McMaster University

Advanced Optimization Laboratory

Title:

How good are interior point methods?

Klee-Minty cubes tighten iteration-complexity bounds

Authors:

Antoine Deza, Eissa Nematollahi and Tamás Terlaky

AdvOl-Report No. 2004/20

December 2004, Hamilton, Ontario, Canada

How good are interior point methods? Klee-Minty cubes tighten iteration-complexity bounds.

Dedicated to Professor Emil Klafszky on the occasion of his 70th birthday

Antoine Deza Eissa Nematollahi Tamás Terlaky

December 3, 2004

Abstract

By refining a variant of the Klee-Minty example that forces the central path to visit all the vertices of the Klee-Minty n-cube, we exhibit a nearly worst-case example for path-following interior point methods. Namely, while the theoretical iteration-complexity upper bound is $O(2^n n^{\frac{5}{2}})$, we prove that solving this n-dimensional linear optimization problem requires at least $2^n - 1$ iterations.

Key words: Linear programming, interior point method, worst-case iteration-complexity.
MSC2000 Subject Classification: Primary: 90C05; Secondary: 90C51, 90C27, 52B12

1 Introduction

While the *simplex method*, introduced by Dantzig [1], works very well in practice for linear optimization problems, in 1972 Klee and Minty [5] gave an example for which the simplex method takes an exponential number of iterations. More precisely, they considered a maximization problem over an n-dimensional "squashed" cube and proved that a variant of the simplex method visits all its 2^n vertices. Thus, the time complexity is not polynomial for the worst case, as $2^n - 1$ iterations are necessary for this n-dimensional linear optimization problem. The pivot rule used in the Klee-Minty example was the most negative reduced cost but variants of the Klee-Minty n-cube allow to prove exponential running time for most pivot rules; see [10] and the references therein. The Klee-Minty worst-case example partially stimulated the search for a polynomial algorithm and, in 1979, Khachiyan's [4] *ellipsoid method* proved that linear programming is indeed polynomially solvable. In 1984, Karmarkar [3] proposed a more efficient polynomial algorithm that sparked the research on polynomial *interior point methods*. In short, while the simplex method goes along the edges of the polyhedron corresponding to the feasible region, interior point methods pass through the interior of this polyhedron. Starting at the *analytic center*, most interior point methods follow the so-called *central path* and converge to the analytic center of the optimal face; see e.g., [6, 8, 9, 13, 14]. In 2004, Deza, Nematollahi, Peyghami and Terlaky [2] showed that, by carefully adding an exponential number of redundant constraints to the Klee-Minty n-cube, the central path can be severely distorted. Specifically, they provided an example for which path-following interior point methods have to take $2^n - 2$ sharp turns as the central path passes within an arbitrarily small neighborhood of the corresponding vertices of the Klee-Minty cube before converging to the optimal solution. This example yields a theoretical lower bound for the number of iterations needed for path-following interior point methods: The number of iterations is at least the number of sharp turns; that is, the iteration-complexity lower bound is $\Omega(2^n)$. On the other hand, the theoretical iteration-complexity upper bound is $O(\sqrt{N}L)$ where N and L respectively denote the number of constraints and the bit-length of the input-data. The iteration-complexity upper bound for the highly redundant Klee-Minty n-cube of [2] is $O(2^{3n}nL) = O(2^{9n}n^4)$, as $N = O(2^{6n}n^2)$ and $L = O(2^{6n}n^3)$ for this example. Therefore, these $2^n - 1$ sharp turns yield an $\Omega(\sqrt[6]{\frac{N}{\ln^2 N}})$ iteration-complexity lower bound. In this paper we show that a refined problem with the same $\Omega(2^n)$ iteration-complexity lower bound exhibits a nearly worst-case iteration-complexity as the complexity upper bound is $O(2^n n^{\frac{5}{2}})$. In other words, this new example, with $N = O(2^{2n}n^3)$, essentially closes the iteration-complexity gap with an $\Omega(\sqrt{\frac{N}{\ln^3 N}})$ lower bound and an $O(\sqrt{N}\ln N)$ upper bound.

2 Notations and the Main Results

We consider the following Klee-Minty variant where ε is a small positive factor by which the unit cube $[0,1]^n$ is squashed.

$$\begin{array}{rll} \min & x_n & \\ \text{subject to} & 0 \leq x_1 \leq 1 & \\ & \varepsilon\, x_{k-1} \leq x_k \leq 1 - \varepsilon\, x_{k-1} & \text{for } k = 2, \ldots, n. \end{array}$$

The above minimization problem has $2n$ constraints, n variables and the feasible region is an n-dimensional cube denoted by C. Some variants of the simplex method take $2^n - 1$ iterations to solve this problem as they visit all the vertices ordered by the decreasing value of the last coordinate x_n starting from $v^{\{n\}} = (0, \ldots, 0, 1)$ till the optimal value $x_n^* = 0$ is reached at the origin $v^{\emptyset}$.

While adding a set h of redundant inequalities does not change the feasible region, the analytic center χ^h and the central path are affected by the addition of redundant constraints. We consider redundant inequalities induced by hyperplanes parallel to the n facets of C containing the origin. The constraint parallel to the facet $H_1 : x_1 = 0$ is added h_1 times at a distance d_1 and the constraint parallel to the facet $H_k : x_k = \varepsilon x_{k-1}$ is added h_k times at a distance d_k for $k = 2, \ldots, n$. The set h is denoted by the integer-vector $h = (h_1, \ldots, h_n)$, $d = (d_1, \ldots, d_n)$, and the redundant linear optimization problem is defined by

$$\begin{array}{rll} \min & x_n & \\ \text{subject to} & 0 \leq x_1 \leq 1 & \\ & \varepsilon\, x_{k-1} \leq x_k \leq 1 - \varepsilon\, x_{k-1} & \text{for } k = 2, \ldots, n \\ & 0 \leq d_1 + x_1 & \text{repeated } h_1 \text{ times} \\ & \varepsilon\, x_1 \leq d_2 + x_2 & \text{repeated } h_2 \text{ times} \\ & \vdots & \vdots \\ & \varepsilon\, x_{n-1} \leq d_n + x_n & \text{repeated } h_n \text{ times.} \end{array}$$

