An analogue of the Klee-Walkup result for Sonnevend's curvature of the central path

Murat Mut   Tamás Terlaky

Department of Industrial and Systems Engineering
Lehigh University
Bethlehem, PA

April 7, 2013

Abstract

For linear optimization (LO) problems, we consider a curvature integral first introduced by Sonnevend et al. (1991). Our main result 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 $n = 2m$. This also implies that the worst cases of LO problems for path-following algorithms can be reconstructed for 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.

1 Introduction

We first introduce our notation and recall the basics of interior-point methods (IPM).
Let A be a $m \times n$ matrix of full rank. For $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$, consider the primal and dual linear optimization (LO) problems,

$$\begin{array}{llll} \min & c^T x & & \\ \text{s.t.} & Ax & = b & \\ & x & \geq 0, & \end{array} \qquad \begin{array}{lll} \max & b^T y & \\ \text{s.t.} & A^T y + s & = c \\ & s & \geq 0, \end{array} \tag{1}$$

where $x, s \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ are vectors of variables. The sets of feasible solutions in (1) are referred to as primal and dual sets, respectively. Provided that the LO problems in (1) have strictly positive primal and dual solutions, the central path $(x(\mu), y(\mu), s(\mu))$, $\mu > 0$ exists and satisfies the equations:

$$\begin{aligned} Ax &= b, \quad x > 0, \\ A^T y + s &= c, \quad s > 0, \\ xs &= \mu e, \end{aligned} \tag{2}$$

where uv denotes $[u_1 v_1, \ldots, u_n v_n]^T$ for $u, v \in \mathbb{R}^n$, and e is the all-one vector.
It is well-known that as $\mu \to 0$, the points $(x(\mu), y(\mu), s(\mu))$ converge to a primal and dual optimal solution for (1). Path-following algorithms traverse the central path with a certain proximity measure until a primal-dual optimal solution for (1) is reached.
Our main focus in this paper is a curvature integral $\displaystyle\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu$, where $\kappa(\mu) = \|\mu \dot{x} \dot{s}\|^{1/2}$, which was first introduced by Sonnevend et al. [7]. This integral will be referred to as the *Sonnevend curvature*. In terms of this curvature integral, Stoer et al. [8] rigorously provided a complexity bound for an algorithm, which is a variant of what is now known as the Mizuno-Todd-Ye predictor-corrector (MTY P-C) algorithm.
Our main result for the Sonnevend curvature $\displaystyle\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu$ 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. [3]. 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 paper, 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 [7], 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.
The idea of using a sequence of polytopes whose size and dimension increase by one was first used by Klee-Walkup [4] in the context of diameter of a polytope. The diameter of a polytope is the maximum of the shortest edge path's between any two 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 [4], 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 rest of the paper is organized as follows. In Section 2 we give the background information for Sonnevend's curvature $\int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu$. In Section 3, we present our main results and conclude the paper with further remarks in Section 4.

2 The Sonnevend curvature of the central path

Sonnevend's curvature is closely related to the iteration-complexity of a variant of the MTY predictor-corrector algorithm which was introduced in [7]. Let $\kappa(\mu) = \|\mu \dot{x} \dot{s}\|^{1/2}$. Stoer et al. [8] proved that their predictor-corrector algorithm has a complexity bound, which can be expressed in terms of $\kappa(\mu)$.

Theorem 2.1. *Let N be the number of iterations of Algorithm 2.1 [8] to reduce the barrier parameter from μ_1 to μ_0. Then*

$$C_3 \int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu - 1 \leq N \leq C_1 \int_{\mu_0}^{\mu_1} \frac{\kappa(\mu)}{\mu} d\mu + C_2 \log\left(\frac{\mu_1}{\mu_0}\right) + 2 \tag{3}$$

for some "universal" constants C_1, C_2, and C_3 that depend only on the neighborhood of the central path.

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

Proposition 2.2. *[7, 10] The following holds.*

1. *We have* $\kappa(\mu) = \left\| \frac{\mu \dot{s}(\mu)}{s(\mu)} - \left(\frac{\mu \dot{s}(\mu)}{s(\mu)}\right)^2 \right\|^{\frac{1}{2}}$.

