A lower bound on the optimal self-concordance parameter of convex cones

Roland Hildebrand *

June 19, 2011

Abstract

Let $K \subset \mathbb{R}^n$ be a regular convex cone, let $e_1, \dots, e_n \in \partial K$ be linearly independent points on the boundary of a compact affine section of the cone, and let $x^* \in K^o$ be a point in the relative interior of this section. For $k = 1, \dots, n$, let l_k be the line through the points e_k and x^*, let y_k be the intersection point of l_k with ∂K opposite to e_k, and let z_k be the intersection point of l_k with the linear subspace spanned by all points e_l, $l = 1, \dots, n$ except e_k. We give a lower bound on the self-concordance parameter ν of logarithmically homogeneous self-concordant barriers $F : K^o \to \mathbb{R}$ on K in terms of the projective cross-ratios $q_k = (e_k, x^*; y_k, z_k)$. The previously known lower bound of Nesterov and Nemirovski can be obtained from our result as a special case. As an application, we construct an optimal barrier for the epigraph of the $||\cdot||_\infty$-norm in $\mathbb{R}^n$ and compute lower bounds on the optimal self-concordance parameters for the power cone and the epigraph of the $||\cdot||_p$-norm in $\mathbb{R}^2$.

1 Introduction

In modern convex optimization, interior point methods are the primary tool to solve conic programs. A central role in solution algorithms for conic programs over some regular (with nonempty interior, containing no lines) convex cone K is assigned to a smooth real-valued convex function $F : K^o \to \mathbb{R}$ on the interior of the cone, the *barrier*. In order to be useful for optimization, the barrier has to satisfy certain properties [5, Section 2.3]. The second and third derivative have to satisfy the self-concordance relation

$$F'''(x)[h,h,h] \le 2(F''(x)[h,h])^{3/2} \qquad \forall\ x \in K^o,\ h \in T_x K^o, \tag{1}$$

with h running through the tangent space at x. The function F has to tend to infinity as its argument tends to the boundary of the cone,

$$\lim_{x \to \partial K} F(x) = +\infty, \tag{2}$$

and it has to satisfy the logarithmic homogeneity condition

$$F(\alpha x) = -\nu \log \alpha + F(x) \qquad \forall\ \alpha > 0,\ x \in K^o. \tag{3}$$

A smooth function $F : K^o \to \mathbb{R}$ satisfying conditions (1,2,3) is called a *logarithmically homogeneous self-concordant* barrier for the cone K. The real constant ν is called the *concordance parameter* of the barrier F.

The lower the concordance parameter of a barrier, the faster are the interior point algorithms based on this barrier. For conic optimization problems over a cone K, it is therefore desirable to have barriers on K with a concordance parameter as small as possible. We call a logarithmically homogeneous self-concordant barrier on K *optimal* if it has the lowest possible concordance parameter.

Optimality of a barrier F is proven by verifying properties (1,2,3) and showing that the concordance parameter ν of F is equal to a lower bound ν_* on this parameter for the given cone. For general cones,

*LJK, Université Grenoble 1 / CNRS, 51 rue des Mathématiques, BP53, 38041 Grenoble cedex 09, France (roland.hildebrand@imag.fr, tel. +33 476 635714, fax +33 476 631263).

all lower bounds on the concordance parameter which are available today are based on a result of Nesterov and Nemirovski [5, Sect. 2.3.4]. Namely, if for some boundary point $z \in \partial K$ of the cone there exists a neighbourhood U of z and affine half-spaces $A_1, \dots, A_k \subset \mathbb{R}^n$ with $z \in \partial A_j$, $j = 1, \dots, k$, such that the normals to the half-spaces at z are linearly independent and the intersection $U \cap K$ equals the intersection $U \cap A_1 \cap \dots \cap A_k$[1], then a lower bound on the concordance parameter of any self-concordant barrier on K is given by $\nu_* = k$. Based on this result, Güler and Tunçel proved that the minimum over the Carathéodory numbers of all points in the interior of K also is a lower bound on the concordance parameter [2, Prop. 4.1]. In this way, the standard barriers for the symmetric cones used in linear, conic quadratic, and semi-definite programming are shown to be optimal. Optimal barriers can be constructed also for general homogeneous cones, with the concordance parameter equal to the rank of the cone [2, Theorem 4.1].

In this contribution, we provide a new lower bound on the optimal concordance parameter of a general cone (Theorems 4.4, 5.4). For n-dimensional cones, this lower bound is contained in the interval $[2, n]$ (Corollary 3.2 and Theorem 5.7). From our result, a slightly stronger bound than that in [5, Sect. 2.3.4] follows as special case (Theorem 6.1). In contrast to the previously known bounds, our results are non-trivial also for "round" cones, i.e., cones with a smooth boundary. The bound is a function of the projective cross-ratio of geometric objects that are constructed from n boundary points of the cone in general position and an interior point. As an application, we compute lower bounds on the optimal self-concordance parameters of the epigraph of the $||\cdot||_\infty$-norm (Corollary 6.2), the power cone (Corollary 7.1), and the epigraph of the $||\cdot||_p$-norm in $\mathbb{R}^2$ (Corollary 7.2).

