{KM}(2, \rho(2))$, see Figure 5.1. By $\mathcal{T}_\delta(m)$, we denote the union $\mathcal{T}_\delta^L(m) \cup \mathcal{T}_\delta^U(m) \cup \mathcal{C}_\delta(m)$. Then for $m \geq 2$, define $\mathcal{T}_\delta^U(m+1) = \{y : \mathbb{R}^{m+1} : \bar{s}_{m+1} \leq \delta,\ (y_1, \ldots, y_m) \in \mathcal{T}_\delta(m)\}$ and $\mathcal{T}_\delta^L(m+1) = \{y : \mathbb{R}^{m+1} : s_{m+1} \leq \delta,\ (y_1, \ldots, y_m) \in \mathcal{T}_\delta(m)\}$. Notice that $\mathcal{T}_\delta^U(3)$ is a tube that corresponds to the upper facet of $\mathcal{KM}(3, \rho(3))$ where $y_3 = 1 - \rho y_2$. Similarly $\mathcal{T}_\delta^L(3)$ is a tube that corresponds to the lower facet of $\mathcal{KM}(3, \rho(3))$

where $y_3 = \rho y_2$. Also these upper and lower facets are $\mathcal{KM}(2, \rho(3))$ cubes themselves, see Figure 5.3. Hence by identifying the first m coordinates of $(y_1, \ldots, y_m, y_{m+1})$ inside $\mathcal{KM}((m+1), \rho(m+1))$ with $(y_1, \ldots, y_m) \in \mathcal{KM}(m, \rho(m+1))$, and considering the assumption that δ is decreasing in m, we can write $\mathcal{T}_\delta^U(m+1) \subset \mathcal{T}_\delta(m)$ and $\mathcal{T}_\delta^L(m+1) \subset \mathcal{T}_\delta(m)$, see Figure 5.4.

We also define a δ-neighborhood of a vertex of $\mathcal{KM}(m, \rho(m))$ by whether exactly one of $s_i \leq \delta$ or $\bar{s}_i \leq \delta$ for each $i = 1, \ldots, m$ in (5.1). Figure 5.1 displays the δ-neighborhoods of the vertices $\mathcal{V}(2) = \{v^1, v^2, v^3, v^4\}$ of the $\mathcal{KM}(m, \rho(m))$ cube for $m = 2$.

The following proposition is essentially Proposition 2.2 in Nematollahi and Terlaky (2008b).

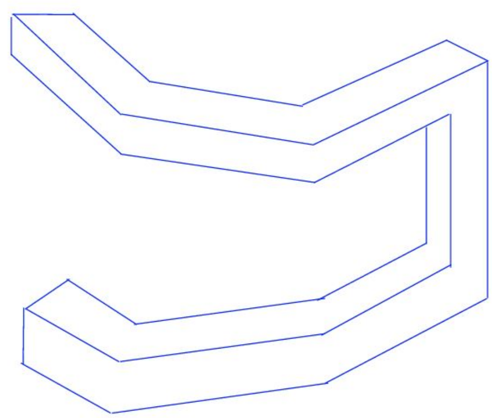

Figure 5.4: Illustration of the tube $\mathcal{T}_\delta(m)$ for $m = 3$.

Proposition 5.1.1. *In (5.1), one can choose the parameters in such a way that the central path $\mathcal{CP}(m)$ in $\mathcal{KM}(m, \rho(m))$ stays inside the tube $\mathcal{T}_\delta(m)$. In particular, we can choose $\rho = \frac{m}{2(m+1)}$, $\delta \leq \frac{1}{4(m+1)}$ so that $n = \mathcal{O}(m2^{2m})$. As μ decreases, the central path visits the δ-neighborhoods of the vertices given in the order (5.3). Moreover, the number of inequalities n is linear in $\frac{1}{\delta}$.*

Proof. See Proposition 2.2 in Nematollahi and Terlaky (2008b). □

Now for $\mathcal{KM}(m, \rho(m))$, we identify two regions R_δ^U and R_δ^L within tube $\mathcal{T}_\delta(m)$ in such a way that going from R_δ^U to R_δ^L (an vice versa) with line segments staying inside tube $\mathcal{T}_\delta(m)$ requires $\Omega(2^{m-1})$ number of iterations. Let $R_\delta^U := \{y \in \mathcal{KM}(m, \rho) : s_1 \leq \delta,\ s_2 \leq \delta, \ldots, s_{m-1} \leq \delta,\ \bar{s}_m \leq \delta\}$ and $R_\delta^L := \{y \in \mathcal{KM}(m, \rho) : s_1 \leq \delta,\ s_2 \leq \delta, \ldots, s_{m-1} \leq \delta,\ s_m \leq \delta\}$. We have the following.

Proposition 5.1.2. *For $\mathcal{KM}(m, \rho(m))$, let $y^U \in R_\delta^U$ and $y^L \in R_\delta^L$. Then staying inside the tube $\mathcal{T}_\delta(m)$, one requires at least 2^{m-1} line segments to reach y^U from y^L and vice versa.*

Proof. With the parameters chosen as in Proposition 5.1.1, we first show $\mathcal{T}_\delta^U(m)$ and $\mathcal{T}_\delta^L(m)$ do not intersect for any m. Suppose , to the contrary, that there is a $y \in \mathcal{T}_\delta^U(m) \cap \mathcal{T}_\delta^L(m)$. From the definition of $\mathcal{T}_\delta^U(m)$ and $\mathcal{T}_\delta^L(m)$, we have $\bar{s}_m = 1 - \rho y_{m-1} - y_m \leq \delta$ and $s_m = y_m - \rho y_{m-1} \leq \delta$. Adding these two inequalities, we get $1 - 2\rho y_{m-1} \leq 2\delta$. By the choice of ρ and δ, it is easy to see that, this will lead to the contradiction $y_{m-1} > 1$. Hence $\mathcal{T}_\delta^U(m) \cap \mathcal{T}_\delta^L(m) = \emptyset$.

The rest of the proof is by induction on m. For $m = 2$, let $y^U \in R_\delta^U$ and $y^L \in R_\delta^L$ with chosen $\delta \leq \frac{1}{4(m+1)}$. Then for y^U, we have that $s_1 = y_1 \leq \delta$ and $\bar{s}_2 \leq \delta$ implies $y_2 \geq 1 - \delta - \rho\delta \geq 1 - 2\delta = \frac{5}{6}$. Also, for y^L we have that $s_1 = y_1 \leq \delta$ and $s_2 \leq \delta$ implies $y_2 \leq \delta + \rho y_1 \leq 2\delta = \frac{1}{6}$. Clearly staying inside the tube $\mathcal{T}_\delta(m)$, it takes at least 2 iterations to reach a point with $y_2 \leq \frac{1}{6}$ from a point with $y_2 \geq \frac{5}{6}$, see Figure 5.1.

As inductive step, suppose that for any points in R_δ^U and R_δ^L belonging to the cube $\mathcal{KM}(m-1, \rho(m-1))$, one requires at least 2^{m-2} steps to reach the point in R_δ^L from the other point in R_δ^U with line segments staying inside $\mathcal{T}_\delta(m-1)$. Let $y^U \in R_\delta^U$ and

$y^L \in R^L_\delta$ inside $\mathcal{T}_\delta(m) \subset \mathcal{KM}(m, \rho(m))$. We distinguish two points p^1 and p^2 such that

$$p^1 \in \{y \in \mathcal{KM}(m, \rho(m)) : s_1 \leq \delta,\ s_2 \leq \delta, \ldots, \bar{s}_{m-1} \leq \delta,\ \bar{s}_m \leq \delta\}$$

and

$$p^2 \in \{y \in \mathcal{KM}(m, \rho(m)) : s_1 \leq \delta,\ s_2 \leq \delta, \ldots, \bar{s}_{m-1} \leq \delta,\ s_m \leq \delta\}.$$

Note that the point p^1 belongs to the δ-neighborhood of the vertex point $v^{2^{m-1}} = (0, 0, \ldots, 0, 1, 1)$ and the point p^2 belongs to the δ-neighborhood of the vertex point $v^{2^{m-1}+1} = (0, 0, \ldots, 0, 1, 0)$. Then, using the inductive definition of $\mathcal{T}^U_\delta(m)$ and $\mathcal{T}^L_\delta(m)$, it is easy to see that $y^U, p^1 \in \mathcal{T}^U_\delta(m)$ and $p^2, y^L \in \mathcal{T}^L_\delta(m)$. By inductive hypothesis, one needs at least 2^{m-2} line segments to reach p^1 from y^U staying inside the tube $\mathcal{T}^U_\delta(m) \subset \mathcal{T}_\delta(m-1)$. Similarly, one needs at least 2^{m-2} line segments to reach y^L from p^2 staying inside the tube $\mathcal{T}^L_\delta(m) \subset \mathcal{T}_\delta(m-1)$. Moreover since by the first part of the proof, we have $\mathcal{T}^U_\delta(m) \cap \mathcal{T}^L_\delta(m) = \emptyset$, it follows that to reach y^L from y^U, one needs to traverse within $\mathcal{T}_\delta(m-1)$ twice, each time requiring at least 2^{m-2} steps. This proves that one requires at least 2^{m-1} line segments to reach y^U from y^L, and the proof is complete. □

5.2 Neighborhood of the KM cube central path

In Section 5.1, we showed that with redundant constraints $n = \mathcal{O}(m2^{2m})$, the central path $\mathcal{CP}(m)$ stays inside a tube $\mathcal{T}_\delta(m)$. Moreover, we proved that starting from a point in R^U_δ close to the analytic center of $\mathcal{KM}(m, \rho(m))$, it will take at least 2^{m-1} line segments to reach a point in R^L_δ close to the optimal solution of (5.1). However, path-following IPMs including the MTY predictor-corrector algorithm, uses the neighborhood $\mathcal{N}(\beta)$ as opposed to the tube neighborhood $\mathcal{T}_\delta(m)$ we used in Section 5.1. In this section we analyze the $\mathcal{N}(\beta)$ neighborhood for the $\mathcal{KM}(m, \rho(m))$, and prove that for $\beta = \Omega(\frac{1}{m+1})$, we have $\mathcal{N}(\beta) \subset \mathcal{T}_\delta(m)$. In other words, with appropriately chosen neighborhood parameters of $\mathcal{KM}(m, \rho(m))$, all the iterates of the MTY predictor-corrector algorithm stays inside tube $\mathcal{T}_\delta(m)$. Hence we can draw the conclusion that for $\mathcal{KM}(m, \rho(m))$, the MTY predictor-corrector algorithm requires at least $\Omega(2^{m-1})$ iterations when the neighborhood $\mathcal{N}(\beta)$ is used with $\beta = \Omega(\frac{1}{m+1})$.