2. *We have* $\frac{\mu \dot{s}(\mu)}{s(\mu)} = Me$, *where* $M = S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1}$ *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* $\left\| \frac{\mu \dot{s}(\mu)}{s(\mu)} \right\| \leq \sqrt{n}$ *and* $\kappa(\mu) \leq \sqrt{n}$ *implying that*

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

Monterio et al. [5] 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

$$\bar{\chi}_A := \sup_D \{\|A^T(ADA^T)^{-1}AD\|\}, \tag{4}$$

where D ranges over the set of positive diagonal matrices. It is known that for matrix with integer entries ([9], Lemma 24), $\log(\bar{\chi}_A) = L_A$, where L_A is the input bit length of A. Then we have the following bound for Sonnevend's curvature.

Theorem 2.3. *[5] We have*

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

Theorem 2.3 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 2.4. *Given $A \in \mathbb{R}^{m \times n}$, define*

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

3 Main results

In this section, we introduce the construction used in [3]. First assume $n > 2m$. We will later reduce the case $m < n < 2m$ to this case. Consider the LO problem

$$\max\{b^T y : y \in \mathcal{P}\}, \text{ where } \mathcal{P} = \{y \in \mathbb{R}^m : A^T y \le c\} \text{ is a polytope.} \tag{5}$$

Without loss of generality, we may assume that:

(A1) The analytic center y^* of $\mathcal{P}$ is the origin, and

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

First, whenever $y^* \neq 0$, we can always shift a general polytope $\mathcal{P}$ with the transformation $A^T y \le c - A^T y^*$ 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 assuming $y^* = 0$, the analytic center being an interior point in $\mathcal{P}$ implies that $c > 0$. Second, if we rescale our LO problem with a positive diagonal matrix as

$$\begin{array}{lrl} \min & (Dc)^T x & \\ \text{s.t.} & ADx & = b \\ & x & \ge 0\ , \end{array} \qquad \begin{array}{lrl} \max & b^T y & \\ \text{s.t.} & DA^T y + Ds & = Dc \\ & s & \ge 0\ ; \end{array}$$

then the rescaled central path becomes $(\bar{x}(\mu), \bar{y}(\mu), \bar{s}(\mu)) = \big(D^{-1}x(\mu), y(\mu), Ds(\mu)\big)$ implying that $\kappa(\mu)$ does not change. Since $c > 0$ by assumption, by choosing D with $De = c^{-1}$, we can make $c = e$.

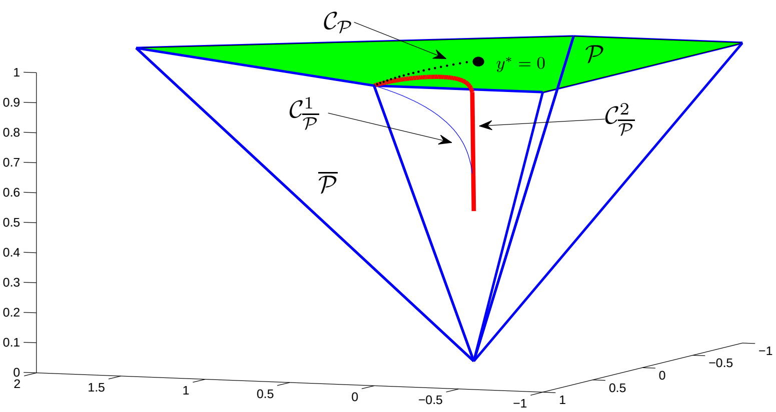

Figure 1: The dotted path $\mathcal{C}_\mathcal{P}$ is the central path of the original polytope $\mathcal{P}$. The figure shows how the central path $\mathcal{C}_{\overline{\mathcal{P}}}$ is changing with θ. A smaller θ_1 leads to the path $\mathcal{C}^1_{\overline{\mathcal{P}}}$, while $\mathcal{C}^2_{\overline{\mathcal{P}}}$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

We now associate problem (5) with a sequence of LO problems parameterized by $\theta > 0$ as follows:

$$\begin{aligned} \max \quad & b^T y + \theta z \\ \begin{bmatrix} A^T & -e_{n\times 1} \\ 0_{1\times m} & 1 \end{bmatrix} \begin{bmatrix} \bar{y} \\ z \end{bmatrix} + \begin{bmatrix} \bar{s} \\ \bar{s}_{n+1} \end{bmatrix} &= \begin{bmatrix} 0_{n\times 1} \\ 1 \end{bmatrix} \\ \bar{s}, \bar{s}_{n+1} &\geq 0. \end{aligned} \tag{6}$$