The remainder of the paper is structured as follows. In the next section we introduce the projective cross-ratio and consider some of its elementary properties. In Section 3 we prove an auxiliary result, which essentially applies to 2-dimensional cones. In Section 4 we deduce the lower bound on the optimal concordance parameters of general cones and in the subsequent section we investigate its properties. In Section 6 we deduce the bounds in [5, Sect. 2.3.4] from our result, and in the last section we apply our results to the power cone and the epigraph of the $||\cdot||_p$-norm. Finally, in the appendix we construct an optimal barrier for the epigraph of the $||\cdot||_\infty$-norm.

2 The projective cross-ratio

Let $x_1, x_2, x_3, x_4 \in \mathbb{R} \cup \{\infty\}$ be four distinct points in the 1-point compactification of the real line. The *cross-ratio* of the quadruple (x_1, x_2, x_3, x_4) is defined as the number

$$(x_1, x_2; x_3, x_4) = \frac{(x_1 - x_3)(x_2 - x_4)}{(x_1 - x_4)(x_2 - x_3)} \in \mathbb{R},$$

where the differences containing the value ∞ are cancelled in the event that one of the points in the quadruple is ∞. This function can be extended continuously to a $(\mathbb{R} \cup \{\infty\})$-valued function on the set of quadruples of points in $\mathbb{R} \cup \{\infty\}$ of which no three coincide.

The cross-ratio is invariant under projective transformations of $\mathbb{R} \cup \{\infty\}$, and can hence also be considered as a $(\mathbb{R} \cup \{\infty\})$-valued function on quadruples of points on the real projective line $\mathbb{R}P^1$. Alternatively, it can be considered as a $(\mathbb{R} \cup \{\infty\})$-valued function on quadruples of 1-dimensional linear subspaces of $\mathbb{R}^2$, as the set of such subspaces is isomorphic to $\mathbb{R}P^1$.

As can be easily checked, the cross-ratio possesses the symmetry

$$(x_1, x_2; x_3, x_4) = \frac{1}{(x_2, x_1; x_3, x_4)}, \tag{4}$$

with the values 0 and ∞ being considered as reciprocal.

In the next two sections we consider the cross-ratio as defined on quadruples of coplanar lines through the origin, while in the example sections it will be more convenient to consider cross-ratios of quadruples of collinear points.

[1] Actually, in [5, Prop. 2.3.6] it is required that K is polyhedral, but this is not used in the proof.

3 2-dimensional linear sections

In this section we prove an auxiliary result which essentially provides a construction of the optimal barrier on a 2-dimensional convex cone under the condition that the direction of the gradient of the barrier at some interior point of the cone is fixed to some value.

Let $K \subset \mathbb{R}^n$, $n \geq 2$, be a regular convex cone, $x^* \in K^o$ a point in the interior of K, l a line containing x^* and intersecting the boundary ∂K of the cone in the points e, y. Denote by L the 2-dimensional linear hull of l and by l_{x^*}, l_e, l_y the 1-dimensional linear subspaces spanned by the vectors x^*, e, y, respectively. Let $F : K^o \to \mathbb{R}$ be a logarithmically homogeneous self-concordant barrier on K with concordance parameter ν, and define $p^* = -F'(x^*)$. Note that p^* is a linear functional on $\mathbb{R}^n$, located in the interior of the dual cone K^*. The kernel of p^* is an $n-1$-dimensional linear subspace $L_{p^*} \subset \mathbb{R}^n$, which intersects K in the origin. Denote the 1-dimensional intersection of L_{p^*} with the 2-dimensional subspace L by l_{p^*}. Then $l_{x^*}, l_e, l_y, l_{p^*}$ are four mutually distinct coplanar lines, and we can define their projective cross-ratio $r = (l_e, l_y; l_{x^*}, l_{p^*})$. The arrangement of the lines implies that $-\infty < r < 0$.

Lemma 3.1. *Assume above notations. Then ν is bounded from below by $\nu_* = \frac{2}{1+\frac{|r+1|}{r-1}}$.*

Proof. Let $K_L = K \cap L$ and let F_L be the restriction of F on the interior K_L^o of K_L. Then F_L is a self-concordant barrier for the 2-dimensional cone K_L with the same concordance parameter ν as K. Moreover, the linear functional $p_L^* = -F_L'(x^*)$ is the restriction of the linear functional p^* on L, and hence its kernel coincides with l_{p^*}. The assertion of the lemma for the cone K thus reduces to the assertion for the cone K_L.

In order to avoid unnecessary notations, we shall assume without restriction of generality that $n = 2$ and hence $K = K_L$. Let $\gamma_* = \frac{\nu_*-2}{\sqrt{\nu_*-1}} = \frac{|r+1|}{\sqrt{-r}}$ and let $\lambda_\pm^* = -\frac{\gamma_*}{2} \pm \sqrt{\frac{\gamma_*^2}{4}+1}$ be the roots of the quadratic equation $\lambda^2 + \gamma_*\lambda - 1 = 0$. Then $-r\gamma_*^2 = (r+1)^2$ and hence $r = -\frac{\gamma_*^2+2}{2} \pm \gamma_*\sqrt{\frac{\gamma_*^2}{4}+1} = -(\lambda_\mp^*)^2$. Since $\gamma_* \geq 0$, we have $\lambda_-^* \leq -1$ and $0 < \lambda_+^* \leq 1$. Therefore $r = -(\lambda_-^*)^2$ if $r \leq -1$ and $r = -(\lambda_+^*)^2$ if $r \geq -1$. Note that in the former case we have $(l_y, l_e; l_{x^*}, l_{p^*}) = r^{-1} = -(\lambda_+^*)^2$ by (4).