In order to find the largest β for which $\mathcal{N}(\beta) \subset \mathcal{T}_\delta(m)$, we will use the weighted paths. The following lemma is basically Lemma 4.1 in Stoer and Zhao (1993).

Lemma 5.2.1. *Fix μ and let $w > 0$ such that $\|w - e\|_2 \leq \epsilon$. Let $(x^w(\mu), y^w(\mu), s^w(\mu))$ denote w-weighted path which is the solution of (2.4). Let $\Delta s_i(\mu) = s_i^w(\mu) - s_i(\mu)$ where $s_i(\mu)$ is the central path point for $i = 1, \dots, n$. Then we have $\left|\frac{\Delta s_i(\mu)}{s_i(\mu)}\right| \leq 2\epsilon$ for $i = 1, \dots, n$.*

When we apply the information in Lemma 5.2.1 to $\mathcal{KM}(m, \rho(m))$, we obtain the following result.

Lemma 5.2.2. *There exists a $\mathcal{KM}(m, \rho(m))$ with $n = \mathcal{O}(m2^{2m})$ such that all the w-weighted paths with $\|w - e\|_2 \leq \beta := \frac{\delta}{4}$ stays inside the tube $\mathcal{T}_\delta(m)$ when $\delta \leq \frac{1}{4(m+1)}$.*

Proof. Let $\delta \leq \frac{1}{4(m+1)}$. Then from Proposition 5.1.1, there exists $\mathcal{KM}(m, \rho(m))$ with $n = \mathcal{O}(m2^{2m})$ so that the central path stays inside the tube $\mathcal{T}_{\frac{\delta}{2}}(m)$. Choose $\beta = \frac{\delta}{4}$ for $\mathcal{KM}(m, \rho(m))$ so that $\|w - e\|_2 \leq \beta$. Since for all the slacks, we have $s_i(\mu) \leq 1$ or $\bar{s}_i(\mu) \leq 1$, Lemma 5.2.1 implies that $s_i^w(\mu) \leq s_i(\mu) + \frac{\delta}{2}$ and $\bar{s}_i^w(\mu) \leq \bar{s}_i(\mu) + \frac{\delta}{2}$. Then whenever $s_i(\mu) \leq \frac{\delta}{2}$ or $\bar{s}_i(\mu) \leq \frac{\delta}{2}$, we have $\bar{s}_i^w(\mu) \leq \delta$ and $s_i^w(\mu) \leq \delta$. Since a tube $\mathcal{T}_\delta(m)$ with a general δ inside $\mathcal{KM}(m, \rho(m))$ is determined by these slacks, it follows that all w-weighted paths stay inside the tube $\mathcal{T}_\delta(m)$ with $\delta \leq \frac{1}{4(m+1)}$. This concludes the proof. □

The following lemma proves a result analogous to Lemma 5.2.2 tailored for R_δ^U and R_δ^L.

Lemma 5.2.3. *Let $\delta \le \frac{1}{4(m+1)}$ and fix $\beta := \frac{\delta}{4}$. Suppose that $y(\mu_1) \in R^U_{\delta/2}$ for some μ_1. Then $\mathcal{N}(\beta, \mu_1) \subset R^U_\delta$. Similarly, if for some μ_0, $y(\mu_0) \in R^L_{\delta/2}$, then $\mathcal{N}(\beta, \mu_0) \subset R^L_\delta$.*

Proof. Suppose that for some μ_1, $y(\mu_1) \in R^U_{\delta/2}$, i.e., $s_1(\mu) \le \frac{\delta}{2},\ s_2(\mu) \le \frac{\delta}{2}, \ldots, s_{m-1}(\mu) \le \frac{\delta}{2}, \bar{s}_m(\mu) \le \frac{\delta}{2}$. Let $y \in \mathcal{N}(\beta, \mu_1)$. Then, for $w := \frac{xs}{\mu_1}$ we have $\|w - e\|_2 \le \beta$. Since for all the slacks in $\mathcal{KM}(m, \rho(m))$, we have $s_i(\mu) \le 1$ or $\bar{s}_i(\mu) \le 1$, Lemma 5.2.1 implies that $s^w_i(\mu) \le s_i(\mu) + \frac{\delta}{2}$ and $\bar{s}^w_i(\mu) \le \bar{s}_i(\mu) + \frac{\delta}{2}$. Then, whenever $s_i(\mu) \le \frac{\delta}{2}$ or $\bar{s}_i(\mu) \le \frac{\delta}{2}$, we have $\bar{s}^w_i(\mu) \le \delta$ and $s^w_i(\mu) \le \delta$. This proves $y \in R^U_\delta$, which implies $\mathcal{N}(\beta, \mu_1) \subset R^U_\delta$. The proof of the rest of the claim can be done analogously. □

In the rest of this section, we aim to find an interval $[\mu_0, \mu_1]$ and an upper bound for $\log(\frac{\mu_1}{\mu_0})$ such that for some δ and β, the neighborhoods satisfy $\mathcal{N}(\beta, \mu_1) \subset R^U_\delta$ and $\mathcal{N}(\beta, \mu_0) \subset R^L_\delta$.

Let $\delta \le \frac{1}{4(m+1)}$ and $(y_1(\mu_1), \ldots, y_m(\mu_1))$ be a central path $\mathcal{CP}(m)$ point such that $s_1(\mu) = \frac{\delta}{2},\ s_2(\mu) \le \frac{\delta}{2}, \ldots, \bar{s}_m(\mu) \le \frac{\delta}{2}$. Note that any point satisfying $s_1(\mu) = \frac{\delta}{2},\ s_2(\mu) \le \frac{\delta}{2}, \ldots, \bar{s}_m(\mu) \le \frac{\delta}{2}$ is inside the $\frac{\delta}{2}$-neighborhood of the vertex point $(0, 0, \ldots, 0, 1)$, hence Proposition 5.1.1 guarantees the existence of a central path point $(y_1(\mu_1), \ldots, y_m(\mu_1))$. Then, by using Theorem 3.7 in Nematollahi and Terlaky (2008b), one can show that $\mu_1 \le \frac{\rho^{m-1}\delta}{2}$. Let us fix $\mu_1 = \frac{\rho^{m-1}\delta}{2}$ and let $\beta := \frac{\delta}{4}$. Then Lemma 5.2.3 implies that the neighborhood $\mathcal{N}(\beta, \mu_1)$ stays inside the region R^U_δ. Hence, any point inside the neighborhood $\mathcal{N}(\beta, \mu_1)$ also stays inside the region R^U_δ.

The next step is to find a μ_0 such that the neighborhood $\mathcal{N}(\beta, \mu_0)$is within the region R^L_δ. Let $(y_1(\mu_0), \ldots, y_m(\mu_0))$ be the central path point such that $y_m = \frac{\rho^{m-1}\delta}{2}$. Note that since the objective function in (5.1) is $-y_m$, a central point satisfying $y_m = \frac{\rho^{m-1}\delta}{2}$ exists and is unique. Since from (5.1), we have $\rho y_i \le y_{i+1}$ for $i = 1, \ldots, (m-1)$, we obtain $y_1(\mu_0) \le \frac{\delta}{2},\ y_2(\mu_0) \le \frac{\delta}{2}, \ldots, y_m(\mu_0) \le \frac{\delta}{2}$, which in turn implies that $s_1(\mu_0) \le \frac{\delta}{2},\ s_2(\mu_0) \le \frac{\delta}{2}, \ldots, s_m(\mu_0) \le \frac{\delta}{2}$. Then, using Lemma 5.2.3 once again, it follows for

μ_0 that the neighborhood $\mathcal{N}(\beta, \mu_0)$ stays inside the region R^L_δ for $\beta = \frac{\delta}{4}$. For the central path (1.5), the duality gap $c^T x(\mu) - b^T y(\mu) = n\mu$. It is well-known (see e.g., Roos et al. (2006)) that $b^T y(\mu)$ is monotonically increasing and $c^T x(\mu)$ is monotonically decreasing along the central path. In our case, $b^T y(\mu) = -y_m(\mu)$ is increasing to 0 and $c^T x(\mu)$ is monotonically decreasing to 0, i.e., $c^T x(\mu) > 0$ for all $\mu > 0$. Then $n\mu = c^T x(\mu) - b^T y(\mu) > y_m(\mu)$ implies that $\mu > \frac{y_m(\mu)}{n}$ for any point on the central path. Hence for the central path point for which $y_m(\mu) = \frac{\rho^{m-1}\delta}{2}$, it follows that $\mu_0 > \frac{\rho^{m-1}\delta}{2n}$. Then using the fact that $n = \mathcal{O}(m2^{2m})$, we have $\log(\frac{\mu_1}{\mu_0}) = \mathcal{O}(m)$. The following corollary summarizes our findings.