By analogy with the unit cube $[0,1]^n$, we denote the vertices of the Klee-Minty cube C by using a subset S of $\{1,\dots,n\}$. For $S \subset \{1,\dots,n\}$, a vertex v^S of C is defined by

$$v_1^S = \begin{cases} 1, & \text{if } 1 \in S \\ 0, & \text{otherwise} \end{cases}$$
$$v_k^S = \begin{cases} 1-\varepsilon v_{k-1}^S, & \text{if } k \in S \\ \varepsilon v_{k-1}^S, & \text{otherwise} \end{cases} \qquad k=2,\dots,n.$$

The δ-neighborhood $\mathcal{N}_\delta(v^S)$ of a vertex v^S is defined, with the convention $x_0 = 0$, by

$$\mathcal{N}_\delta(v^S) = \left\{ x \in C : \begin{cases} 1 - x_k - \varepsilon x_{k-1} \le \varepsilon^{k-1}\delta, & \text{if } k \in S \\ x_k - \varepsilon x_{k-1} \le \varepsilon^{k-1}\delta, & \text{otherwise} \end{cases} \qquad k=1,\dots,n \right\}.$$

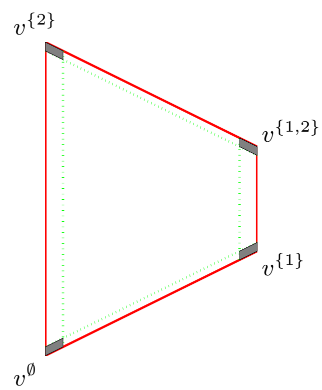

Figure 1: The δ-neighborhoods of the 4 vertices of the Klee-Minty 2-cube.

In this paper we focus on the following problem $\mathcal{C}_\delta^n$ defined by

$\varepsilon = \frac{n}{2(n+1)}$,
$d = n(2^{n+4},\dots,2^{n-k+5},\dots,2^5)$,
$h = \left(\lfloor \frac{2^{2n+8}(n+1)^n}{\delta n^{n-1}} - \frac{2^{n+7}(n+1)}{\delta} \rfloor, \dots, \lfloor \frac{2^{2n+8}(n+1)^{n+k-1}}{\delta n^{n+k-2}} - \frac{2^{n+k+6}(n+1)^{2k-1}}{\delta n^{2k-2}} \rfloor, \dots, \lfloor 3\frac{2^{2n+6}(n+1)^{2n-1}}{\delta n^{2n-2}} \rfloor \right)$,

where $0 < \delta \le \frac{1}{4(n+1)}$.

Note that we have: $\varepsilon + \delta < \frac{1}{2}$; that is, the δ-neighborhoods of the 2^n vertices are non-overlapping, and that h is, up to a floor operation, linearly dependent on δ^{-1}. Proposition 2.1 states that, for $\mathcal{C}_\delta^n$, the central path takes at least $2^n - 2$ turns before converging to the origin as it passes through the δ-neighborhood of all the 2^n vertices of the Klee-Minty n-cube; See Section 3.2 for the proof. Note that the proof given in Section 3.2 yields a slightly stronger result than Proposition 2.1: In addition to pass through the δ-neighborhood of all the vertices, the central path is bent along the edge-path followed by the simplex method. We set $\delta = \frac{1}{4(n+1)}$ in Propositions 2.3 and 2.4 in order to exhibit the sharpest bounds. The corresponding linear optimization problem $\mathcal{C}_{1/4(n+1)}^n$ depends only on the dimension n.

Proposition 2.1. *The central path $\mathcal{P}$ of $\mathcal{C}^n_\delta$ intersects the δ-neighborhood of each vertex of the n-cube.*

Since the number of iterations required by path-following interior point methods is at least the number of sharp turns, Proposition 2.1 yields a theoretical lower bound for the iteration-complexity for solving this n-dimensional linear optimization problem.

Corollary 2.2. *For $\mathcal{C}^n_\delta$, the iteration-complexity lower bound of path-following interior point methods is $\Omega(2^n)$.*

Since the theoretical iteration-complexity upper bound for path-following interior point methods is $O(\sqrt{N}L)$, where N and L respectively denote the number of constraints and the bit-length of the input-data, we have:

Proposition 2.3. *For $\mathcal{C}^n_{1/4(n+1)}$, the iteration-complexity upper bound of path-following interior point methods is $O(2^n n^{\frac{3}{2}} L)$; that is, $O(2^{3n} n^{\frac{11}{2}})$.*

Proof. We have $N = 2n + \sum_{k=1}^n h_k = 2n + \sum_{k=1}^n n^2 \left(2^{2n+10}\left(\frac{n+1}{n}\right)^{n+k} - 2^{n+k+8}\left(\frac{n+1}{n}\right)^{2k}\right)$ and, since $\sum_{k=1}^n \left(\frac{n+1}{n}\right)^{n+k} \leq ne^2$, we have $N = O(2^{2n}n^3)$ and $L \leq N \ln d_1 = O(2^{2n}n^4)$. □