Let us now introduce a coordinate system in $\mathbb{R}^2$ such that the lines l_{x^*}, l_{p^*} are parallel to the basis vectors $(1,0)^T, (0,1)^T$, respectively. Let further the lines l_e, l_y be parallel to the vectors $(1, \lambda_\pm^*)^T$, respectively, if $r \geq -1$, and to the vectors $(1, \lambda_\mp^*)^T$, respectively, if $r < -1$. This is always possible because

$$(\lambda_+^*, \lambda_-^*; 0, \infty) = \frac{(\lambda_+^* - 0)(\lambda_-^* - \infty)}{(\lambda_-^* - 0)(\lambda_+^* - \infty)} = \frac{\lambda_+^*}{\lambda_-^*} = -(\lambda_+^*)^2 = \begin{cases} (l_y, l_e; l_{x^*}, l_{p^*}), & r < -1; \\ (l_e, l_y; l_{x^*}, l_{p^*}), & r \geq -1, \end{cases}$$

and the cross-ratio of a quadruple of distinct lines through the origin of a plane is the only invariant of the quadruple with respect to the general linear group of this plane. Let us further scale the coordinates by a homothety such that $x^* = (1,0)^T$. In this coordinate system the cone K is given by the set $\{x = (x_1, x_2)^T \mid x_1 \geq 0,\ \lambda_-^* x_1 \leq x_2 \leq \lambda_+^* x_1\}$, and $p^* = -F'(x^*) = (\nu, 0)^T$ due to the identity $\langle p^*, x^* \rangle = \nu$ [5, eq. (2.3.13)].

Consider the 1-dimensional Lie group $A(t) = \exp(t \cdot a) = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$ generated by the Lie algebra element $a = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Let $x(t) = A(t)x^* = (1,t)^T$ and consider the scalar $f(t) = \nu^{-1}F(x(t))$. We shall now establish a differential inequality which is satisfied by the function $f(t)$. Formulas [5, eq. (2.3.12–14)] and its derivatives yield the relations $F'(x)[x] = -\nu$, $F''(x)[x,\cdot] = -F'(x)$, $F'''(x)[x,\cdot,\cdot] = -2F''(x)$. At $x(t) = A(t)x^*$ we thus have

$$F' = \nu A^T(-t)\begin{pmatrix} -1 \\ \alpha \end{pmatrix}, \quad F'' = \nu A^T(-t)\begin{pmatrix} 1 & -\alpha \\ -\alpha & \beta \end{pmatrix} A(-t),$$

$$F'''[A(t)h,\cdot,\cdot] = -2\nu A^T(-t)\left[\begin{pmatrix} 1 & -\alpha \\ -\alpha & \beta \end{pmatrix} h_1 + \begin{pmatrix} -\alpha & \beta \\ \beta & \mu \end{pmatrix} h_2\right] A(-t) \tag{5}$$

for some $\alpha, \beta, \mu \in \mathbb{R}$. Here $\beta > \alpha^2$, because $F'' \succ 0$, and $h = (h_1, h_2)^T \in \mathbb{R}^2$ is an arbitrary vector. Condition (1), applied to the vector $A(t)h$, then yields

$$-h_1^3 + 3\alpha h_1^2 h_2 - 3\beta h_1 h_2^2 - \mu h_2^3 \le \sqrt{\nu}(h_1^2 - 2\alpha h_1 h_2 + \beta h_2^2)^{3/2}.$$

Setting $h = (\alpha - \sqrt{\frac{\beta-\alpha^2}{\nu-1}}, 1)^T$, we obtain after some calculations

$$-(\alpha^3 + 3\alpha(\beta - \alpha^2) + \mu) \le \gamma(\beta - \alpha^2)^{3/2}, \tag{6}$$

where $\gamma = \frac{\nu-2}{\sqrt{\nu-1}}$. Let us compute the derivatives of $f(t)$. By (5) and using the fact that $\dot{x} = (0, 1)^T$ is constant, we have

$$\dot{f} = \nu^{-1} F'(x)[ax] = \alpha, \quad \ddot{f} = \nu^{-1} F''(x)[ax, ax] = \beta, \quad f^{(3)} = \nu^{-1} F'''(x)[ax, ax, ax] = -2\mu.$$

Inserting this into (6), we get

$$4\dot{f}^3 - 6\dot{f}\ddot{f} + f^{(3)} \le 2\gamma(\ddot{f} - \dot{f}^2)^{3/2}. \tag{7}$$