Corollary 5.2.4. *Let the neighborhood parameters be given as* $\beta_0 < \beta_1 = \frac{1}{16(m+1)}$ *for the MTY predictor-corrector algorithm. Then there exists a* $\mathcal{KM}(m, \rho(m))$ *with* $n = \mathcal{O}(m2^{2m})$ *for which MTY predictor-corrector algorithm requires at least* $\Omega(2^{m-1})$ *predictor steps to reduce the barrier parameter from* μ_1 *to* μ_0*, where* $\log(\frac{\mu_1}{\mu_0}) = \mathcal{O}(m)$.

Proof. Fix $\delta := \frac{1}{4(m+1)}$ and $\beta_1 = \frac{\delta}{4} = \frac{1}{16(m+1)}$. We know from Lemma 5.2.2 that, there exists a $\mathcal{KM}(m, \rho(m))$ with $n = \mathcal{O}(m2^{2m})$ such that $\mathcal{N}(\beta) \subset \mathcal{T}_\delta(m)$. Further, Lemma 5.2.3 shows that there is an interval $[\mu_0, \mu_1]$ such that the neighborhoods $\mathcal{N}(\beta, \mu_1) \subset R^U_\delta$ and $\mathcal{N}(\beta, \mu_0) \subset R^L_\delta$. Now starting from an iterate (x^1, y^1, s^1) and μ_1 such that $(x^1, y^1, s^1) \in \mathcal{N}(\beta, \mu_1) \subset R^U_\delta$, in order to reach an iterate (x^0, y^0, s^0) and μ_0 such that $(x^0, y^0, s^0) \in \mathcal{N}(\beta, \mu_0) \subset R^L_\delta$; Proposition 5.1.2 and Proposition 5.2.2 imply that one needs $\Omega(2^{m-1})$ steps. Since the number of corrector steps is constant, it follows that the number of predictor steps is $\Omega(2^{m-1})$. Moreover the discussion after Lemma 5.2.3 proves that, we can choose the interval $[\mu_0, \mu_1]$ so that $\log(\frac{\mu_1}{\mu_0}) = \mathcal{O}(m)$. This completes the proof. □

5.3 A worst-case iteration-complexity lower bound for the Sonnevend curvature

In Section 5.2, we established that there is an interval $[\mu_0, \mu_1]$ such that the MTY predictor-corrector algorithm requires $\Omega(2^{m-1})$ iterations to reduce the barrier parameter from μ_1 to μ_0 for the enlarged neighborhood $\mathcal{N}(\beta_1)$, where $\beta_1 = \Omega(\frac{1}{m+1})$. In this section, our goal is to obtain a lower bound for the Sonnevend curvature using the tools from the previous section. To this end, we need to examine the constants in Theorem 3.2.1 more closely.

Lemma 5.3.1. *Let β_1 be the enlarged neighborhood constant so that $\beta_1 \leq \dfrac{1}{400}$ and let N be the number of iterations of the MTY predictor-corrector algorithm to reduce the barrier parameter from μ_1 to μ_0. Then*

$$N \leq \frac{4\sqrt{2}}{\sqrt{\beta_1}} \int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu + \frac{1}{2\log(1+\frac{\sqrt{\beta_1}}{4})} \log\left(\frac{\mu_1}{\mu_0}\right). \tag{5.4}$$

Proof. See Theorem 2.4 and its proof in Stoer and Zhao (1993). □

The following theorem shows that on the interval $[\mu_0, \mu_1]$, the total Sonnevend curvature is in comparable order to the number of sharp turns of the central path.

Theorem 5.3.2. *There is an integer $m_0 > 0$ such that for any $m \geq m_0$, there exists a $\mathcal{KM}(m, \rho(m))$ and interval $[\mu_0, \mu_1]$ such that the Sonnevend curvature satisfies*

$$\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu = \Omega\left(\left(\frac{\sqrt{n}}{\sqrt{\log n}\sqrt{\log(n+1)}} - \frac{8\sqrt{\log n + 1}}{\log(2)}\right) \log\left(\frac{\mu_1}{\mu_0}\right)\right). \tag{5.5}$$

Proof. Let $\beta_1 = \frac{1}{16(m+1)}$ and let us choose the parameters of $\mathcal{KM}(m, \rho(m))$ as $\rho = \frac{m}{2(m+1)}$, and $\delta = \frac{1}{8(m+1)}$ so that $n = \mathcal{O}(m2^{2m})$. Write $n = \tau m 2^{2m}$ for some constant

$\tau > 0$ and we calculate $\log\left(\frac{\mu_1}{\mu_0}\right) = \log n = \log\tau + \log m + 2m$. This shows that for large enough m, $\log\left(\frac{\mu_1}{\mu_0}\right) = \mathcal{O}(m)$. Since we can extend the interval $[\mu_0, \mu_1]$ so that it still includes all the sharp turns, we will assume that $\log\left(\frac{\mu_1}{\mu_0}\right) = \mathcal{O}(m)$. Then Corollary 5.2.4 applies and we have $N \geq 2^{m-1}$. Now using the bound $\log(1+\omega) \geq (\log 2)\omega$ for $0 \leq \omega \leq 1$, from (5.5) we get the inequality

$$\frac{1}{2\log(1+\frac{\sqrt{\beta_1}}{4})} \leq \frac{8\sqrt{m+1}}{\log 2}.$$

Using the fact that $m \leq \log n$, a straightforward calculation shows that

$$\frac{\displaystyle\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu}{\log\left(\frac{\mu_1}{\mu_0}\right)} = \Omega\left(\frac{\sqrt{n}}{\sqrt{\log n}\sqrt{\log(n+1)}} - \frac{8\sqrt{\log n + 1}}{\log(2)}\right). \tag{5.6}$$

This completes the proof. □

Corollary 5.3.3. *For any $\epsilon > 0$, there is an integer $m_0 > 0$ such that for any $m \geq m_0$, there exists a $\mathcal{KM}(m, \rho(m))$ and interval $[\mu_0, \mu_1]$ such that* $\displaystyle\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu \geq n^{\left(\frac{1}{2}-\epsilon\right)} \log\left(\frac{\mu_1}{\mu_0}\right)$, *where* $\log\left(\frac{\mu_1}{\mu_0}\right) = \mathcal{O}(m)$.

Proof. The claim follows from Theorem 5.3.2 for sufficiently large m. □

Remark 5.3.4. *Corollary 5.3.3 yields a negative answer to the question raised by Stoer and Zhao (1993), i.e., whether there exists an $\alpha < \frac{1}{2}$ with $\log\left(\frac{\mu_1}{\mu_0}\right) = \Omega(1)$ such that* $\displaystyle\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu \leq n^{\alpha} \log\left(\frac{\mu_1}{\mu_0}\right)$ *for the class of LO problems.*

5.4 An iteration-complexity lower bound for the MTY predictor-corrector algorithm with constant neighborhood parameter

In practice, the MTY predictor-corrector algorithm operates in a larger neighborhood where β_1 is a constant. In order to conclude an iteration-complexity lower bound for the MTY predictor-corrector algorithm with constant neighborhood parameter β_1, by using Theorem 3.2.1 we need to show that for $\mathcal{KM}(m, \rho(m))$, there is a constant $\nu > 0$ with $\kappa(\mu) \geq \nu$ for $\mu \in [\mu_0, \mu_1]$. While this appears to hold numerically, proving it is much more difficult. To go around this difficulty, we exploit a trick introduced by Sonnevend et al. (1991). The idea is to use one dimensional LO problems, where it is easier to calculate the central path and its corresponding $\kappa(\mu)$; and to use LO problems with the scaled objectives with block diagonal constraints. For the details, we refer the reader to Appendix Section A.

Recall that by Corollary 5.3.3, we know there exists a $\mathcal{KM}(m, \rho(m))$ and an interval $[\mu_0, \mu_1]$ such that $\displaystyle\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu \geq n^{\left(\frac{1}{2}-\epsilon\right)} \log\left(\frac{\mu_1}{\mu_0}\right)$. Here $n = \mathcal{O}(m2^{2m})$ and $\frac{\mu_1}{\mu_0} = \mathcal{O}(\log n)$. Now by using Lemma A.4 and Proposition A.2, we can embed $\mathcal{KM}(m, \rho(m))$ in a block diagonal LO problem at the expense of increasing the size of the problem by at most $\overline{n} := n + \mathcal{O}(m + \log m)$. Denote by $\overline{\mathcal{KM}}(\overline{m})$ this hybrid construction in which $\mathcal{KM}(m, \rho(m))$ is embedded. Since $\overline{n} = \mathcal{O}(n)$, we have the following:

Theorem 5.4.1. *For any $\epsilon > 0$, there exists a positive integer m_0 such that for any $\overline{m} \geq m_0$, there exists an LO problem $\overline{\mathcal{KM}}(\overline{m})$ and an interval $[\mu_0, \mu_1]$ with the following properties:*

- $\dfrac{\mu_1}{\mu_0} = \mathcal{O}\left(m2^{2m}\right)$.
- *Let $\beta_0 < \beta_1 \leq \frac{1}{400}$ be the constant neighborhood $\mathcal{N}(\beta)$ parameters. Then the MTY*

predictor-corrector algorithm on this neighborhood requires $\Omega\left(n^{\left(\frac{1}{2}-\epsilon\right)}\log\left(\frac{\mu_1}{\mu_0}\right)\right)$ *predictor steps.*

Proof. Consider the $\mathcal{KM}(m,\rho(m))$ from Corollary 5.3.3. Then by using Lemma A.4 and Proposition A.2, we can embed $\mathcal{KM}(m,\rho(m))$ in a block diagonal LO problem with size $\overline{n} := n+\mathcal{O}(m+\log m)$ and $\overline{m} = \mathcal{O}(m)$. Note that since the interval $[\mu_0,\mu_1]$ comes from $\mathcal{KM}(m,\rho(m))$, the first claim in the theorem follows from Corollary 5.3.3. Also since for $\overline{\mathcal{KM}}(\overline{m})$, its corresponding $\overline{\kappa}(\mu) \geq \nu$ for some constant $\nu > 0$ for all $\mu \in [\mu_0,\mu_1]$, Theorem 3.2.1 implies the first claim. This completes the proof. □

Chapter 6

Condition numbers for Sonnevend's curvature

6.1 A strongly polynomial bound for Sonnevend's curvature

In Section 3.2, we mention that the condition numbers $\overline{\chi}_A^*$ and $\overline{\chi}_A$ can have arbitrarily different orders of magnitude. In this section, we will show that for $m = 1$ and $m = n-1$, the Sonnevend curvature $\int_0^\infty \frac{\kappa(\mu)}{\mu} d\mu$ admits a strongly polynomial upper bound. To this end, the following lemma summarizes important properties of the condition number $\overline{\chi}_A$.