Noticing that the only two vertices with last coordinates smaller than or equal to ε^{n-1} are $v^{\emptyset}$ and $v^{\{1\}}$, with $v^{\emptyset}_n = 0$ and $v^{\{1\}}_n = \varepsilon^{n-1}$, the stopping criterion can be replaced by: stopping duality gap smaller than ε^n with the corresponding central path parameter at the stopping point being $\mu^* = \frac{\varepsilon^n}{N}$. Additionally, one can check that by setting the central path parameter to $\mu^0 = 1$, we obtain a starting point which belongs to the interior of the δ-neighborhood of the highest vertex $v^{\{n\}}$, see Section 3.3 for a detailed proof. In other words, a path-following algorithm using a standard ϵ-precision as stopping criterion can stop when the duality gap is smaller than ε^n as the optimal vertex is identified, see [8]. The corresponding iteration-complexity bound $O(\sqrt{N} \log \frac{N}{\epsilon})$ yields, for our construction, a precision-independent iteration-complexity $O(\sqrt{N} \ln \frac{N\mu^0}{N\mu^*}) = O(\sqrt{N}n)$ and Proposition 2.3 can therefore be strengthened to:

Proposition 2.4. *For $\mathcal{C}^n_{1/4(n+1)}$, the iteration-complexity upper bound of path-following interior point methods is $O(2^n n^{\frac{5}{2}})$.*

Remark 2.5.

(*i*) For $\mathcal{C}^n_{1/4(n+1)}$, by Corollary 2.2 and Proposition 2.4, the order of the iteration-complexity of path-following interior point methods is between 2^n and $2^n n^{\frac{5}{2}}$ or, equivalently, between $\sqrt{\frac{N}{\ln^3 N}}$ and $\sqrt{N} \ln N$.

(*ii*) The k-th coordinate of the vector d corresponds to the scalar d defined in [2] for dimension $n - k + 3$.

(*iii*) Other settings for d and h ensuring that the central path visits all the vertices of the Klee-Minty n-cube are possible. For example, d can be set to $(1.1, 22)$ in dimension 2.

(*iv*) Our results apply to path-following interior point methods but not to other interior point methods such as Karmarkar's original projective algorithm [3].

Remark 2.6.

(*i*) Megiddo and Schub [7] proved, for affine scaling trajectories, a result with a similar flavor as our result for the central path, and noted that their approach does not extend to projective scaling. They considered the non-redundant Klee-Minty cube.

(*ii*) Todd and Ye [11] gave an $\Omega(\sqrt[3]{N})$ iteration-complexity lower bound between two updates of the central path parameter μ.

(*iii*) Vavasis and Ye [12] provided an $O(N^2)$ upper bound for the number of approximately straight segments of the central path.

3 Proofs of Proposition 2.1 and Proposition 2.4

3.1 Preliminary Lemmas

Lemma 3.1. *With* $b = \frac{4}{\delta}(1,\dots,1)$, $\varepsilon = \frac{n}{2(n+1)}$, $d = n(2^{n+4},\dots,2^{n-k+5},\dots,2^5)$, $\tilde{h} = \left(\frac{2^{2n+8}(n+1)^n}{\delta n^{n-1}} - \frac{2^{n+7}(n+1)}{\delta},\dots,\frac{2^{2n+8}(n+1)^{n+k-1}}{\delta n^{n+k-2}} - \frac{2^{n+k+6}(n+1)^{2k-1}}{\delta n^{2k-2}},\dots,3\frac{2^{2n+6}(n+1)^{2n-1}}{\delta n^{2n-2}}\right)$ *and*

$$A = \begin{pmatrix} \frac{1}{d_1+1} & \frac{-\varepsilon}{d_2} & 0 & 0 & \dots & 0 & 0 \\ \frac{-1}{d_1} & \frac{2\varepsilon}{d_2+1} & \frac{-\varepsilon^2}{d_3} & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \frac{-1}{d_1} & 0 & 0 & \frac{2\varepsilon^{k-1}}{d_k+1} & \frac{-\varepsilon^k}{d_{k+1}} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ \frac{-1}{d_1} & 0 & 0 & 0 & \dots & \frac{2\varepsilon^{n-2}}{d_{n-1}+1} & \frac{-\varepsilon^{n-1}}{d_n} \\ \frac{-1}{d_1} & 0 & 0 & 0 & \dots & 0 & \frac{2\varepsilon^{n-1}}{d_n+1} \end{pmatrix},$$

we have $A\tilde{h} \geq \frac{3b}{2}$.

Proof. As $\varepsilon = \frac{n}{2(n+1)}$ and $d = n(2^{n+4},\dots,2^{n-k+5},\dots,2^5)$, $\tilde{h}$ can be rewritten as $\tilde{h} = \frac{4}{\delta}\left(d_1(\frac{4}{\varepsilon^n} - \frac{1}{\varepsilon}),\dots,\frac{d_k}{\varepsilon^{k-1}}(\frac{4}{\varepsilon^n} - \frac{1}{\varepsilon^k}),\dots,\frac{d_n}{\varepsilon^{n-1}}\frac{3}{\varepsilon^n}\right)$ and $A\tilde{h} \geq \frac{3b}{2}$ can be rewritten as

$$\begin{aligned} \frac{4}{\delta}\frac{d_1}{d_1+1}(\frac{4}{\varepsilon^n} - \frac{1}{\varepsilon}) - \frac{4}{\delta}(\frac{4}{\varepsilon^n} - \frac{1}{\varepsilon^2}) &\geq \frac{6}{\delta} \\ -\frac{4}{\delta}(\frac{4}{\varepsilon^n} - \frac{1}{\varepsilon}) + \frac{4}{\delta}\frac{2d_k}{d_k+1}(\frac{4}{\varepsilon^n} - \frac{1}{\varepsilon^k}) - \frac{4}{\delta}(\frac{4}{\varepsilon^n} - \frac{1}{\varepsilon^{k+1}}) &\geq \frac{6}{\delta} \quad \text{for } k = 2,\dots,n-1 \\ -\frac{4}{\delta}(\frac{4}{\varepsilon^n} - \frac{1}{\varepsilon}) + \frac{4}{\delta}\frac{2d_n}{d_n+1}\frac{3}{\varepsilon^n} &\geq \frac{6}{\delta}, \end{aligned}$$