We would like to bound the function $f(t)$ by the solution of the differential equation which is obtained when one assumes equality in (7). This is accomplished by introducing the variables $q_\pm = -\frac{\sqrt{\ddot{f}-\dot{f}^2}}{\lambda_\mp} - \dot{f}$, where $\lambda_\pm = -\frac{\gamma}{2} \pm \sqrt{\frac{\gamma^2}{4} + 1}$ are the roots of the quadratic equation $\lambda^2 + \gamma\lambda - 1 = 0$. These variables satisfy the relation $q_+ > q_-$ and $\dot{f}, \ddot{f}$ can be recovered from them by the formula

$$\dot{f} = -\frac{\lambda_+}{\lambda_+ - \lambda_-} q_- + \frac{\lambda_-}{\lambda_+ - \lambda_-} q_+, \qquad \ddot{f} = \frac{\lambda_+}{\lambda_+ - \lambda_-} q_-^2 - \frac{\lambda_-}{\lambda_+ - \lambda_-} q_+^2. \tag{8}$$

We have $\ddot{f} - \dot{f}^2 = \frac{(q_+ - q_-)^2}{\gamma^2+4}$, and hence $(\ddot{f} - \dot{f}^2)^{3/2} = \frac{(q_+-q_-)^3}{(\gamma^2+4)^{3/2}}$. A simple calculus then shows that $2\gamma(\ddot{f} - \dot{f}^2)^{3/2} - 4\dot{f}^3 + 6\dot{f}\ddot{f} = -\frac{2\lambda_+}{\lambda_+-\lambda_-} q_-^3 + \frac{2\lambda_-}{\lambda_+-\lambda_-} q_+^3$ and hence by (7)

$$f^{(3)} \le -\frac{2\lambda_+}{\lambda_+ - \lambda_-} q_-^3 + \frac{2\lambda_-}{\lambda_+ - \lambda_-} q_+^3. \tag{9}$$

Recall that $p^* = (\nu, 0)^T$ and hence $\dot{f}(0) = 0$. Hence by (8) we have

$$\lambda_+ q_-(0) = \lambda_- q_+(0). \tag{10}$$

Differentiating the first relation in (8) with respect to t and expressing $\ddot{f}$ by the second relation, we obtain

$$\frac{\lambda_+}{\lambda_+ - \lambda_-}(q_-^2 + \dot{q}_-) - \frac{\lambda_-}{\lambda_+ - \lambda_-}(q_+^2 + \dot{q}_+) = 0. \tag{11}$$

Differentiating the second relation in (8) and inserting into (9) yields the inequality

$$\frac{\lambda_+}{\lambda_+ - \lambda_-} q_-(q_-^2 + \dot{q}_-) - \frac{\lambda_-}{\lambda_+ - \lambda_-} q_+(q_+^2 + \dot{q}_+) \le 0. \tag{12}$$

Combining (11) with (12) yields

$$\frac{\lambda_+}{\lambda_+ - \lambda_-}(q_- - q_+)(q_-^2 + \dot{q}_-) \le 0, \quad \frac{\lambda_-}{\lambda_+ - \lambda_-}(q_- - q_+)(q_+^2 + \dot{q}_+) \le 0,$$

which by the relation $q_+ > q_-$ gives

$$q_-^2 + \dot{q}_- \ge 0, \qquad q_+^2 + \dot{q}_+ \le 0. \tag{13}$$

The solution of the differential equation $q^2 + \dot{q} = 0$ is given by $q(t) = \frac{1}{t+q^{-1}(0)}$. The differential inequalities (13) then yield the bounds

$$q_-(t) \ge \frac{1}{t + q_-^{-1}(0)}, \quad t \ge 0; \qquad q_+(t) \ge \frac{1}{t + q_+^{-1}(0)}, \quad t \le 0. \tag{14}$$

By (10) and the condition $q_+ > q_-$ we have $q_-(0) < 0$ and $q_+(0) > 0$. Note that $\lim_{t\to\lambda_+^*} f(t) = +\infty$, hence $\lim_{t\to\lambda_+^*} \dot f(t) = +\infty$ and $\lim_{t\to\lambda_+^*} q_-(t) = -\infty$. The right-hand side in the first relation in (14) tends to $-\infty$ for $t \to -q_-^{-1}(0)$, and hence $\lambda_+^* \geq -q_-^{-1}(0)$. Likewise, $\lim_{t\to\lambda_-^*} f(t) = +\infty$, hence $\lim_{t\to\lambda_-^*} \dot f(t) = -\infty$ and $\lim_{t\to\lambda_-^*} q_+(t) = +\infty$. The right-hand side in the second relation in (14) tends to $+\infty$ for $t \to -q_+^{-1}(0)$, and hence $\lambda_-^* \geq -q_+^{-1}(0)$.