Lemma 6.1.1. Monteiro and Tsuchiya (2008) *Let $A \in \mathbb{R}^{m\times n}$ be a matrix of full rank. Then,*

1. $\overline{\chi}_A = \max_B \ \|B^{-1}A\|_2$ *where B is a non-singular submatrix of A.*

2. *For any non-singular matrix $G \in \mathbb{R}^{m\times m}$, $\overline{\chi}_{GA} = \overline{\chi}_A$.*

3. $\log(\overline{\chi}_A) = \mathcal{O}(L_A)$ *where* L_A *is the input bit length of* A *when the matrix has all integer entries.*

4. $\overline{\chi}_A = \overline{\chi}_F$ *for any* $F \in \mathbb{R}^{(n-m)\times n}$ *such that* $\mathcal{N}(A) = \mathcal{R}(F^T)$.

The following proposition is one of the main results of this section.

Proposition 6.1.2. *Let* $A \in \mathbb{R}^{m\times n}$ *be a matrix of full rank where either* $m = 1$ *or* $m = n-1$. *Then* $\log(\overline{\chi}_A^*) = \mathcal{O}(n)$ *which implies that*

$$\int_0^\infty \frac{\kappa(\mu)}{\mu} d\mu = \mathcal{O}(n^{4.5}). \tag{6.1}$$

Proof. Suppose that $A = [a_1, a_2, \ldots, a_n]$ is a non-zero matrix of a single row. Let D be a positive diagonal matrix with entries $d_i = \frac{1}{|a_i|}$ for $a_i \neq 0$ and $d_i = 1$ for $a_i = 0$, $i = 1, \ldots, n$. Then each entry of AD is either 1, -1 or 0. Then, using the definition $\overline{\chi}_A^* := \inf_D\{\overline{\chi}_{AD}\}$ and Lemma 6.1.1 part *3.*, we have $\log(\overline{\chi}_A^*) \leq \log(\overline{\chi}_{AD}) = \mathcal{O}(L_{AD})$. Since the matrix AD has only 1, -1 or 0 entries, Definition 2.1.1 gives $L_{AD} = \mathcal{O}(n)$. The, for any $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ so that the central path (1.5) exists, and equation (3.18) implies that the Sonnevend curvature satisfies $\int_0^\infty \frac{\kappa(\mu)}{\mu} d\mu = \mathcal{O}(n^{4.5})$. This proves the claim for the case $m = 1$.
Now suppose $A \in \mathbb{R}^{m\times n}$ is a matrix of full rank with $m = n-1$. Let $F \in \mathbb{R}^{1\times n}$ be any matrix such that $\mathcal{N}(A) = \mathcal{R}(F^T)$. By the first part of the proof, we know that there is a positive diagonal matrix D such that $\log(\overline{\chi}_{FD}) = \mathcal{O}(n)$. Now it is easy to verify that $\mathcal{N}(AD) = \mathcal{R}((FD)^T)$. Then Lemma 6.1.1 part *4.* implies that $\overline{\chi}_A^* \leq \overline{\chi}_{AD} = \overline{\chi}_{FD} = \mathcal{O}(2^n)$. This implies that $\log(\overline{\chi}_A^*) = \mathcal{O}(n)$ yielding (6.1). This completes the proof. □

Now the question arises whether the type of argument in Proposition 6.1.2, i.e., using the condition number $\overline{\chi}_A^*$, could be extended to the case of general (m, n) to yield a strongly polynomial bound for $\displaystyle\int_0^\infty \frac{\kappa(\mu)}{\mu} d\mu$. Unfortunately, the answer is negative as illustrated by the following example. It shows that there is a 2×4 matrix A for which $\overline{\chi}_A^*$ could be arbitrarily large.

Example 6.1.3. *Let* $A = [I_{2\times 2}, U]$ *where* $U = \begin{bmatrix} r & r \\ r & r + \frac{1}{r^2} \end{bmatrix}$ *for* $r > 0$. *Recall that* $\overline{\chi}_A^* := \inf\{\overline{\chi}_{AD}\}$ *for* D *being a strictly positive diagonal matrix. Then for any* $\epsilon > 0$ *given, there is a diagonal matrix* $D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}$, *with* D_1, D_2 *being* 2×2 *diagonal matrices such that* $\overline{\chi}_A^* \geq \overline{\chi}_{AD} - \epsilon$. *Then by Lemma 6.1.1 part 1.,* $\overline{\chi}_{AD} = \overline{\chi}_{D_1^{-1}AD_1}$ *and* $\overline{\chi}_{AD} = \overline{\chi}_{D_2^{-1}U^{-1}AD_2}$. *Now we calculate* $D_1^{-1}AD_1 = [I_{2\times 2}, D_1^{-1}UD_2]$ *with*

$$D_1^{-1}UD_2 = \begin{bmatrix} \frac{d_3}{d_1}r & \frac{d_4}{d_1}r \\ \frac{d_3}{d_2}r & \frac{d_4}{d_2}(r + \frac{1}{r^2}) \end{bmatrix}, \tag{6.2}$$

and $D_2^{-1}U^{-1}AD_2 = [D_2^{-1}U^{-1}D_1, I_{2\times 2}]$ *with*

$$D_2^{-1}U^{-1}D_1 = \begin{bmatrix} \frac{d_1}{d_3}(r^2 + \frac{1}{r}) & \frac{d_2}{d_3}(-r^2) \\ \frac{d_1}{d_4}(-r^2) & \frac{d_2}{d_4}(r + \frac{1}{r^2}) \end{bmatrix}. \tag{6.3}$$

Now if $\frac{d_3}{d_1} \geq 1$, *then from (6.2), we have* $||D_1^{-1}UD_2||_2 \geq ||D_1^{-1}UD_2||_{\max} \geq r$. *On the other hand, if* $\frac{d_3}{d_1} < 1$, *from (6.3), we have* $||D_2^{-1}U^{-1}D_1||_2 \geq ||D_2^{-1}U^{-1}D_1||_{\max} \geq (r^2 + \frac{1}{r})$. *Hence in either case, we have* $\overline{\chi}_{AD} \geq r$. *Since* $\overline{\chi}_{AD} \leq \overline{\chi}_A^* + \epsilon$, *we conclude that* $\overline{\chi}_A^*$ *does not have a strongly polynomial upper bound in general.*

Chapter 7

Volumetric barrier and path

Following the notation of Section 1.2, let $\mathcal{D} = \{y \in \mathbb{R}^m : A^T y \leq c\}$ with $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $y \in \mathbb{R}^m$. We assume that the feasible set $\mathcal{D}$ has a nonempty interior and is bounded. Let $F(y) = -\sum_{i=1}^{n} \log(c_i - a_i^T y)$ be the logarithmic barrier function, and $H(y) = \nabla^2 F(y)$ be the Hessian of $F(y)$. We know $F(y)$ is a strictly convex function, so the Hessian $H(y)$ is positive-definite. Define $V(y) = \text{logdet } H(y)$. The function $V(y)$ is called the volumetric barrier function for $y \in P$ and is known to be strictly convex as well Vaidya (1989). The volumetric barrier $V(y)$ is another self-concordant barrier function.

In Atkinson and Vaidya (1995) Atkinson and Vaidya introduces a cutting plane algorithm using the volumetric barrier function. This algorithm and its iteration-complexity bound is further improved in Anstreicher (1997, 1998). Anstreicher Anstreicher (2000) extends the volumetric barrier approach to the semidefinite case. The volumetric barrier IPMs algorithm Anstreicher (1996) has an iteration-complexity upper bound $\mathcal{O}(n^{1/4} m^{1/2} \log n/\epsilon)$. Notice that when $n \gg m$, this bound is significantly better than the classical one $\mathcal{O}(\sqrt{n} \log n/\epsilon)$ for the logarithmic barrier function.

Given the LO problem in the form with $b \neq 0$:

$$\begin{aligned} \max \quad & b^T y \\ \text{s.t.} \quad & A^T y \leq c, \end{aligned} \tag{7.1}$$

for $\mu > 0$. Consider the volumetric barrier problem:

$$\begin{aligned} \min \quad & -b^T y + \mu V(y) \\ \text{s.t.} \quad & A^T y < c. \end{aligned} \tag{7.2}$$

Let $G_\mu(y) = -b^T y + \mu V(y)$ and $y(\mu)$ be the (unique) optimal solution of $G_\mu(y)$. The optimal points $y(\mu)$ parameterized by μ form an analytic curve called the volumetric path. As $\mu \to 0$, $y(\mu)$ converges to an optimal solution of (7.1).