which is equivalent to

$$\begin{aligned} \left(\frac{1}{\varepsilon^2} - \frac{1}{\varepsilon} - \frac{3}{2}\right) d_1 &\geq \frac{4}{\varepsilon^n} - \frac{1}{\varepsilon^2} + \frac{3}{2} \\ \left(\frac{1}{\varepsilon^{k+1}} - \frac{2}{\varepsilon^k} + \frac{1}{\varepsilon} - \frac{3}{2}\right) d_k &\geq \frac{8}{\varepsilon^n} - \frac{1}{\varepsilon^{k+1}} - \frac{1}{\varepsilon} + \frac{3}{2} \quad \text{for } k = 2,\dots,n-1 \\ \left(\frac{2}{\varepsilon^n} + \frac{1}{\varepsilon} - \frac{3}{2}\right) d_n &\geq \frac{4}{\varepsilon^n} - \frac{1}{\varepsilon} + \frac{3}{2}. \end{aligned}$$

As $\frac{1}{\varepsilon^2} - \frac{1}{\varepsilon} - \frac{3}{2} \geq \frac{1}{2}$, $\frac{1}{\varepsilon} - \frac{3}{2} \geq 0$, $\frac{1}{\varepsilon^2} - \frac{3}{2} \geq 0$ and $\frac{1}{\varepsilon^{k+1}} + \frac{1}{\varepsilon} - \frac{3}{2} \geq 0$, the above system is implied by

$$\begin{array}{rcll} \frac{1}{2}d_1 & \geq & \frac{4}{\varepsilon^n} & \\ \left(\frac{1}{\varepsilon^{k+1}} - \frac{2}{\varepsilon^k}\right) d_k & \geq & \frac{8}{\varepsilon^n} & \text{for } k = 2, \ldots, n-1 \\ \frac{2}{\varepsilon^n} d_n & \geq & \frac{4}{\varepsilon^n}, & \end{array}$$

as $\frac{1}{\varepsilon^{k+1}} - \frac{2}{\varepsilon^k} = \frac{2}{n\varepsilon^k}$ and $\frac{1}{\varepsilon^{n-k}} = 2^{n-k}\left(1 + \frac{1}{n}\right)^{n-k} \leq 2^{n-k+2}$, the above system is implied by

$$\begin{array}{lll} d_1 & \geq 2^{n+5} & \\ d_k & \geq n2^{n-k+4} & \text{for } k = 2, \ldots, n-1 \\ d_n & \geq 2 & \end{array}$$

which is true since $d = n(2^{n+4}, \ldots, 2^{n-k+5}, \ldots, 2^5)$. □

Corollary 3.2. *With the same assumptions as in Lemma 3.1 and $h = \lfloor \tilde{h} \rfloor$, we have $Ah \geq b$.*

Proof. Since $0 \leq \tilde{h}_k - h_k < 1$ and $d_k = n2^{n-k+5}$, we have:

$$\begin{array}{rll} \frac{\tilde{h}_1 - h_1}{d_1+1} - \frac{(\tilde{h}_2 - h_2)\varepsilon}{d_2} & \leq \frac{2}{\delta} & \\ -\frac{\tilde{h}_1 - h_1}{d_1} + \frac{2(\tilde{h}_k - h_k)\varepsilon^{k-1}}{d_k+1} - \frac{(\tilde{h}_{k+1} - h_{k+1})\varepsilon^k}{d_{k+1}} & \leq \frac{2}{\delta} & \text{for } k = 2, \ldots, n-1 \\ -\frac{\tilde{h}_1 - h_1}{d_1} + \frac{2(\tilde{h}_n - h_n)\varepsilon^{n-1}}{d_n+1} & \leq \frac{2}{\delta} & \end{array}$$

thus, $A(\tilde{h} - h) \leq \frac{b}{2}$, which implies, since $A\tilde{h} \geq \frac{3b}{2}$ by Lemma 3.1, that $Ah \geq b$. □

Corollary 3.3. *With the same assumptions as in Lemma 3.1 and $h = \lfloor \tilde{h} \rfloor$, we have: $\frac{h_k \varepsilon^{k-1}}{d_k+1} \geq \frac{h_{k+1}\varepsilon^k}{d_{k+1}} + \frac{4}{\delta}$ for $k = 1, \ldots, n-1$.*

Proof. For $k = 1, \ldots, n-1$, one can easily check that the first k inequalities of $Ah \geq b$ imply $\frac{h_k \varepsilon^{k-1}}{d_k+1} \geq \frac{h_{k+1}\varepsilon^k}{d_{k+1}} + \frac{4}{\delta}$. □

The analytic center $\chi^n = (\xi_1^n, \ldots, \xi_n^n)$ of $\mathcal{C}_\delta^n$ is the unique solution to the problem consisting of maximizing the product of the slack variables:

$$\begin{array}{ll} s_1 = x_1 & \\ s_k = x_k - \varepsilon x_{k-1} & \text{for } k = 2, \ldots, n \\ \bar{s}_1 = 1 - x_1 & \\ \bar{s}_k = 1 - \varepsilon x_{k-1} - x_k & \text{for } k = 2, \ldots, n \\ \tilde{s}_1 = d_1 + s_1 & \text{repeated } h_1 \text{ times} \\ \vdots & \vdots \\ \tilde{s}_n = d_n + s_n & \text{repeated } h_n \text{ times.} \end{array}$$

Equivalently, χ^n is the solution of the following maximization problem:

$$\max_x \sum_{k=1}^{n} (\ln s_k + \ln \bar{s}_k + h_k \ln \tilde{s}_k),$$

i.e., with the convention $x_0 = 0$,

$$\max_x \sum_{k=1}^{n} \Big(\ln(x_k - \varepsilon x_{k-1}) + \ln(1 - \varepsilon x_{k-1} - x_k) + h_k \ln(d_k + x_k - \varepsilon x_{k-1}) \Big).$$