The feasible set for the problem (6) can be written as $\overline{\mathcal{P}} = \{A^T\bar{y} \leq ze,\ z \leq 1\}$.
Let $\bar{A} = \begin{bmatrix} A & 0_{m\times 1} \\ -e_{1\times n} & 1 \end{bmatrix}$. The associated central path equations for (6) are

$$A^T \frac{\bar{y}(\mu)}{z(\mu)} + \frac{\bar{s}(\mu)}{z(\mu)} = e, \qquad A\bar{s}(\mu)^{-1} = \frac{b}{\mu}, \tag{7}$$

$$\frac{1}{\bar{s}_{n+1}(\mu)} = \frac{1}{1 - z(\mu)} = \sum_{i=1}^{n} \frac{1}{\bar{s}_i(\mu)} + \frac{\theta}{\mu}. \tag{8}$$

Note that $\bar{y}$, $\bar{s}$ and $\bar{z}$ in (7) and (8) are functions of both μ and θ. We will usually drop θ or μ, when no confusion arises.

Denote the central path of $\mathcal{P}$ and $\overline{\mathcal{P}}$ by $\mathcal{C}_\mathcal{P}$ and $\mathcal{C}_{\overline{\mathcal{P}}}$, respectively.

Intuitively a large θ should force $z \cong 1$ in such a way that, the central path $\mathcal{C}_{\overline{\mathcal{P}}}$ first follows an almost straight line from the analytic center to the face $\mathcal{P} \times \{1\}$ and then stays close to the central path $\mathcal{C}_\mathcal{P}$. The following proposition, first proved in [3], shows that this is indeed the case.

Proposition 3.1. *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* $\bar{y}(\mu) \to y(\mu)$ *uniformly;*

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

Proof. Claim *1.* is the same as Proposition 2.1 in [3] (see also the remark following it). Statement *2.* follows from the first part since $\bar{s}_{n+1}(\mu) = 1 - z(\mu)$ and $\bar{s}(\mu) = z(\mu) - A^T\bar{y}(\mu)$. □

The following proposition shows that if $z(\mu)$ in (7) and (8) is known, then $\bar{y}(\mu)$ is completely determined by the central path $\mathcal{C}_{\mathcal{P}}$.

Proposition 3.2. *Let $z(\mu)$ satisfy the central path equations (7),(8). Then*

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

Proof. Direct substition into (7), with the choice of $\mu' = \dfrac{\mu}{z(\mu)}$, shows that the vectors $\bar{y}(\mu) = z(\mu) y\left(\dfrac{\mu}{z(\mu)}\right)$ and $\bar{s}(\mu) = z(\mu) s\left(\dfrac{\mu}{z(\mu)}\right)$ satisfy the equations (7), which are the central path equations for (5). Since the solution is unique, the claim follows. □

Note that Proposition 3.1 and Proposition 3.2 show that for a fixed interval $[\mu_0, \mu_1]$, parameter θ can be chosen large enough so that the central paths $\mathcal{C}_{\mathcal{P}}$ and $\mathcal{C}_{\overline{\mathcal{P}}}$ become close to each other on that interval. Hence, it is natural to expect that Sonnevend's curvature for $\mathcal{C}_{\mathcal{P}}$ and $\mathcal{C}_{\overline{\mathcal{P}}}$ on the same interval should have similar order of magnitudes.

Proposition 3.3. *Let $\bar{\kappa}(\mu)$ correspond to the central path $\mathcal{C}_{\overline{\mathcal{P}}}$. Then, on the fixed interval $[\mu_0, \mu_1]$, we have $\dot{\bar{s}}(\mu) \to \begin{bmatrix} \dot{s}(\mu) \\ 0 \end{bmatrix}$ uniformly as $\theta \to \infty$. Consequently, as $\theta \to \infty$, $\bar{\kappa}(\mu) \to \kappa(\mu)$ on $[\mu_0, \mu_1]$ uniformly as well.*

Proof. It is well-known, see [6] e.g., that for system (2), we have $\dot{s} = \dfrac{1}{\mu} A^T(AS^{-2}A^T)^{-1}AS^{-1}e$. Now we calculate $U := \bar{A}\begin{bmatrix} \bar{S}^{-1} & 0 \\ 0 & s_{n+1}^{-1} \end{bmatrix} = \begin{bmatrix} A\bar{S}^{-1} & 0 \\ -\bar{s}^{-1} & (s_{n+1})^{-1} \end{bmatrix}$, which gives

$$UU^T = \begin{bmatrix} A\bar{S}^{-2}A^T & , & -A\bar{s}^{-1} \\ (-A\bar{s}^{-1})^T & , & e^T\bar{s}^{-2} + \frac{1}{\bar{s}_{n+1}^2} \end{bmatrix}. \tag{9}$$