We thus get $\frac{\lambda_+^*}{\lambda_+} \geq -\frac{1}{\lambda_+ q_-(0)}$, $-\frac{\lambda_-^*}{\lambda_-} \geq \frac{1}{\lambda_- q_+(0)}$. Combining with (10), this yields $\frac{\lambda_+^*}{\lambda_+} \geq \frac{\lambda_-^*}{\lambda_-}$. Finally, inserting the values $\lambda_- = -\sqrt{\nu - 1}$, $\lambda_+ = \frac{1}{\sqrt{\nu-1}}$, $\lambda_-^* = -\sqrt{\nu_* - 1}$, $\lambda_+^* = \frac{1}{\sqrt{\nu_*-1}}$, we arrive at the desired conclusion. □

Corollary 3.2. *Let $K \subset \mathbb{R}^n$, $n \geq 2$, be a regular convex cone, and $F : K^o \to \mathbb{R}$ a logarithmically homogeneous self-concordant barrier on K with concordance parameter ν. Then $\nu \geq 2$.* □

If $n = 2$ and the bound on the concordance parameter given by Lemma 3.1 is saturated, then $\lambda_+^* = \lambda_+ = -q_-^{-1}(0)$ and $\lambda_-^* = \lambda_- = -q_+^{-1}(0)$. Inequalities (14) must be saturated too and $\dot f(t) = -\frac{\lambda_+}{\lambda_+-\lambda_-}\frac{1}{t-\lambda_+} + \frac{\lambda_-}{\lambda_+-\lambda_-}\frac{1}{t-\lambda_-}$. This can be integrated, yielding $f(t) = -\frac{\lambda_+}{\lambda_+-\lambda_-}\log(t-\lambda_+) + \frac{\lambda_-}{\lambda_+-\lambda_-}\log(t-\lambda_-) + const$, thus determining the barrier F up to an additive constant. It is not hard to check that F is invariant under the action of the Lie group generated by the Lie algebra element $a = \begin{pmatrix} 0 & 1 \\ 1 & -\gamma \end{pmatrix}$. This group acts transitively on the set of rays constituting the interior of the cone K, and hence the expression ν_* in Lemma 3.1 is independent of the choice of the interior point x^* in this case.

4 Main result

In this section a lower bound on the concordance parameter ν for a logarithmically homogeneous self-concordant barrier $F : K^o \to \mathbb{R}$ on a given regular convex cone $K \subset \mathbb{R}^n$ is obtained from Lemma 3.1 by the following consideration. Let $x^* \in K^o$ be an interior point of K, denote $p^* = -F'(x^*)$, and let $L_1, \dots, L_m$ be 2-dimensional linear subspaces containing x^*. For fixed p^* and for each $i = 1, \dots, m$, application of Lemma 3.1 to the intersection $K_{L_i} = L_i \cap K$ gives rise to a lower bound $\nu_i^*(p^*)$ on ν, and hence $\nu \geq \max_{i=1,\dots,m} \nu_i^*(p^*)$. If m is smaller than the dimension n of the cone K, then for subspaces $L_1, \dots, L_m$ in general position p^* can be chosen such that $\max_{i=1,\dots,m} \nu_i^*(p^*) = 2$, and no information is gained with respect to Corollary 3.2. If, however, $m \geq n$, then $\min_{p^*} \max_{i=1,\dots,m} \nu_i^*(p^*) > 2$ in general, except in the case that K is the Lorentz cone. The goal of this section is to solve the minimax problem $\min_{p^*} \max_{i=1,\dots,m} \nu_i^*(p^*)$ for the case $m = n$. We will work under the following assumption.

Assumption 4.1. *Let $K \subset \mathbb{R}^n$ be a regular convex cone, let $x^* = (x_1^*, \dots, x_n^*)^T \in K^o$, and let $L_1, \dots, L_n \subset \mathbb{R}^n$ be 2-dimensional linear subspaces containing x^*. Assume that there exist linearly independent vectors $e_i \in L_i \cap \partial K$, $i = 1, \dots, n$, and let $y_i \in L_i \cap \partial K$ be such that $L_i = \mathrm{span}\{e_i, y_i\}$, $i = 1, \dots, n$. Let l_k be the line through the points e_k and y_k, and let z_k be the intersection point of l_k with the $(n-1)$-dimensional linear subspace $\hat L_k$ spanned by all points e_l, $l = 1, \dots, n$ except e_k. Let $l_{e_k}, l_{y_k}, l_{z_k}$, and l_{x^*} be the linear hulls of e_k, y_k, z_k, and x^*, respectively. Then for each $k = 1, \dots, n$ the 4 lines $l_{e_k}, l_{y_k}, l_{z_k}, l_{x^*}$ are in L_k and hence coplanar, and no three of them are identical. Let $q_k = [l_{e_k}, l_{x^*}; l_{y_k}, l_{z_k}]$ be their cross-ratio.*

Now let $F : K^o \to \mathbb{R}$ be a logarithmically homogeneous self-concordant barrier with parameter ν, set $p^* = -F'(x^*)$, and let $l_{p_k^*}$ be the 1-dimensional intersection of the plane L_k with the kernel of the linear functional p^*. We shall express the cross-ratio $r_k = [l_{e_k}, l_{y_k}; l_{x^*}, l_{p_k^*}]$, which determines the bound $\nu_k^*(p^*)$ from Lemma 3.1, in terms of the cross-ratios q_k and the linear functional p^*.

Introduce a coordinate system in $\mathbb{R}^n$ with basis vectors equal to e_k. Let p_k^* be the elements of p^* in these coordinates. Note that p^* is in the interior of the dual cone K^*, and hence $\langle p^*, e_k\rangle = p_k^* > 0$ for all $k = 1, \dots, n$.