The optimality conditions (the gradient is equal to zero at optimality) for this concave maximization problem give:

$$\left\{ \begin{array}{rl} \frac{1}{\sigma_k^n} - \frac{\varepsilon}{\sigma_{k+1}^n} - \frac{1}{\bar{\sigma}_k^n} - \frac{\varepsilon}{\bar{\sigma}_{k+1}^n} + \frac{h_k}{\tilde{\sigma}_k^n} - \frac{h_{k+1}\varepsilon}{\tilde{\sigma}_{k+1}^n} = 0 & \text{for } k = 1, \dots, n-1 \\ \frac{1}{\sigma_n^n} - \frac{1}{\bar{\sigma}_n^n} + \frac{h_n}{\tilde{\sigma}_n^n} = 0 & \\ \sigma_k^n > 0, \bar{\sigma}_k^n > 0, \tilde{\sigma}_k^n > 0 & \text{for } k = 1, \dots, n, \end{array} \right. \tag{1}$$

where

$$\begin{array}{ll} \sigma_1^n = \xi_1^n & \\ \sigma_k^n = \xi_k^n - \varepsilon \xi_{k-1}^n & \text{for } k = 2, \dots, n \\ \bar{\sigma}_1^n = 1 - \xi_1^n & \\ \bar{\sigma}_k^n = 1 - \varepsilon \xi_{k-1}^n - \xi_k^n & \text{for } k = 2, \dots, n \\ \tilde{\sigma}_k^n = d_k + \sigma_k^n & \text{for } k = 1, \dots, n. \end{array}$$

The following lemma states that, for $\mathcal{C}_\delta^n$, the analytic center χ^n belongs to the neighborhood of the vertex $v^{\{n\}} = (0, \dots, 0, 1)$.

Lemma 3.4. *For $\mathcal{C}_\delta^n$, we have: $\chi^n \in \mathcal{N}_\delta(v^{\{n\}})$.*

Proof. Adding the n-th equation of (1) multiplied by $-\varepsilon^{n-1}$ to the j-th equation of (1) multiplied by ε^{j-1} for $j = k \dots, n-1$, we have, for $k = 1, \dots, n-1$,

$$\frac{\varepsilon^{k-1}}{\sigma_k^n} - \frac{\varepsilon^{k-1}}{\bar{\sigma}_k^n} - \frac{2\varepsilon^{n-1}}{\sigma_n^n} - 2\sum_{i=k}^{n-2} \frac{\varepsilon^i}{\bar{\sigma}_{i+1}^n} + \frac{h_k \varepsilon^{k-1}}{\tilde{\sigma}_k^n} - \frac{2h_n \varepsilon^{n-1}}{\tilde{\sigma}_n^n} = 0,$$

implying:

$$\frac{2h_n \varepsilon^{n-1}}{\tilde{\sigma}_n^n} - \frac{h_k \varepsilon^{k-1}}{\tilde{\sigma}_k^n} = \frac{\varepsilon^{k-1}}{\sigma_k^n} - \left(\frac{\varepsilon^{k-1}}{\bar{\sigma}_k^n} + \frac{2\varepsilon^{n-1}}{\sigma_n^n} + 2\sum_{i=k}^{n-2} \frac{\varepsilon^i}{\bar{\sigma}_{i+1}^n} \right) \le \frac{\varepsilon^{k-1}}{\sigma_k^n},$$

which implies, since $\tilde{\sigma}_n^n \le d_n + 1$, $\tilde{\sigma}_k^n \ge d_k$ and $\frac{h_1}{d_1} \ge \frac{h_k \varepsilon^{k-1}}{d_k}$ by Corollary 3.3,

$$\frac{2h_n \varepsilon^{n-1}}{d_n + 1} - \frac{h_1}{d_1} \le \frac{\varepsilon^{k-1}}{\sigma_k^n},$$

implying, since $\frac{2h_n\varepsilon^{n-1}}{d_n+1} - \frac{h_1}{d_1} \geq \frac{1}{\delta}$ by Corollary 3.2, $\sigma_k^n \leq \varepsilon^{k-1}\delta$ for $k = 1, \ldots, n-1$. The n-th equation of (1) implies: $\frac{h_n\varepsilon^{n-1}}{\tilde{\sigma}_n^n} \leq \frac{\varepsilon^{n-1}}{\bar{\sigma}_n^n}$; that is, since $\tilde{\sigma}_n^n < d_n + 1$ and $\frac{h_n\varepsilon^{n-1}}{d_n+1} \geq \frac{1}{\delta}$ by Corollary 3.2, we have: $\frac{1}{\delta} \leq \frac{h_n\varepsilon^{n-1}}{d_n+1} \leq \frac{\varepsilon^{n-1}}{\bar{\sigma}_n^n}$, implying: $\bar{\sigma}_n^n \leq \varepsilon^{n-1}\delta$. □

The central path $\mathcal{P}$ of $\mathcal{C}_\delta^n$ can be defined as the set of analytic centers $\chi^n(\alpha) = (x_1^n, \ldots, x_{n-1}^n, \alpha)$ of the intersection of the hyperplane $H_\alpha : x_n = \alpha$ with the feasible region of $\mathcal{C}_\delta^n$ where $0 < \alpha \leq \xi_n^n$, see [8]. These intersections $\Omega(\alpha)$ are called the *level sets* and $\chi^n(\alpha)$ is the solution of the following system:

$$\left\{ \begin{array}{rl} \frac{1}{s_k^n} - \frac{\varepsilon}{s_{k+1}^n} - \frac{1}{\bar{s}_k^n} - \frac{\varepsilon}{\bar{s}_{k+1}^n} + \frac{h_k}{\tilde{s}_k^n} - \frac{h_{k+1}\varepsilon}{\tilde{s}_{k+1}^n} = 0 & \text{for } k = 1, \ldots, n-1 \\ s_k^n > 0, \bar{s}_k^n > 0, \tilde{s}_k^n > 0 & \text{for } k = 1, \ldots, n-1, \end{array} \right. \tag{2}$$

where

$$\begin{array}{rcll} s_1^n &=& x_1^n & \\ s_k^n &=& x_k^n - \varepsilon x_{k-1}^n & \text{for } k = 2, \ldots, n-1 \\ s_n^n &=& \alpha - \varepsilon x_{n-1} & \\ \bar{s}_1^n &=& 1 - x_1^n & \\ \bar{s}_k^n &=& 1 - \varepsilon x_{k-1}^n - x_k^n & \text{for } k = 2, \ldots, n-1 \\ \bar{s}_n^n &=& 1 - \alpha - \varepsilon x_{n-1}^n & \\ \tilde{s}_k^n &=& d_k + s_k^n & \text{for } k = 1, \ldots, n. \end{array}$$