Compared to the vast amount of literature on logarithmic barrier methods in LO, the volumetric barrier function is relatively less studied possibly due to its inferior performance in computational practice and its more involved analysis. In this section, we will prove that certain basic properties that hold for the logarithmic barrier do not hold for the volumetric barrier, see Mut and Terlaky (2012).

7.1 Basic Properties

In the next two propositions we prove certain fundamental properties of the volumetric path. The next proposition deals with the monotonicity of the objective on the volumetric path, see Roos et al. (2006).

Proposition 7.1.1. *For $\mu_1 < \mu_2$, we have the following:*

1. $y(\mu_1) \neq y(\mu_2)$.
2. $b^T y(\mu_1) > b^T y(\mu_2)$.

Proof.

1. Let $\mu_1 < \mu_2$ and suppose, to the contrary, that $y(\mu_1) = y(\mu_2) = \overline{y}$. The first order optimality conditions for (7.2) give

$$\nabla(-b^T y + \mu_1 V(y)) = 0 \Longrightarrow -b + \mu_1 \nabla V(y) = 0$$

$$\nabla(-b^T y + \mu_1 V(y)) = 0 \Longrightarrow -b + \mu_2 \nabla V(y) = 0,$$

that implies

$$\nabla V(y) = \frac{1}{\mu_1} b = \frac{1}{\mu_2} b,$$

which is a contradiction for $b \neq 0$.

2. Let y^1 and y^2 be the optimal solutions of G_{μ_1} and G_{μ_2}, respectively. Since G_μ is strictly convex, for $\mu > 0$, we have

$$G_{\mu_1}(y^1) < G_{\mu_1}(y^2)$$

$$G_{\mu_2}(y^2) < G_{\mu_2}(y^1),$$

which implies

$$-b^T y^1 + \mu_1 V(y^1) < -b^T y^2 + \mu_1 V(y^2) \tag{7.3}$$

$$-b^T y^2 + \mu_2 V(y^2) < -b^T y^1 + \mu_2 V(y^1). \tag{7.4}$$

By multiplying the inequalities (7.3) by μ_2, and (7.4) by μ_1, respectively, and adding the resulting inequalities, after cancellations one gets

$$-\mu_2 b^T y^1 - \mu_1 b^T y^2 < -\mu_2 b^T y^2 - \mu_1 b^T y^1,$$

which implies that $b^T y^1 > b^T y^2$. □

Next we examine the relationship between the points $y(\mu)$, $\mu > 0$ on the volumetric path and the so-called volumetric center of the level sets $b^T y = \alpha$.

Definition 7.1.2.

1. *A level set* $\mathcal{L}_\alpha$ *for* (7.1) *is the set* $\{y \in \mathbb{R}^n : b^T y = \alpha,\ A^T y \leq c\}$

2. *The volumetric center* $\hat{y}$ *of a (bounded) level set* $\mathcal{L}_\alpha$ *is defined to be the (unique) minimizer of the volumetric function* $V(y)$ *over* $\mathcal{L}_\alpha$.

Proposition 7.1.3. *Let* $\mu > 0$ *and* $\hat{y} = y(\mu)$ *be the optimal solution of* (7.2) *with* $b^T \hat{y} = \alpha$ *for some* α. *Then* $\hat{y}$ *is the volumetric center of* $\mathcal{L}_\alpha$.

Proof. Consider the following problems:

$$(\S)\quad \begin{array}{ll} \min & V(y) \\ \text{s.t} & b^T y = \alpha \\ & A^T y < c. \end{array} \qquad\qquad (\S\S)\quad \begin{array}{ll} \min & -\frac{b^T y}{\mu} + V(y) \\ \text{s.t} & A^T y < c. \end{array}$$

Let $\overline{y}$ and $\hat{y}$ be the optimal solutions of (§) and (§§), respectively. The first order optimality conditions for (§) give

$$\nabla V(\overline{y}) + \lambda b = 0, \quad b^T \overline{y} = \alpha \ , \tag{7.5}$$

where λ is the (unique) Lagrange multiplier, and for (§§) give

$$-\frac{b}{\mu} + \nabla V(\hat{y}) = 0 \ . \tag{7.6}$$

Since by assumption $b^T \hat{y} = \alpha$, $\hat{y}$ satisfies (7.5) with the choice of $\lambda = -\frac{1}{\mu}$. Since (§) and (§§) have unique optimal solutions, it follows that $\overline{y} = \hat{y}$. This completes the proof. □

7.2 Limit point of the volumetric path

Let $y^* = \lim_{\mu \to 0} y(\mu)$ be an optimal solution of (7.1) with the corresponding optimal objective value $\alpha^* = b^T y^*$. From Proposition 7.1.3, one sees that as α decreases to α^*, the volumetric centers of the level sets $\mathcal{L}_{\alpha^*}$ converge to y^*. Thus a natural question arises about whether y^* is the volumetric center ofthe optimal level set $\mathcal{L}_{\alpha^*}$. Observe that since certain constraints have to be active in the optimal level set $\mathcal{L}_{\alpha^*}$, the volumetric barrier function $V(y)$ is not defined on $\mathcal{L}_{\alpha^*}$. Hence in order to define the volumetric center of $\mathcal{L}_{\alpha^*}$, one needs to identify the constraints that are inactive at y^*, i.e. the constraints which hold with strict inequality in the relative interior of $\mathcal{L}_{\alpha^*}$. Let I be the set of inactive constraints of $A^T y \leq c$ in the relative interior of the optimal level set $\mathcal{L}_{\alpha^*}$. Let $\overline{F}(y) = -\sum_{i \in I} \log(c_i - a_i^T y)$ and $\overline{V}(y)$ be defined accordingly. The volumetric center of the optimal level set $\mathcal{L}_{\alpha^*}$ is defined as the unique minimizer of

$$
\begin{aligned}
\min \quad & \overline{V}(y) \\
\text{s.t} \quad & b^T y = \alpha^* \\
& a_i^T y = c_i, \;\; i \notin I \\
& a_i^T y < c_i, \;\; i \in I
\end{aligned}
\tag{7.7}
$$

It is known, see e.g. Roos et al. (2006) that for the logarithmic barrier function, the central path converges to the analytic center of the optimal level set. In particular, for a linear optimization problem in the standard form, the volumetric barrier function reduces to the logarithmic barrier function, hence in this case the volumetric path converges to the volumetric center of the optimal level set also. A natural question to ask is whether this extends to the problems in the form of (7.1).

As the following example illustrates, this fact fails to hold for (7.1).

Example 6.2.1.

Let the rows of the matrix $A^T \in \mathbb{R}^{5\times 2}$ be given by the vectors $a_1^T = (-1, 0)$, $a_2^T = (0.1, 1)$, $a_3^T = (1, 0)$, $a_4^T = (0.1, -1)$, $a_5^T = (0, -1)$ with the objective vector $b^T = -(0, 1)$ and $c^T = (0, 1, 1, 0, 0.1)$. The optimal objective value is $\alpha^* = 0.1$. For a polyhedral set of the form $P = \{y : A^T y \le c\}$, Vaidya (1989) showed that the Hessian $H(y) = \nabla^2 F(y)$ of the logarithmic barrier function is computed as $H(y) = \sum_{i=1}^m H_i(y)$, where $H_i(y) = a_i a_i^T/(c_i - a_i^T y)^2$. For our example $n = 5$ and

$$H_1(y) = \frac{1}{y_1^2}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ H_2(y) = \frac{1}{(1-0.1y_1-y_2)^2}\begin{bmatrix} 0.01 & 0.1 \\ 0.1 & 1 \end{bmatrix},\ H_3(y) = \frac{1}{(1-y_1)^2}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

$$H_4(y) = \frac{1}{(-0.1y_1+y_2)^2}\begin{bmatrix} 0.01 & -0.1 \\ -0.1 & 1 \end{bmatrix},\ H_5(y) = \frac{1}{(y_2-0.1)^2}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

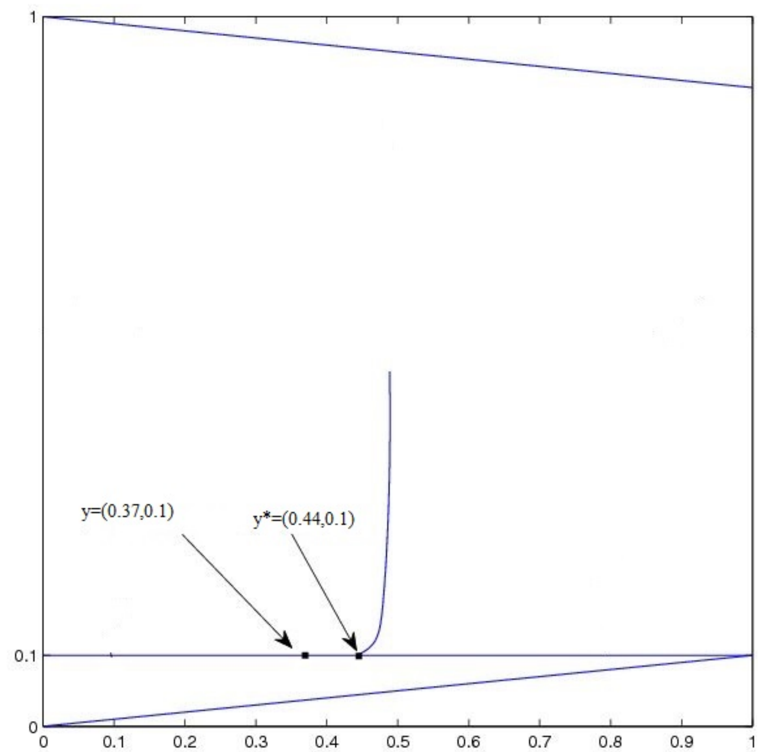

Figure 7.1: The volumetric center of the optimal level set is (0.37,0.1), while the volumetric path converges to (0.44, 0.1).