In each of the planes L_k, consider the affine line $\tilde l_k$ through x^* which is parallel to l_{e_k}. Introduce a real parameter λ on $\tilde l_k$, putting the number λ in correspondence with the point $x^* + \lambda e_k$. Then the intersection point of $\tilde l_k$ with l_{e_k} is the infinitely remote point and can be assigned the parameter

$\lambda = +\infty$. The intersection point of $\tilde{l}_k$ with l_{x^*} has parameter value $\lambda = 0$. The coordinate vector of the intersection point of $\tilde{l}_k$ with l_{z_k} has a zero at its k-th entry and thus has parameter value $\lambda = -x_k^*$. Let λ_k be the parameter value corresponding to the intersection point of $\tilde{l}_k$ with l_{y_k}. Since the segment of $\tilde{l}_k$ corresponding to the parameter values $\lambda \in [0,\infty)$ lies in the interior of K, we have $\lambda_k < 0$. Finally, the intersection point of $\tilde{l}_k$ with $l_{p_k^*}$ is given by the equation $\langle p^*, x^* + \lambda e_k\rangle = 0$, which yields the value $\hat{\lambda}_k = -\frac{\langle p^*,x^*\rangle}{p_k^*}$ for the parameter of this intersection point. Note that $\hat{\lambda}_k < \lambda_k$, because p^* has to be positive on y_k. We then obtain the cross-ratios

$$q_k = [l_{e_k}, l_{x^*}; l_{y_k}, l_{z_k}] = \frac{(\infty - \lambda_k)(0 - (-x_k^*))}{(\infty - (-x_k^*))(0 - \lambda_k)} = -\frac{x_k^*}{\lambda_k}, \tag{15}$$

$$r_k = [l_{e_k}, l_{y_k}; l_{x^*}, l_{p_k^*}] = \frac{(\infty - 0)(\lambda_k - \hat{\lambda}_k)}{(\infty - \hat{\lambda}_k)(\lambda_k - 0)} = 1 - \frac{\hat{\lambda}_k}{\lambda_k} = 1 + \frac{\langle p^*, x^*\rangle}{p_k^*\lambda_k}, \tag{16}$$

from which we get

$$\frac{r_k+1}{r_k-1} = 1 + \frac{2p_k^*\lambda_k}{\langle p^*, x^*\rangle}. \tag{17}$$

Lemma 4.2. *The relation $\sum_{k\,:\,q_k>0} q_k > 1$ holds.*

Proof. Let $I_+ = \{k \,|\, q_k > 0\}$ and $Q_+ = \sum_{k\in I_+} q_k$.

If $I_+ = \emptyset$, then x^* is in the nonpositive orthant by (15). Thus $-x^*$ is also in K as a linear combination of the basis vectors e_k with nonnegative coefficients. This contradicts the regularity of K.

Hence $I_+ \neq \emptyset$. For every k, $x^* + \lambda_k e_k$ is a positive multiple of y_k and hence equals a nonzero vector in ∂K. Consider the point $s = \sum_{k\in I_+} q_k(x^* + \lambda_k e_k) = (Q_+ - 1)x^* + \left(x^* - \sum_{k\in I_+} x_k^* e_k\right)$, where the second relation comes from (15). As a linear combination of nonzero elements in K, with positive coefficients, the point s is also a nonzero vector in K. The vector $x^* - \sum_{k\in I_+} x_k^* e_k$ has only nonpositive components and is hence in $-K$. It follows that $(Q_+ - 1)x^* = s - \left(x^* - \sum_{k\in I_+} x_k^* e_k\right)$ is a nonzero vector in K, which implies $Q_+ > 1$ and proves the lemma. □

By Lemma 3.1, the concordance parameter ν of the barrier F is bounded from below by the expression $\max_{k=1,\dots,n} \frac{2}{1-\left|\frac{r_k+1}{r_k-1}\right|}$. This bound still depends on the negative gradient p^* of F at x^*. A lower bound on the concordance parameter of an arbitrary logarithmically homogeneous self-concordant barrier on K is then given by

$$\nu_* = \min_{p^*}\max_k \frac{2}{1 - \left|\frac{r_k+1}{r_k-1}\right|} = \frac{2}{1 - \min_{p^*}\max_k \left|\frac{r_k+1}{r_k-1}\right|}, \tag{18}$$

where p^* is subject to the constraints $\langle p^*, e_k\rangle > 0$, $\langle p^*, y_k\rangle > 0$ for all $k = 1,\dots,n$. By (17) this transforms into

$$\nu_* = \frac{2}{1 - \min_{p^*}\max_k \left|1 + \frac{2p_k^*\lambda_k}{\langle p^*,x^*\rangle}\right|}, \tag{19}$$

where the components of p^* have to satisfy the requirements $p_k^* > 0$ and $-\frac{\langle p^*,x^*\rangle}{p_k^*} = \hat{\lambda}_k < \lambda_k$. Equivalently, $0 < \frac{p_k^*}{\langle p^*,x^*\rangle} < -\frac{1}{\lambda_k}$.