Lemma 3.5. *For $\mathcal{C}_\delta^n$, $C_\delta^k = \{x \in C : \ \bar{s}_k \geq \varepsilon^{k-1}\delta, s_k \geq \varepsilon^{k-1}\delta\}$ and $\hat{C}_\delta^k = \{x \in C : \ \bar{s}_{k-1} \leq \varepsilon^{k-2}\delta, s_{k-2} \leq \varepsilon^{k-3}\delta, \ldots, s_1 \leq \delta\}$, we have: $C_\delta^k \cap \mathcal{P} \subseteq \hat{C}_\delta^k$ for $k = 2, \ldots, n$.*

Proof. Let $x \in C_\delta^k \cap \mathcal{P}$. Adding the $(k-1)$-st equation of (2) multiplied by $-\varepsilon^{k-2}$ to the i-th equation of (1) multiplied by ε^{i-1} for $i = j \ldots, k-2$, we have, for $k = 2, \ldots, n-1$,

$$-\frac{2h_{k-1}\varepsilon^{k-2}}{\tilde{s}_{k-1}^n} + \frac{h_j\varepsilon^{j-1}}{\tilde{s}_j^n} + \frac{h_k\varepsilon^{k-1}}{\tilde{s}_k^n} + \frac{\varepsilon^{j-1}}{s_j^n} + \frac{\varepsilon^{k-1}}{s_k^n} + \frac{\varepsilon^{k-1}}{\bar{s}_k^n} - \left(\frac{2\varepsilon^{k-2}}{s_{k-1}^n} + \frac{\varepsilon^{j-1}}{\bar{s}_j^n} + 2\sum_{i=j}^{k-3} \frac{\varepsilon^i}{\bar{s}_{i+1}^n} \right) = 0,$$

which implies, since $\tilde{s}_{k-1}^n < d_{k-1} + 1$, $\tilde{s}_j^n > d_j$, $\tilde{s}_k^n > d_k$ and $s_k^n \geq \varepsilon^{k-1}\delta$ and $\bar{s}_k^n \geq \varepsilon^{k-1}\delta$ as $x \in C_\delta^k$,

$$\frac{2h_{k-1}\varepsilon^{k-2}}{d_{k-1}+1} - \frac{h_j\varepsilon^{j-1}}{d_j} - \frac{h_k\varepsilon^{k-1}}{d_k} \leq \frac{\varepsilon^{j-1}}{s_j^n} + \frac{2}{\delta},$$

implying, since $\frac{h_1}{d_1} \geq \frac{h_j\varepsilon^{j-1}}{d_j}$ by Corollary 3.3,

$$-\frac{h_1}{d_1} + \frac{2h_{k-1}\varepsilon^{k-2}}{d_{k-1}+1} - \frac{h_k\varepsilon^{k-1}}{d_k} \leq \frac{\varepsilon^{j-1}}{s_j^n} + \frac{2}{\delta},$$

that is, as $\frac{3}{\delta} \leq -\frac{h_1}{d_1} + \frac{2h_{k-1}\varepsilon^{k-2}}{d_{k-1}+1} - \frac{h_k\varepsilon^{k-1}}{d_k}$ by Corollary 3.2: $s_j^n \leq \varepsilon^{j-1}\delta$. Considering the $(k-1)$-st equation of (2), we have

$$\frac{h_{k-1}\varepsilon^{k-2}}{\tilde{s}_{k-1}^n} - \frac{h_k\varepsilon^{k-1}}{\tilde{s}_k^n} = \frac{\varepsilon^{k-2}}{\bar{s}_{k-1}^n} + \frac{\varepsilon^{k-1}}{s_k^n} + \frac{\varepsilon^{k-1}}{\bar{s}_k^n} - \frac{\varepsilon^{k-2}}{s_{k-1}^n},$$

which implies, since $\tilde{s}^n_{k-1} < d_{k-1} + 1$, $\tilde{s}^n_k > d_k$ and $s^n_k \geq \varepsilon^{k-1}\delta$ and $\bar{s}^n_k \geq \varepsilon^{k-1}\delta$ as $x \in C^k_\delta$,

$$\frac{h_{k-1}\varepsilon^{k-2}}{d_{k-1}+1} - \frac{h_k\varepsilon^{k-1}}{d_k} \leq \frac{\varepsilon^{k-2}}{\bar{s}^n_{k-1}} + \frac{2}{\delta},$$

which implies, since $\frac{3}{\delta} \leq \frac{h_{k-1}\varepsilon^{k-2}}{d_{k-1}+1} - \frac{h_k\varepsilon^{k-1}}{d_k}$ by Corollary 3.3, that $\bar{s}^n_{k-1} \leq \varepsilon^{k-2}\delta$ and, therefore, $x \in \hat{C}^k_\delta$. ∎

3.2 Proof of Proposition 2.1

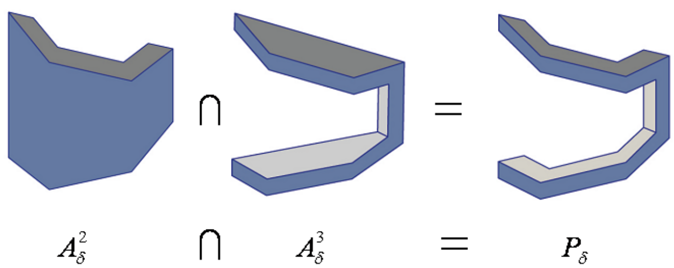

Figure 2: The set P_δ for the Klee-Minty 3-cube.