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

$$(UU^T)^{-1} = \begin{bmatrix} (A\bar{S}^{-2}A^T)^{-1} + \frac{W_1}{r} & , & \frac{W_2}{r} \\ (\frac{W_2}{r})^T & , & \frac{1}{r} \end{bmatrix}, \tag{10}$$

where $r = e^T\bar{s}^{-2} + \frac{1}{\bar{s}_{n+1}^2} - (A\bar{s}^{-1})^T(A\bar{S}^{-2}A^T)^{-1}A\bar{s}^{-1}$, $W_2 = (A\bar{S}^{-2}A^T)^{-1}A\bar{s}^{-1}$, and $W_1 = W_2W_2^T$. Then, since $\bar{s} \to s$ as $\theta \to \infty$, it follows that the terms W_1 and W_2 converge to finite limits that are only determined by (5). Then, in terms of $\bar{s}_{n+1}$, we get $\frac{1}{r} = \mathcal{O}(\bar{s}_{n+1}^2)$. Thus, we conclude that

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

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

$$\begin{aligned} (UU^T)^{-1}\bar{A}\begin{bmatrix} \bar{s}^{-1} \\ \bar{s}_{n+1}^{-1} \end{bmatrix} &= (UU^T)^{-1}\begin{bmatrix} A\bar{s}^{-1} \\ -e^T\bar{s}^{-1} + \bar{s}_{n+1}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} (A\bar{S}^{-2}A^T)^{-1}A\bar{s}^{-1} + \mathcal{O}(\bar{s}_{n+1}) \\ \mathcal{O}(\bar{s}_{n+1}) \end{bmatrix}. \end{aligned} \tag{11}$$

Finally, from (11), we obtain

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

Taking the limit in θ, we get

$$\dot{\bar{s}}(\mu) = \frac{1}{\mu}\bar{A}^T(UU^T)^{-1}\bar{A}\begin{bmatrix} \bar{s}^{-1} \\ \bar{s}_{n+1}^{-1} \end{bmatrix} \to \begin{bmatrix} \frac{1}{\mu}A^T(AS^{-2}A^T)^{-1}As^{-1} \\ 0 \end{bmatrix} = \begin{bmatrix} \dot{s} \\ 0 \end{bmatrix}. \tag{12}$$

Since from Proposition 2.2, all the terms in $\bar{\kappa}(\mu)$ converge uniformly, we conclude that $\bar{\kappa}(\mu) \to \kappa(\mu)$ uniformly on $[\mu_0, \mu_1]$ as $\theta \to \infty$.

□

Corollary 3.4. *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 $\mathcal{C}_{\overline{\mathcal{P}}}$. Then, for any $\epsilon > 0$, there is a LO problem of size $(m+1, n+1)$ with the Sonnevend curvature*

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

Proof. By Proposition 3.3, we can choose a θ large enough so that $\kappa(\mu)$ and $\bar{\kappa}(\mu)$ is arbitrarily close to each other on $[\mu_0, \mu_1]$. Hence the claim follows. □

We proved that on a fixed interval $[\mu_0, \mu_1]$, one can always make the Sonnevend curvature of $\mathcal{C}_{\overline{\mathcal{P}}}$ and $\mathcal{C}_{\mathcal{P}}$ 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 $\mathcal{C}_{\mathcal{P}}$ stays small on $[\mu_1, \mu_2]$, it can be made as large as a constant for $\mathcal{C}_{\overline{\mathcal{P}}}$ on the same interval by increasing θ. To this end, the following proposition provides important tools. First, we need 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 μ and θ.

Proposition 3.5. *As $\mu \to \infty$ one has,*

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

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

Proof.

1. From Proposition 3.2, we have $\bar{s}(\mu, \theta) = z(\mu, \theta) \left(e - A^T y(\frac{\mu}{z(\mu, \theta)}) \right)$. Since by assumption, the analytic center of $\mathcal{P}$ is $y^* = 0$, we have $y(\mu) \to 0$ as $\mu \to \infty$. This proves the claim.