Introduce variables $\alpha_k = 1 + \frac{2p_k^*\lambda_k}{\langle p^*,x^*\rangle} \in (-1,1)$. These variables have to satisfy the additional requirement $\sum_{k=1}^n \frac{\alpha_k - 1}{2\lambda_k} x_k^* = 1$, which by (15) is equivalent to $\sum_{k=1}^n \frac{1-\alpha_k}{2} q_k = 1$. We shall now solve the minimax problem

$$\min\left\{\max_k |\alpha_k| \;\middle|\; \alpha_k \in (-1,1),\ \sum_{k=1}^n \frac{1-\alpha_k}{2} q_k = 1\right\}. \tag{20}$$

Lemma 4.3. *The value of problem (20) is given by $\frac{|\sum_k q_k - 2|}{\sum_k |q_k|} < 1$.*

Proof. First note that $\sum_k |q_k| \geq \sum_k q_k$ and hence $\sum_k q_k - 2 < \sum_k |q_k|$. On the other hand, by Lemma 4.2 we have $\sum_k (q_k + |q_k|) > 2$ and hence $2 - \sum_k q_k < \sum_k |q_k|$. It follows that $\frac{|\sum_k q_k - 2|}{\sum_k |q_k|} \in [0,1)$. It is easily checked that this value is attained by the solution $\alpha_k = \alpha_k^* = \frac{\sum_l q_l - 2}{\sum_l |q_l|} \operatorname{sgn} q_k \in (-1,1)$.

On the other hand, every feasible vector $\alpha = (\alpha_1, \dots, \alpha_n)^T$ has to satisfy the constraint $\sum_k \alpha_k q_k = \sum_k q_k - 2$. It follows that $(\max_k |\alpha_k|) \sum_k |q_k| \geq \sum_k |\alpha_k||q_k| \geq |\sum_k q_k - 2|$, which yields $\max_k |\alpha_k| \geq \frac{|\sum_k q_k - 2|}{\sum_k |q_k|}$ and thus proves optimality of the solution $\alpha_k = \alpha_k^*$. □

Inserting the optimal value of (20) into (19), we obtain the following theorem.

Theorem 4.4. *Under Assumption 4.1, the concordance parameter ν of any logarithmically homogeneous self-concordant barrier F on the cone K is bounded from below by the quantity*

$$\nu_* = \frac{2}{1 - \frac{|\sum_k q_k - 2|}{\sum_k |q_k|}}.$$

5 Properties of the lower bound

In this section we will consider the bound given by Theorem 4.4 in more detail.

Lemma 5.1. *Assume the conditions of Assumption 4.1. Let $I \subset \{1, \dots, n\}$ be such that the vectors in the set $\{y_k \mid k \in I\} \cup \{e_k \mid k \notin I\}$ are linearly independent. Then the bound in Theorem 4.4 is invariant under an interchange of the points e_k and y_k for all $k \in I$.*

Proof. The exchange of y_k and e_k leaves the subspace L_k and the line $l_{p_k^*}$ invariant and hence by (4) leads to the transformation $r_k \mapsto r_k^{-1}$ in (16). This in turn leads to the transformation $\frac{r_k+1}{r_k-1} \mapsto -\frac{r_k+1}{r_k-1}$. The set of constraints $\langle p^*, e_k \rangle > 0$, $\langle p^*, y_k \rangle > 0$ is also left invariant, and hence by (18) the value of ν_* is left unchanged. □

Remark 5.2. The assertion of Lemma 5.1 cannot be easily inferred directly from the explicit expression of the bound ν_* in Theorem 4.4, because the exchange of e_k and y_k for one index k changes the lines l_{z_l} for all $l \neq k$ and hence all cross-ratios q_l in (15) are changed.

Lemma 5.3. *Assume the conditions of Assumption 4.1. There exists a subset $I \subset \{1, \dots, n\}$ of indices with complement $\bar{I}$ such that the set $\{e_k \mid k \in \bar{I}\} \cup \{y_k \mid k \in I\}$ is linearly independent, and x^* is a linear combination of the vectors in this set with nonnegative coefficients.*

Proof. The assertion of the lemma will follow from the statement that we can render all entries of x^* nonnegative by exchanging the vectors e_k and y_k for a number of indices k and adapting the coordinate system accordingly. Let us prove this statement.

Suppose there exists an index k such that $x_k^* < 0$. Recall that z_k is the unique point on l_k situated in the linear subspace $\hat{L}_k$ spanned by all e_l except e_k. Therefore, linear dependence of the vectors $e_1, \dots, e_{k-1}, y_k, e_{k+1}, \dots, e_n$ is equivalent to the relation $z_k = y_k$. Since $x_k^* < 0$, the line l_{x^*} intersects l_k in a point opposite to e_k with respect to z_k. But y_k is opposite to e_k with respect to l_{x^*} in L_k, hence $z_k \neq y_k$. This proves that we can exchange the roles of e_k and y_k without violating the condition of linear independence of the points e_l.