For $k = 2, \ldots, n$, while C^k_δ, defined in Lemma 3.5, can be seen as the central part of the cube C, the sets $T^k_\delta = \{x \in C : \ \bar{s}_k \leq \varepsilon^{k-1}\delta\}$ and $B^k_\delta = \{x \in C : \ s_k \leq \varepsilon^{k-1}\delta\}$, can be seen, respectively, as the top and bottom part of C. Clearly, we have $C = T^k_\delta \cup C^k_\delta \cup B^k_\delta$ for each $k = 2, \ldots, n$. Using the set $\hat{C}^k_\delta$ defined in Lemma 3.5, we consider the set $A^k_\delta = T^k_\delta \cup \hat{C}^k_\delta \cup B^k_\delta$ for $k = 2 \ldots, n$, and, for $0 < \delta \leq \frac{1}{4(n+1)}$, we show that the set $P_\delta = \bigcap_{k=2}^n A^k_\delta$, see Figure 2, contains the central path $\mathcal{P}$. By Lemma 3.4, the starting point χ^n of $\mathcal{P}$ belongs to $\mathcal{N}_\delta(v^{\{n\}})$. Since $\mathcal{P} \subset C$ and $C = \bigcap_{k=2}^n (T^k_\delta \cup C^k_\delta \cup B^k_\delta)$, we have:

$$\mathcal{P} = C \cap \mathcal{P} = \bigcap_{k=2}^{n} (T^k_\delta \cup C^k_\delta \cup B^k_\delta) \cap \mathcal{P} = \bigcap_{k=2}^{n} (T^k_\delta \cup (C^k_\delta \cap \mathcal{P}) \cup B^k_\delta) \cap \mathcal{P},$$

that is, by Lemma 3.5,

$$\mathcal{P} \subseteq \bigcap_{k=2}^{n} (T^k_\delta \cup \hat{C}^k_\delta \cup B^k_\delta) = \bigcap_{k=2}^{n} A^k_\delta = P_\delta$$

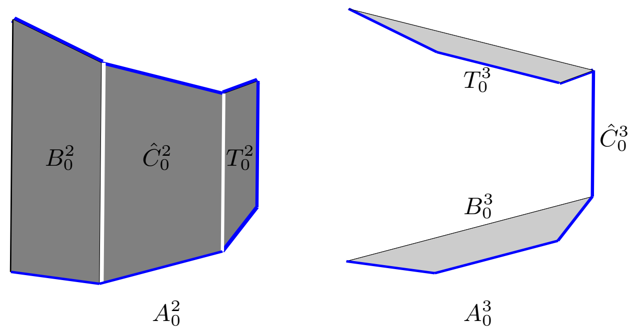

Figure 3: The sets A_0^2 and A_0^3 for the Klee-Minty 3-cube.

□

Remark that the sets $C_\delta^k, \hat{C}_\delta^k, T_\delta^k, B_\delta^k$ and A_δ^k can be defined for $\delta = 0$, see Figure 3, and that the corresponding set $P_0 = \bigcap_{k=2}^n A_0^k$ is precisely the path followed by the simplex method on the original Klee-Minty problem as it pivots along the edges of C. The set P_δ is a δ-sized (cross section) tube along the path P_0. See Figure 4 illustrating how P_0 starts at $v^{\{n\}}$, decreases with respect to the last coordinate x_n and ends at $v^\emptyset$.

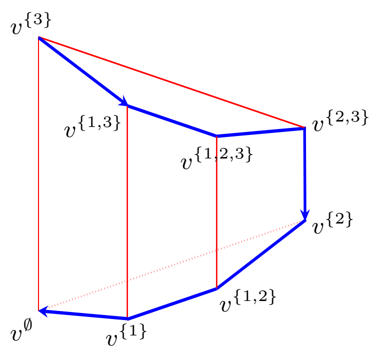

Figure 4: The path P_0 followed by the simplex method for the Klee-Minty 3-cube.

3.3 Proof of Proposition 2.4

We consider the point $\bar{x}$ of the central path which lies on the boundary of the δ-neighborhood of the highest vertex $v^{\{n\}}$. This point is defined by: $s_1 = \delta, s_k \leq \varepsilon^{k-1}\delta$ for $k = 2, \ldots, n-1$ and $s_{2n} \leq \varepsilon^n \delta$. Note that the notation s_k for the central path (perturbed complementarity) conditions, $y_k s_k = \mu$ for $k = 1, \ldots, p_n$, is consistent with the slacks introduced after Corollary 3.3 with $s_{n+k} = \bar{s}_k$ for $k = 1, \ldots, n$ and $s_{p_i+k} = \tilde{s}_k$ for $k = 1, \ldots, h_{i+1}$ and $i = 0, \ldots, n-1$. Let $\bar{\mu}$ denote the central path parameter corresponding to $\bar{x}$. In the following, we prove that $\bar{\mu} \leq \varepsilon^{n-1}\delta$ which implies that any point of the central path with corresponding parameter $\mu \geq \bar{\mu}$ belong to the interior of the δ-neighborhood of the highest vertex $v^{\{n\}}$. In particular, it implies that by setting the central path parameter to $\mu^0 = 1$, we obtain a starting point which belongs to the interior of the δ-neighborhood of the vertex $v^{\{n\}}$.