2. Since the analytic center of $\mathcal{P}$ is $y^* = 0$, we conclude $s_i(\mu, \theta) \leq n$ for large μ with $i = 1, \ldots, n$. From (8) and Proposition 3.2, we have

$$\frac{1}{1 - z(\mu, \theta)} - \frac{1}{z(\mu, \theta)} \left(\sum_{i=1}^{n} \frac{1}{s_i(\frac{\mu}{z(\mu,\theta)})} \right) = \frac{\theta}{\mu} > 0, \tag{13}$$

which implies

$$z(\mu, \theta) > \frac{\displaystyle\sum_{i=1}^{n} \frac{1}{s_i(\frac{\mu}{z(\mu,\theta)})}}{1 + \displaystyle\sum_{i=1}^{n} \frac{1}{s_i(\frac{\mu}{z(\mu,\theta)})}}. \tag{14}$$

Since $s_i(\mu, \theta) \leq n$ for large μ, $i = 1, \ldots, n$, (14) yields $\displaystyle\sum_{i=1}^{n} \frac{1}{s_i(\mu, \theta)} \geq 1$ for large μ. Then from (14), and using the fact that $z(\mu, \theta) \leq 1$, we obtain $z(\mu, \theta) > \frac{1}{2}$ for large μ.

3. Let $i = 1, \ldots, n$. Differentiating the equation $\bar{s}_i(\mu, \theta) = z(\mu, \theta) s_i \left(\frac{\mu}{z(\mu, \theta)} \right)$ from Proposition 3.2 gives,

$$\dot{\bar{s}}_i(\mu, \theta) = \dot{z}(\mu, \theta) s \left(\frac{\mu}{z(\mu, \theta)} \right) + z(\mu, \theta) \dot{s}_i \left(\frac{\mu}{z(\mu, \theta)} \right) \left(\frac{1}{z(\mu, \theta)} - \frac{\mu \dot{z}(\mu, \theta)}{z(\mu, \theta)^2} \right). \tag{15}$$

Using (15),

$$\frac{\mu \dot{\bar{s}}_i(\mu, \theta)}{\bar{s}_i(\mu, \theta)} = \frac{\mu \dot{z}(\mu, \theta)}{z(\mu, \theta)} + \frac{\mu \dot{s}_i \left(\frac{\mu}{z(\mu,\theta)} \right)}{s_i \left(\frac{\mu}{z(\mu,\theta)} \right)} \left(\frac{1}{z(\mu, \theta)} - \frac{\mu \dot{z}(\mu, \theta)}{z(\mu, \theta)^2} \right). \tag{16}$$

Proposition 2.2 part *2.* implies that $\dfrac{\mu \dot{s}_i \left(\frac{\mu}{z(\mu,\theta)} \right)}{s_i \left(\frac{\mu}{z(\mu,\theta)} \right)} \to 0$ as $\mu \to \infty$. Also Proposition 2.2 part *3.* implies that

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

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 (16) that, $\dfrac{\mu \dot{\bar{s}}_i(\mu, \theta)}{\bar{s}_i(\mu, \theta)} \to \dfrac{\mu \dot{z}(\mu, \theta)}{z(\mu, \theta)}$ for $i = 1, \ldots, n$ as $\mu \to \infty$. Note also that all the bounds come from problem (5), and therefore independent of θ. This proves that the bounds in statements *1.* and *3.* are uniform in θ.

$\square$

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

Lemma 3.6. *As $\mu \to \infty$, we have*

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

Moreover the bound is uniform in θ.

Proof. From (8), we have $\bar{s}_{n+1}(\mu,\theta) = \dfrac{1}{\dfrac{\theta}{\mu} + \sum_{i=1}^{n} \dfrac{1}{\bar{s}_i(\mu,\theta)}}$. Then one has

$$\log(\bar{s}_{n+1}(\mu,\theta)) = -\log\left(\frac{\theta}{\mu} + \sum_{i=1}^{n} \frac{1}{\bar{s}_i(\mu,\theta)}\right). \tag{17}$$

By differentiating (17) and multiplying by μ, we get

$$\frac{\mu\dot{\bar{s}}_{n+1}(\mu,\theta)}{\bar{s}_{n+1}(\mu,\theta)} = \frac{\dfrac{\theta}{\mu} + \sum_{i=1}^{n} \dfrac{\mu\dot{\bar{s}}_i(\mu,\theta)}{\bar{s}_i(\mu,\theta)^2}}{\dfrac{\theta}{\mu} + \sum_{i=1}^{n} \dfrac{1}{\bar{s}_i(\mu,\theta)}}. \tag{18}$$