From Proposition (7.1.3), one can see that the volumetric path converges to

$$y^* = \lim_{\epsilon \to 0} y(\epsilon),$$

where

$$\begin{aligned} y(\epsilon) = \text{argmin} \quad & \text{logdet } H(y) \\ & y_2 = 0.1 + \epsilon. \end{aligned}$$

Now, logdet $H(y)$ is computed as

$$\begin{aligned} &\log[(60000\epsilon^4 y_1^2 - 60000\epsilon^4 y_1 + 30000\epsilon^4 - 64000\epsilon^3 y_1^2 + 64000\epsilon^3 y_1 - 32000\epsilon^3 + 600\epsilon^2 y_1^4 \\ &-1200\epsilon^2 y_1^3 + 26200\epsilon^2 y_1^2 - 25600\epsilon^2 y_1 + 12800\epsilon^2 + 160\epsilon y_1^4 - 3200\epsilon y_1^3 + 6080\epsilon y_1^2 \\ &-4480\epsilon y_1 + 1440\epsilon + 4y_1^6 - 66y_1^5 + 401y_1^4 - 800y_1^3 + 722y_1^2 - 342y_1 + 81)/ \\ &(\epsilon^2 y_1^2 (y_1-1)^2 (10\epsilon - y_1 + 1)^2 (10\epsilon + y_1 - 9)^2)] \ . \end{aligned}$$

Let $h(y_1, \epsilon) = \log(\epsilon^2 \det H(y))$. Clearly for $\epsilon > 0$ fixed, the minimizer of the function log det $H(y)$ is the same as the minimizer of $h(y_1, \epsilon)$. Note that

$$\begin{aligned} \lim_{\epsilon \to 0} h(y_1, \epsilon) &= \log \frac{(4y_1^6 - 66y_1^5 + 401y_1^4 - 800y_1^3 + 722y_1^2 - 342y_1 + 81)}{y_1^2 (y_1-1)^4 (y_1-9)^2} \\ &= \log \frac{(4y_1^4 - 58y_1^3 + 281y_1^2 - 180y_1 + 81)}{y_1^2 (y_1-1)^2 (y_1-9)^2}. \end{aligned}$$

Denote this limit by $g(y_1)$. We will argue that the first coordinate of the limit point of the volumetric path $y^* = \lim_{\epsilon \to 0} y(\epsilon)$ is the minimizer of $g(y_1)$.

First the unique minimizer of $g(y_1)$ can be computed as $y^* = 0.44248$. Let $g_k(y_1) = h(y_1, \frac{1}{k})$. We will show that $\lim_{k\to\infty} \text{argmin } g_k(y_1) = y^*$. Suppose by contradiction that $\lim_{k\to\infty} \text{argmin } g_k(y_1) = \overline{y} \neq y^*$ for some $\overline{y}$. Choose an interval $[a,b] \subseteq [0,1]$ containing y^* such that $\overline{y} \notin [a,b]$. Since $g(y_1)$ has minimum at y^*, one can choose an ϵ with $0 < \epsilon < \min\{\frac{g(a)-g(y^*)}{2}, \frac{g(b)-g(y^*)}{2}\}$. Since $g_k(y_1)$ converges uniformly to $g(y_1)$ on the compact interval $[a,b]$, there exists a number $N \in \mathbb{N}$ such that $k \geq N$ implies $g(y) - \epsilon < g_k(y) < g(y) + \epsilon$ for all $y \in [a,b]$. For $k \geq N$ we have,

$$\begin{aligned} g_k(y^*) < g(y^*) + \epsilon < g(a) - \epsilon < g_k(a) \\ g_k(y^*) < g(y^*) + \epsilon < g(b) - \epsilon < g_k(b). \end{aligned} \tag{7.8}$$

Fix $k \geq N$. If $g_k(y_1)$ had a minimizer y not in $[a,b]$, then (7.8) would imply that the points $g_k(a), g_k(b), g_k(y^*)$ and $g_k(y)$ contradict the convexity property of $g_k(y_1)$. This shows that for any $k \geq N$, the unique minimizer of $g_k(y_1)$ must lie in the interval $[a,b]$. Hence this would be a contradiction to the assumption that $\lim_{k\to\infty} \text{argmin } g_k(y_1) = \overline{y} \notin [a,b]$. Thus we obtain $\lim_{k\to\infty} \text{argmin } g_k(y_1) = y^* = 0.44248$ as the limit point of the volumetric path.

On the other hand, at the optimal objective value $\alpha^* = 0.1$ the constraint a_5 is active, and the volumetric center of the optimal level set defined by (7.7) is the unique solution of

$$\begin{aligned} \min \quad & \text{logdet } \overline{H}(y) \\ & y_2 = 0.1, \end{aligned} \tag{7.9}$$

where $\overline{H}(y) = \sum_{i=1}^{4} H_i(y)$. The optimal solution of (7.9) is computed as
$y = (0.37087, 0.1)$, whose first coordinate is not equal to y^*. Thus this counterexample demonstrates that the limit of the volumetric path is not necessarily the volumetric center of the optimal level set.

Chapter 8

Conclusions and future research

In this final chapter, we review the results of the thesis and highlight future research problems.

Recall that the construction (4.2) in Chapter 4, (see also Figure 4.1) achieves two things: i) The curvature of the new central path $\mathcal{C}_{\overline{P}}$ makes an extra sharp turn, and ii) each extra sharp turn obtained increases the Sonnevend curvature by a constant amount. On the other hand, the $\mathcal{KM}(m)$ construction of Chapter 5 shows that each sharp turn of the central path in a properly chosen neighborhood of the central path generates an extra Newton step, which in turn accounts for an increase in Sonnevend's curvature, see Proposition 5.1.2 and Theorem 5.3.2. These two cases suggests that there might be a close relationship between the geometric curvature and the Sonnevend curvature of the central path. Hence,

Problem 1:

In a general setting, investigate whether Sonnevend's curvature and geometric curvature of the central path are in similar orders of magnitude.

A positive relationship along these lines in a general setting would imply a new type of bound for the total geometric curvature of the central path.

The most significant result of Chapter 5 is that we rigorously show the iteration-complexity upper bound $\mathcal{O}\left(\sqrt{n}\log\left(\frac{\mu_1}{\mu_0}\right)\right)$ for the MTY predictor-corrector algorithm is essentially tight. While the MTY predictor-corrector algorithm is an adaptive-step algorithm, it still follows the central path closely. A natural question is to try to extend this result to long step IPMs algorithms and ask whether the relevant iteration-complexity upper bounds for the long step variants are tight.

Problem 2:

Are the iteration-complexity upper bounds for long step IPMs algorithms tight? Can we use or extend the construction (5.1) as a worst-case LO example for these type of algorithms?

A useful way to view the Sonnevend curvature $\int_0^\infty \frac{\kappa(\mu)}{\mu}d\mu$ is through Grassmann manifold, see Theorem B.2 and Theorem B.3 in Appendix B. In light of these theorems, the problem of estimating the worst-case value of the Sonnevend curvature reduces to the following problem:

$$\max\left\{\int_{t_0}^{t_1}\|Me(I-M)e\|_2^{1/2}dt : M' = h(M),\ M(t_0)\in G(n,m)\right\},$$

where $h(M) = Mdiag(Me) + diag(Me)M - 2Mdiag(Me)M$ and $\mu = e^{-t}$. Note that this problem can be cast in purely in the terms of Grassmann manifold. This suggests that it is at least plausible to expect an upper bound for $\int_0^\infty \frac{\kappa(\mu)}{\mu}d\mu$ independent of a condition number $\overline{\chi}^*_A$, possibly a strongly polynomial bound. Clearly due to Theorem 3.2.1, a new bound for $\int_0^\infty \frac{\kappa(\mu)}{\mu}d\mu$ would lead to important implications about the iteration-complexity of IPMs algorithms. Chapter 6 gives a strongly polynomial bound for $\int_0^\infty \frac{\kappa(\mu)}{\mu}d\mu$ when either $m = 1$ or $m = n - 1$.

At this point, we would like to report a surprising numerical behavior. Let $A_{1\times n} = [1, 1, \ldots, 1, 1+\sqrt{n}]$ and form the matrix $M_1 = A^T(AA^T)^{-1}A$. It is possible to show that we have $\|M_1 e(I - M_1)e\|_2^{1/2} = \Omega(\sqrt{n})$. Now for $m < n$, consider

$$\max_U \left\{\|Me(I - M)e\|_2^{1/2} : M = U^T(UU^T)^{-1}U\right\},$$

where U is an $m \times n$ matrix of full rank. Let $M_2 := U^T(UU^T)^{-1}U$ be the optimal solution of this problem. Note that $\text{rank}(M_1) = 1$, while $\text{rank}(M_2) = m$. Now Theorem B.3 implies that the Sonnevend curvatures $\|M(t)e(I - M(t))e\|_2^{1/2}$ for each cases are completely determined from the initial values M_1 and M_2. Figure 8.1 draws both $\|M_1(t)e(I - M_1(t))e\|_2^{1/2}$ and $\|M_2(t)e(I - M_2(t))e\|_2^{1/2}$ for $m = 2$, $n = 4$ versus the horizontal axis t.