Estimation of the central path parameter $\bar{\mu}$:

The formulation of the dual problem of $\mathcal{C}^n_\delta$ is:

$$\max \quad z = -\sum_{k=n+1}^{2n} y_k - \sum_{k=1}^{n} d_k \sum_{i=p_{k-1}+1}^{p_k} y_i$$

$$\text{subject to} \quad y_k - \varepsilon y_{k+1} - y_{n+k} - \varepsilon y_{n+k+1} + \sum_{i=p_{k-1}+1}^{p_k} y_i - \varepsilon \sum_{i=p_k+1}^{p_{k+1}} y_i = 0 \quad \text{for } k = 1, \dots, n-1$$

$$y_n - y_{2n} + \sum_{i=p_{n-1}+1}^{p_n} y_i = 1$$

$$y_k \geq 0 \quad \text{for } k = 1, \dots, p_n,$$

where $p_0 = 2n$ and $p_k = 2n + h_1 + \cdots + h_k$ for $k = 1, \dots, n$.

For $k = 1, \dots, n$, multiplying by ε^{k-1} the k-th equation of the above dual constraints and summing then up, we have:

$$y_1 - y_{n+1} - 2\left(\varepsilon y_{n+2} + \varepsilon^2 y_{n+3} + \cdots + \varepsilon^{n-1} y_{2n}\right) + \sum_{i=2n+1}^{2n+h_1} y_i = \varepsilon^{n-1}$$

which implies

$$2\varepsilon^{n-1} y_{2n} \leq y_1 + \sum_{i=2n+1}^{2n+h_1} y_i$$

implying, since for $i = 2n+1, \dots, 2n+h_1$, $d_1 \leq s_i$ yields $y_i \leq \frac{\bar{\mu}}{d_1}$, that

$$2\varepsilon^{n-1} y_{2n} \leq y_1 + \frac{\bar{\mu} h_1}{d_1} = \frac{\bar{\mu}}{\delta} + \frac{\bar{\mu} h_1}{d_1}.$$

Since for $i = p_{n-1}+1, \dots, p_n$, $s_i = d_n + \bar{x}_n - \varepsilon \bar{x}_{n-1} \leq d_n + 1$ yields $y_i \geq \frac{\bar{\mu}}{d_n+1}$, the last dual constraint implies

$$y_{2n} \geq \sum_{i=p_{n-1}+1}^{p_n} y_i - 1 \geq \frac{\bar{\mu} h_n}{d_n + 1} - 1$$

which, combined with the previously obtained inequality, gives $\bar{\mu}\left(\frac{2h_n\varepsilon^{n-1}}{d_n+1} - \frac{h_1}{d_1} - \frac{1}{\delta}\right) \leq 2\varepsilon^{n-1}$, and, since Corollary 3.2 gives $\frac{2h_n\varepsilon^{n-1}}{d_n+1} - \frac{h_1}{d_1} - \frac{1}{\delta} \geq \frac{2}{\delta}$, we have $\bar{\mu} \leq \varepsilon^{n-1}\delta$.

Acknowledgments. We would like to thank an associate editor and the referees for pointing out the paper [7] and for helpful comments and corrections. Many thanks to Yinyu Ye for precious suggestions and hints which triggered this work and for informing us about the papers [11, 12]. Research supported by an NSERC Discovery grant and a MITACS grant for the last two authors, by the Canada Research Chair program for the first and last authors and by an NSERC Discovery grant for the first author.

References

[1] G. B. Dantzig: Maximization of a linear function of variables subject to linear inequalities. In: T. C. Koopmans (ed.) *Activity Analysis of Production and Allocation*. John Wiley (1951) 339–347.

[2] A. Deza, E. Nematollahi, R. Peyghami and T. Terlaky: The central path visits all the vertices of the Klee-Minty cube. *Optimization Methods and Software* (to appear).

[3] N. K. Karmarkar: A new polynomial-time algorithm for linear programming. *Combinatorica* 4 (1984) 373–395.

[4] L. G. Khachiyan: A polynomial algorithm in linear programming. *Soviet Mathematics Doklady* 20 (1979) 191–194.

[5] V. Klee and G. J. Minty: How good is the simplex algorithm? In: O. Shisha (ed.) *Inequalities III*, Academic Press (1972) 159–175.

[6] N. Megiddo: Pathways to the optimal set in linear programming. In: N. Megiddo (ed.) *Progress in Mathematical Programming: Interior-Point and Related Methods*, Springer-Verlag (1988) 131–158; also in: *Proceedings of the 7th Mathematical Programming Symposium of Japan*, Nagoya, Japan (1986) 1–35.

[7] N. Megiddo and M. Shub: Boundary behavior of interior point algorithms in linear programming. *Mathematics of Operations Research* 14-1 (1989) 97–146.

[8] C. Roos, T. Terlaky and J-Ph. Vial: *Theory and Algorithms for Linear Optimization: An Interior Point Approach*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley (1997).

[9] G. Sonnevend: An "analytical centre" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In: A. Prékopa, J. Szelezsán, and B. Strazicky (eds.) *System Modelling and Optimization: Proceedings of the 12th IFIP-Conference*, Budapest 1985. Lecture Notes in Control and Information Sciences 84 Springer Verlag (1986) 866–876.

[10] T. Terlaky and S. Zhang: Pivot rules for linear programming - a survey. *Annals of Operations Research* 46 (1993) 203–233.

[11] M. Todd and Y. Ye: A lower bound on the number of iterations of long-step and polynomial interior-point linear programming algorithms. *Annals of Operations Research* 62 (1996) 233–252.

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

[13] S. J. Wright: *Primal-Dual Interior-Point Methods*. SIAM Publications (1997).

[14] Y. Ye: *Interior-Point Algorithms: Theory and Analysis*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley (1997).

Antoine Deza, Eissa Nematollahi, Tamás Terlaky
Advanced Optimization Laboratory, Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada.
Email: deza, nematoe, terlaky@mcmaster.ca