Substituting $\bar{s}_{n+1}(\mu,\theta) = 1 - z(\mu,\theta)$ in (18) and using parts *1.* and *3.* in Proposition 3.5, as $\mu \to \infty$, we can write;

$$-\frac{\mu\dot{z}(\mu,\theta)}{1 - z(\mu,\theta)} = \frac{\dfrac{\theta}{\mu} + \dfrac{n\mu\dot{z}(\mu,\theta) + o(1)}{z(\mu,\theta)^2 + o(1)}}{\dfrac{\theta}{\mu} + \dfrac{n}{z(\mu,\theta)} + o(1)}. \tag{19}$$

Rearranging the terms in (19), we have

$$-\mu\dot{z}(\mu,\theta) = \frac{\dfrac{(1 - z(\mu,\theta))\theta}{\mu} + (1 - z(\mu,\theta))\left(\dfrac{n\mu\dot{z}(\mu,\theta) + o(1)}{z(\mu,\theta)^2 + o(1)}\right)}{\dfrac{\theta}{\mu} + \dfrac{n}{z(\mu,\theta)} + o(1)}. \tag{20}$$

To solve (20) 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,

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

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

which proves the claim. Moreover, since all the bounds come from Proposition 3.5, the bound is uniform in θ. This concludes the proof. □

Corollary 3.7. *There exists a* $\tau \geq \frac{\sqrt{19}}{40}\log 2$ *such that*

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

Proof. Let $\epsilon > 0$. Since $\int_0^\infty \frac{\kappa(\mu)}{\mu}d\mu$ is finite by Theorem 2.3, one can find a μ_0 and a 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 3.6, we can also choose a μ_1 such that

$$\left|\frac{\mu\dot{\bar{s}}_{n+1}(\mu,\theta)}{\bar{s}_{n+1}(\mu,\theta)} - \frac{\dfrac{\theta}{\mu}}{\dfrac{\theta}{\mu}+\dfrac{n}{z(\mu,\theta)^2}}\right| \leq \frac{1}{30} \tag{21}$$

for $\mu \geq \mu_1$ 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 $\int_{\mu_0}^{\mu_1} \frac{\bar{\kappa}(\mu)}{\mu}d\mu \geq v - \epsilon$ are satisfied. Note that Corollary 3.4 implies that such a θ' exists. Since by Proposition 3.5, part *2.*, we have $\frac{1}{2} \leq z(\mu,\theta') \leq 1$, it follows that $n \leq \frac{n}{z(\mu,\theta')^2} \leq 4n$ for $n \geq 2$. 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 $\frac{n}{2} \leq \frac{\theta'}{\mu} \leq n$ and $n \leq \frac{n}{z(\mu,\theta')^2} \leq 4n$, which together implies that

$$\frac{1}{10} \leq \frac{\dfrac{\theta'}{\mu}}{\dfrac{\theta'}{\mu}+\dfrac{n}{z(\mu,\theta')^2}} \leq \frac{2}{3} \tag{22}$$

Then for $\mu \in [\mu_2, 2\mu_2]$, (21) and (22) together imply $\frac{1}{20} \leq \frac{\mu\dot{\bar{s}}_{n+1}(\mu,\theta')}{\bar{s}_{n+1}(\mu,\theta')} \leq \frac{7}{10}$. Thus we obtain

$$\left|\left(\frac{\mu\dot{\bar{s}}_{n+1}(\mu,\theta')}{\bar{s}_{n+1}(\mu,\theta')}\right)^2 - \left(\frac{\mu\dot{\bar{s}}_{n+1}(\mu,\theta')}{\bar{s}_{n+1}(\mu,\theta')}\right)\right|^{\frac{1}{2}} \geq \frac{\sqrt{19}}{20} \tag{23}$$

for $\mu \in [\mu_2, 2\mu_2]$. Hence from (23) and Proposition 2.2, part *1.*, we obtain

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

Finally, we have

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

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

Finally we deal with the case when $m < n < 2m$. In this case let $\hat{A} = [A \ \ A]$, $\hat{b} = 2b$, and $\hat{c} = \begin{bmatrix} c \\ c \end{bmatrix}$ so that $\hat{n} = 2n > 2m$. Then the central path is given as $\hat{x}(\mu) = \begin{bmatrix} x(\mu) \\ x(\mu) \end{bmatrix}$, $\hat{y}(\mu) = y(\mu)$, and $\hat{s}(\mu) = \begin{bmatrix} s(\mu) \\ s(\mu) \end{bmatrix}$. From these formulas, it is easy to see 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 2.4.