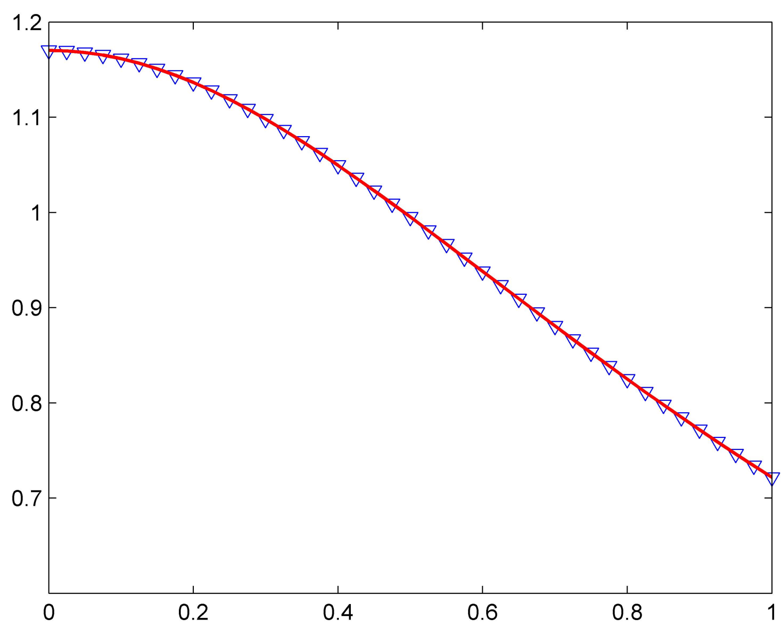

Figure 8.1: The solid curve in the figure results from $M_1(t)$ while the curve given by downward-pointing triangle comes from $M_2(t)$.

As we see the two curves completely overlap. This behavior certainly requires an explanation.

Problem 3:

1. *When* $\|Me(I-M)e\|_2^{1/2}$*, where* $M = A^T(AA^T)^{-1}A$*, is at its global maximum for an* $m \times n$ *matrix of full rank, can the Sonnevend curvature* $\|M(t)e(I - M(t))e\|_2^{1/2}$ *be also expressed with an initial value* $M_0 = \overline{A}^T(\overline{AA}^T)^{-1}\overline{A}$ *for a matrix* $\overline{A}_{1\times n}$*?*

2. *Is part 1. also true for a local maximum of* $\|Me(I-M)e\|_2^{1/2}$*? Or is it true in general?*

Note that whenever one can reduce the estimate of $\int_{t_0}^{t_1} \|Me(I-M)e\|_2^{1/2} dt$ for $\text{rank}(M) = m$ to the case of $\text{rank}(M) = 1$, due to Proposition 6.1.2, we obtain a strongly polynomial bound for the Sonnevend curvature $\int_0^{\infty} \frac{\kappa(\mu)}{\mu} d\mu$. This, in turn (using Theorem 3.2.1), would allow us to categorize the number of iterations in two parts, where the part accounted by the integral $\int_0^{\infty} \frac{\kappa(\mu)}{\mu} d\mu$ has a strongly polynomial bound. This would shed more light on the iteration-complexity of IPMs.

Appendix A

Additional lemmas for Chapter 4 and 5

Lemma A.1. *For large enough r, there is a 1-dimensional LO problem with $(r+1)$ constraints and constants $\tau_1, \tau_2 \geq 0$, for which $\tau_1\sqrt{r} \leq \kappa(\mu) \leq \tau_2\sqrt{r}$ for any $\mu \in \left[\frac{1}{r-\frac{\sqrt{r}}{4}}, \frac{1}{r-\sqrt{r}}\right]$.*

Proof. Consider the problem $\min\{\ y :\ \ y \leq 1$ and, $y \geq 0$ repeated r times$\}$. The construction is given in Sonnevend et al. (1991), p:551. Let $s_0(\mu) = 1 - y(\mu)$. Then it is possible to show that (Sonnevend et al. (1991), p:551), $\dfrac{\dot{s}_0(\mu)}{s_0(\mu)}$ is larger than $\dfrac{r^2}{3\sqrt{r}}$ on the interval $\left[\dfrac{1}{r-\frac{\sqrt{r}}{4}}, \dfrac{1}{r-\sqrt{r}}\right]$ so that $\dfrac{\mu\dot{s}_0(\mu)}{s_0(\mu)} = \Omega(\sqrt{r})$ on $\left[\dfrac{1}{r-\frac{\sqrt{r}}{4}}, \dfrac{1}{r-\sqrt{r}}\right]$. Then from Proposition 3.2.2 part *1.*, we have $\kappa(\mu) = \Omega(\sqrt{r})$ for any $\mu \in \left[\dfrac{1}{r-\frac{\sqrt{r}}{4}}, \dfrac{1}{r-\sqrt{r}}\right]$. □

Proposition A.2. *Consider the LO problems*

$$
\begin{aligned}
\min \quad & (c^1)^T x & \qquad \min \quad & (c^2)^T x \\
\text{s.t.} \quad & A^1 x^1 = b^1 & \qquad \text{s.t.} \quad & A^2 x^2 = b^2 \\
& x^1 \geq 0, & & x^2 \geq 0,
\end{aligned}
\tag{A.1}
$$

with the corresponding $\kappa^1(\mu)$ *and* $\kappa^2(\mu)$ *on the interval* $[\mu_0, \mu_1]$*.*

Then for the problem

$$
\begin{aligned}
\min \quad & c^T x \\
\text{s.t.} \quad & Ax = b \\
& x \geq 0,
\end{aligned}
\tag{A.2}
$$

with the corresponding $\overline{\kappa}(\mu)$*, where* $c = [c^1, c^2]^T$*,* $b = [b^1, b^2]^T$ *and* $A = \begin{bmatrix} A^1 & 0 \\ 0 & A^2 \end{bmatrix}$*.*
Then on the interval $[\mu_0, \mu_1]$*, we have* $\overline{\kappa}(\mu) \geq \kappa^i(\mu)$ *for* $i = 1, 2$*.*

Proof. Let $(x^1(\mu), y^1(\mu), s^1(\mu))$ and $(x^2(\mu), y^2(\mu), s^2(\mu))$ be the central paths in (A.1). Then the term $\overline{\kappa}(\mu)$ for the combined problem (A.2) becomes $\overline{\kappa}(\mu) = ||[\mu \dot{x}^1 \dot{s}^1, \mu \dot{x}^2 \dot{s}^2]||^{\frac{1}{2}} \geq \kappa^i(\mu)$ for $i = 1, 2$ on $[\mu_0, \mu_1]$. □

Proposition A.3. *Let* $\eta > 0$ *and consider the central path* (1.5) *and its* $\kappa(\mu)$*. Let* $(\hat{A}, \hat{b}, \hat{c})$ *be another problem instance, where* $(\hat{A}, \hat{b}, \hat{c}) = (A, \frac{b}{\eta}, c)$ *with its corresponding* $\hat{\kappa}(\mu)$*. Then, we have*

$$
\hat{\kappa}(\mu) = \kappa(\eta\mu), \ \mu \in \left[\frac{\mu_0}{\eta}, \frac{\mu_1}{\eta}\right]. \tag{A.3}
$$

Proof. Using (1.5), it is straightforward to verify that the central path $(\hat{x}(\mu), \hat{y}(\mu), \hat{s}(\mu))$ of the new problem satisfies $\hat{x}(\mu) = \dfrac{x(\eta\mu)}{\eta}$, $\hat{y}(\mu) = y(\eta\mu)$ and $\hat{s}(\mu) = s(\eta\mu)$. Using the definition of $\kappa(\mu)$, we get $\hat{\kappa}(\mu) = \kappa(\eta\mu)$. Hence the claim follows. □

Lemma A.4. *Given an interval $[\mu_0, \mu_1]$ and a constant $\nu > 0$, there exists an LO problem of size $n = \Theta\left(\log(\frac{\mu_1}{\mu_0})\right)$ such that for all $\mu \in [\mu_0, \mu_1]$, we have $\overline{\kappa}(\mu) \geq \nu$. The hidden constant in $n = \Theta\left(\log(\frac{\mu_1}{\mu_0})\right)$ depends on ν.*

Proof. Let a constant $\nu > 0$ and an interval $[\mu_0, \mu_1]$ be given. For the given $\nu > 0$, by Lemma A.1, there exists an LO problem with its $\kappa(\mu) \geq \nu$ on an interval $\mu \in [\alpha_1, \alpha_2]$. By applying Proposition A.3 for $\eta := \dfrac{\alpha_1}{(\dfrac{\alpha_2}{\alpha_2})^i \mu_0}$ for $i = 0, 1, \ldots, k$, we find $k-1$ scaled LO problems with their corresponding $\kappa^i(\mu)$, $i = 0, 1, \ldots, k-1$ such that $\kappa^i(\mu) = \kappa(\eta\mu)$ on $\mu \in [(\frac{\alpha_2}{\alpha_1})^i \mu_0, (\frac{\alpha_2}{\alpha_1})^{i+1} \mu_0]$, for $i = 0, 1, \ldots, k-1$. Then by using Proposition A.2, we can obtain a block diagonal LO problem with its $\overline{\kappa}(\mu) \geq \kappa^i(\mu) \geq \nu$ for $i = 0, 1, \ldots, k-1$ for any $\mu \in \left[\mu_0, (\dfrac{\alpha_2}{\alpha_1})^k \mu_0\right]$. In order to have $\overline{\kappa}(\mu) \geq \nu$ for any $\mu \in [\mu_0, \mu_1]$, it is then enough to have $(\dfrac{\alpha_2}{\alpha_1})^k \mu_0 \geq \mu_1$. This is true if and only if $k \log(\frac{\alpha_2}{\alpha_1}) \geq \log(\frac{\mu_1}{\mu_0})$. Since the ratio $\frac{\alpha_2}{\alpha_1}$ is a constant depending only on the given ν, the number of blocks k needed is $\Theta(\log(\frac{\alpha_2}{\alpha_1}))$. Also since the size of the LO problem with its $\kappa(\mu)$ is a constant which is determined only by ν, the size of the problem $n = \Theta\left(\log(\frac{\mu_1}{\mu_0})\right)$ to achieve $\overline{\kappa}(\mu) \geq \nu$ for all $\mu \in [\mu_0, \mu_1]$. This completes the proof. $\square$

Appendix B

Sonnevend's curvature and Grassmann manifold

Definition B.1. *The Grassmann manifold is defined as*

$$G(n,m) := \{M \in \mathbb{R}^{n\times n} : M^T = M, M^2 = M, \ \mathrm{rank}(M) = m\}.$$

Theorem B.2. Sonnevend et al. (1991) *Consider the central path in* (1.5) *and for any* $\mu > 0$ *let* $M = S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1}$. *Then*

1. $M^T = M$, $M^2 = M$, $\|M\|_2 = 1$ *and* $\mathrm{rank}(M) = m$.

2. $\kappa(\mu) = \|Me(I - M)e\|_2^{1/2}$.

3. $u \in R(AS^{-1})$ *if and only if* $Mu = u$.

4. $u \in N(AS^{-1})$ *if and only if* $Mu = 0$.

5. *For any* $M \in G(n,m)$, *there exists a matrix* $\overline{A} \in \mathbb{R}^{m\times n}$ *of full rank, such that* $M = \overline{A}^T(\overline{A}\overline{A}^T)^{-1}\overline{A}$.

Theorem B.3. Sonnevend et al. (1991) *Define a parametrization* t *on* $(-\infty, \infty)$ *such that* $e^{-t} = \mu$. *Then* $M(t) = S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1}$ *satisfies*

$$\frac{dM}{dt} = h(M),$$

where $h(M) = Mdiag(Me) + diag(Me)M - 2Mdiag(Me)M$. *For any given projection matrix* $M(t_0)$, *this differential equation determines* $M(t)$ *for all* t.

Bibliography

Allamigeon, X., Benchimol, P., Gaubert, S., and Joswig, M. (2014). Long and winding central paths. *arXiv preprint arXiv:1405.4161*.

Anstreicher, K. M. (1996). Large step volumetric potential reduction algorithms for linear programming. *Annals of Operations Research*, 62(1):521–538.

Anstreicher, K. M. (1997). On vaidya's volumetric cutting plane method for convex programming. *Mathematics of Operations Research*, 22(1):63–89.

Anstreicher, K. M. (1998). Towards a practical volumetric cutting plane method for convex programming. *SIAM Journal on Optimization*, 9(1):190–206.

Anstreicher, K. M. (2000). The volumetric barrier for semidefinite programming. *Mathematics of Operations Research*, 25(3):365–380.

Atkinson, D. S. and Vaidya, P. M. (1995). A cutting plane algorithm for convex programming that uses analytic centers. *Mathematical Programming*, 69(1-3):1–43.

Bland, R. G., Goldfarb, D., and Todd, M. J. (1981). The ellipsoid method: a survey. *Operations Research*, 29(6):1039–1091.

Dantzig, G. B. (1965). *Linear Programming and Extensions.* Princeton University Press.

De Loera, J. A., Sturmfels, B., and Vinzant, C. (2012). The central curve in linear programming. *Foundations of Computational Mathematics*, 12(4):509–540.

Dedieu, J.-P., Malajovich, G., and Shub, M. (2005). On the curvature of the central path of linear programming theory. *Foundations of Computational Mathematics*, 5(2):145–171.

Deza, A., Nematollahi, E., Peyghami, R., and Terlaky, T. (2006). The central path visits all the vertices of the Klee–Minty cube. *Optimisation Methods and Software*, 21(5):851–865.

Deza, A., Nematollahi, E., and Terlaky, T. (2008a). How good are interior point methods? Klee–Minty cubes tighten iteration-complexity bounds. *Mathematical Programming*, 113(1):1–14.

Deza, A., Terlaky, T., and Zinchenko, Y. (2008b). Polytopes and arrangements: Diameter and curvature. *Operations Research Letters*, 36(2):215 – 222.

Deza, A., Terlaky, T., and Zinchenko, Y. (2009). A continuous d-step conjecture for polytopes. *Discrete & Computational Geometry*, 41:318–327.

Deza, A. and Xie, F. (2007). Hyperplane arrangements with large average diameter. *Polyhedral Computation*, 48.

Karmarkar, N. (1984). A new polynomial-time algorithm for linear programming. In *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, pages 302–311. ACM.

Khachiyan, L. G. (1979). A polynomial algorithm in linear programming. In *Sov. Math., Dokl.*, volume 20, pages 191–194.

Khachiyan, L. G. (1980). Polynomial algorithms in linear programming. *USSR Computational Mathematics and Mathematical Physics*, 20(1):53–72.

Klee, V. and Minty, G. J. (1972). How good is the simplex algorithm? *Inequalities: Proceedings*, 3:158–172.

Klee, V. and Walkup, D. W. (1967). The d-step conjecture for polyhedra of dimension $d < 6$. *Acta Mathematica*, 117:53–78.

Mizuno, S., Todd, M. J., and Ye, Y. (1993). On adaptive-step primal-dual interior-point algorithms for linear programming. *Mathematics of Operations Research*, 18(4):964–981.

Monteiro, R. D. and Tsuchiya, T. (2005). A new iteration-complexity bound for the mty predictor-corrector algorithm. *SIAM Journal on Optimization*, 15(2):319–347.

Monteiro, R. D. and Tsuchiya, T. (2008). A strong bound on the integral of the central path curvature and its relationship with the iteration-complexity of primal-dual path-following LP algorithms. *Mathematical Programming*, 115:105–149.

Mut, M. and Terlaky, T. (2012). On the volumetric path. *Optimization Letters*, 6(4):687–693.

Nematollahi, E. and Terlaky, T. (2008a). A redundant Klee–Minty construction with all the redundant constraints touching the feasible region. *Operations Research Letters*, 36(4):414–418.

Nematollahi, E. and Terlaky, T. (2008b). A simpler and tighter redundant Klee–Minty construction. *Optimization Letters*, 2(3):403–414.

Nemirovski, A. S. and Yudin, D. B. (1976). Informational complexity and effective methods for the solution of convex extremal problems. *Èkon. Mat. Metody*, 12:357–369.

Nesterov, Y. and Nemirovskii, A. (1994). *Interior-Point Polynomial Algorithms in Convex Programming*, volume 13. SIAM, Philadelphia.

Ohara, A. and Tsuchiya, T. (2007). An information geometric approach to polynomial-time interior-point algorithms: complexity bound via curvature integral. *Research Memorandum, The Institute of Statistical Mathematics*, 1055.

Potra, F. A. (1994). A quadratically convergent predictor–corrector method for solving linear programs from infeasible starting points. *Mathematical Programming*, 67(1-3):383–406.

Renegar, J. (1987). *A Mathematical View of Interior-Point Methods in Convex Optimization*, volume 3. SIAM, Philadelphia.

Renegar, J. (1988). A polynomial-time algorithm, based on newton's method, for linear programming. *Mathematical Programming*, 40(1-3):59–93.

Roos, C., Terlaky, T., and Vial, J.-P. (2006). *Interior Point Methods for Linear Optimization*. Springer Science, Heidelberg/Boston.

Roos, C. and Vial, J.-P. (1988). Analytic centers in linear programming. Technical report, Faculty of Mathematics and Computer Science, TU Delft, NL–2628 BL Delft, The Netherlands.

Shor, N. Z. (1972). Utilization of the operation of space dilatation in the minimization of convex functions. *Cybernetics and Systems Analysis*, 6(1):7–15.

Sonnevend, G. (1985). An analytic center for polyhedrons and new classes of global algorithms for linear (smooth, convex) optimization. Technical report, Dept. of Numerical Analysis, Institute of Mathematics, Eötvös University, 1088, Budapest, Múzeum Körút 6-8, 1985.

Sonnevend, G., Stoer, J., and Zhao, G. (1991). On the complexity of following the central path of linear programs by linear extrapolation II. *Mathematical Programming*, 52:527–553.

Stoer, J. and Zhao, G. (1993). Estimating the complexity of a class of path-following methods for solving linear programs by curvature integrals. *Applied Mathematics and Optimization*, 27:85–103.

Todd, M. J. (1993). A lower bound on the number of iterations of primal-dual interior-point methods for linear programming. In *15th Biennial Conference on Numerical Analysis at Dundee, Numerical Analysis 1993 (G.A. Watson and D.F. Griffiths, eds.), Pitman Research Notes in Mathematics 303*, pages 237–259. Longman Press.

Todd, M. J. and Ye, Y. (1996). A lower bound on the number of iterations of long-step primal-dual linear programming algorithms. *Annals of Operations Research*, 62(1):233–252.

Vaidya, P. M. (1989). A new algorithm for minimizing convex functions over convex sets. In *SFCS '89: Proceedings of the 30th Annual Symposium on Foundations of Computer Science*, pages 338–343, Washington, DC, USA. IEEE Computer Society.

Vavasis, S. A. and Ye, Y. (1996). A primal-dual interior point method whose running time depends only on the constraint matrix. *Mathematical Programming*, 74:79–120.

Zhao, G. (2010). Representing the space of linear programs as the Grassmann manifold. *Mathematical Programming*, 121:353–386.

Index