Corollary 3.8. *Let $A \in \mathbb{R}^{m\times n}$. Then there exists an $\overline{m}$, a matrix $\bar{A} \in \mathbb{R}^{\overline{m}\times 2\overline{m}}$, and a constant τ independent of problem data such that,*

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

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

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

Proof. We give the proof for only the case of $n > 2m$. The case of $m < n < 2m$ is analogous. Let $\epsilon > 0$ and $A \in \mathbb{R}^{m\times n}$ be given. From Definition 2.4, 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 3.7, increasing the size of the problem $n - 2m$ times, we obtain a new problem data $\bar{A} \in \mathbb{R}^{\overline{m}\times 2\overline{m}}$, $b \in \mathbb{R}^{\overline{m}}$ and $c \in \mathbb{R}^{2\overline{m}}$ such that

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

where $\int_0^\infty \frac{\bar{\kappa}(\mu)}{\mu} d\mu$ is the Sonnevend curvature of the new central path and τ is the constant derived in the proof of Corollary 3.7. Using Definition 2.4 once again, it follows that

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

Since ϵ is arbitrarily small, the result follows.

□

In the end, several observations are in order. First, even though we presented the construction (6) for $n > 2m$, the construction (6) is valid for any m and n, and the increase in the Sonnevend curvature is still at least a constant. Second, repeating (6) leads to an $\Omega(n)$ worst-case lower bound for the Sonnevend curvature for a problem data $\bar{A}, \bar{b}, \bar{c}$, 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, by doing the scaling $\hat{b} := \frac{\bar{b}}{\eta}$ by a large η, the same $\Omega(n)$ worst-case iteration complexity can occur on any interval $[\mu^{'}, \mu^{''}]$.

4 Conclusions and further remarks

In order to prove an upper bound for the Sonnevend curvature for general any (m,n), it is sufficient, in light of our main result, to prove it only for the $n = 2m$ case. A similar behavior is known for both the diameter of a polytope [4], and the total geometric curvature of the central path [1, 2, 3]. The lower bound $\Omega(n)$ for the Sonnevend curvature is also analogous to the worst-case lower bound known for the geometric curvature, [1, 2, 3]. An interesting topic for future research would be to investigate the relationship between the Sonnevend curvature and the geometric curvature of the central path in a more general setting.

References

[1] Antoine Deza, Tamás Terlaky, Feng Xie, and Yuriy Zinchenko. Diameter and curvature: Intriguing analogies. *Electronic Notes in Discrete Mathematics*, 31(0):221 – 225, 2008.

[2] Antoine Deza, Tamás Terlaky, and Yuriy Zinchenko. Polytopes and arrangements: Diameter and curvature. *Operations Research Letters*, 36(2):215 – 222, 2008.

[3] Antoine Deza, Tamás Terlaky, and Yuriy Zinchenko. A continuous d-step conjecture for polytopes. *Discrete & Computational Geometry*, 41:318–327, 2009.

[4] Victor Klee and David W. Walkup. The d-step conjecture for polyhedra of dimension $d < 6$. *Acta Mathematica*, 117:53–78, 1967.

[5] Renato D.C. Monteiro and Takashi Tsuchiya. 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, 2008.

[6] Cornelius Roos, Tamás Terlaky, and Jean-Philippe Vial. *Interior Point Methods for Linear Optimization*. Springer, 2005.

[7] György Sonnevend, Joseph Stoer, and Gongyun Zhao. On the complexity of following the central path of linear programs by linear extrapolation II. *Mathematical Programming*, 52:527–553, 1991.

[8] Joseph Stoer and Gongyun Zhao. Estimating the complexity of a class of path-following methods for solving linear programs by curvature integrals. *Applied Mathematics and Optimization*, 27:85–103, 1993.

[9] Stephen A. Vavasis and Yinyu Ye. A primal-dual interior point method whose running time depends only on the constraint matrix. *Mathematical Programming*, 74:79–120, 1996.

[10] Gongyun Zhao. Representing the space of linear programs as the Grassmann manifold. *Mathematical Programming*, 121:353–386, 2010.