Assume without restriction of generality that x^* is situated on the line segment between e_k and y_k. This is equivalent to multiplication of x^* by a positive constant and does not change the signs of the entries x_k^*. We have the explicit expression $z_k = \frac{x^* - x_k^* e_k}{1 - x_k^*}$, deriving from the condition that z_k is the affine combination of e_k and x^* whose k-th entry vanishes. On the other hand, we have $y_k = \frac{x^* + \lambda_k e_k}{1 + \lambda_k}$, deriving from the condition that y_k is the affine combination of e_k and x^* which is a multiple of $x^* + \lambda_k e_k$. It follows that

$$x^* = \frac{x_k^*(1+\lambda_k)}{\lambda_k + x_k^*} y_k + \frac{\lambda_k(1 - x_k^*)}{\lambda_k + x_k^*} z_k = \frac{x_k^*(1+\lambda_k)}{\lambda_k + x_k^*} y_k + \frac{\lambda_k}{\lambda_k + x_k^*}(x^* - x_k^* e_k). \tag{21}$$

Note that $x^* - x_k^* e_k$ is a linear combination of the points $e_1, \dots, e_{k-1}, e_{k+1}, \dots, e_n$, with coefficients being equal to the corresponding entries of x^*. From (21) it then follows that the coefficients $\tilde{x}_l^*$ of x^*, when expressed as a linear combination of the vectors $e_1, \dots, e_{k-1}, y_k, e_{k+1}, \dots, e_n$, are given by

$$\tilde{x}_k^* = \frac{x_k^*(1+\lambda_k)}{\lambda_k + x_k^*}, \qquad \tilde{x}_l^* = \frac{\lambda_k}{\lambda_k + x_k^*} x_l^*, \ l \neq k.$$

Since x^* is situated between y_k and z_k on the line l_k and is different from these points, we have by (21) that $\frac{x_k^*(1+\lambda_k)}{\lambda_k + x_k^*} > 0$, $\frac{\lambda_k(1-x_k^*)}{\lambda_k + x_k^*} > 0$, and hence also $\frac{\lambda_k}{\lambda_k + x_k^*} > 0$. Therefore the sign of $\tilde{x}_l^*$ equals that of x_l^* for $l \neq k$, while $\tilde{x}_k^*$ is positive.

As a consequence, exchanging e_k and y_k has lead to a decrease in the number of negative entries of x^* by one. Repeating this process, we can eliminate all negative entries of x^*. □

Theorem 5.4. *In addition to Assumption 4.1, suppose that x^* is contained in the simplicial cone generated by the vectors e_k. Then the lower bound in Theorem 4.4 is given by*

$$\nu_* = \begin{cases} \frac{\sum_k q_k}{\sum_k q_k - 1}, & \sum_k q_k \leq 2; \\ \sum_k q_k, & \sum_k q_k \geq 2. \end{cases} \tag{22}$$

Proof. Choose a coordinate system in $\mathbb{R}^n$ with basis vectors e_k, $k = 1, \dots, n$. Then by assumption of the theorem all entries of x^* are nonnegative, and by (15) all cross-ratios q_k are nonnegative. Then the bound in Theorem 4.4 simplifies to (22). □

We now consider the situation when $q_k = 0$ for some k. In this case $x^* = z_k$ is contained in the $(n-1)$-dimensional linear subspace $\hat{L}_k \subset \mathbb{R}^n$ spanned by all e_l except e_k. For every $l \neq k$, since both e_l, x^* are in $\hat{L}_k$, we also have $y_l \in \hat{L}_k$. We can then apply the construction of the previous section to the $n-1$ 2-dimensional subspaces $L_1, \dots, L_{k-1}, L_{k+1}, \dots, L_n$ of the $n-1$-dimensional cone $\tilde{K} = K \cap \hat{L}_k$. Since all q_l, $l \neq k$, retain their values in the lower-dimensional space, and $\sum_{l=1}^n q_l = \sum_{l \neq k} q_l$, $\sum_{l=1}^n |q_l| = \sum_{l \neq k} |q_l|$ by $q_k = 0$, Theorem 4.4 will yield the same bound ν_* for the cone $\tilde{K}$ as it has for the cone K. In other words, in the case $q_k = 0$ the bound given by Theorem 4.4 for the cone K is essentially a consequence of a similar bound for the cone $\tilde{K}$, which is a linear section of K with codimension 1.

Lemma 5.5. *Assume the conditions of Theorem 5.4. Then for all $k = 1, \dots, n$ we have $q_k \leq 1$, and for all index sets $I \subset \{1, \dots, n\}$ of cardinality $n-1$ we have $\sum_{k \in I} q_k \geq 1$.*

Proof. Assume the notations of the previous section. Since the simplicial cone K_S generated by the vectors e_k is contained in K, the interval $I_S = l_k \cap K_S$ is contained in the interval $I_K = l_k \cap K$ for every $k = 1, \dots, n$. Note that y_k is the endpoint of I_K opposite to e_k, while z_k is the endpoint of I_S opposite to e_k. Hence y_k is situated on l_k opposite to e_k with respect to z_k (but it may coincide with z_k). Recall that the parameter of the intersection point of l_{e_k} with $\tilde{l}_k$ was $\lambda = +\infty$, the intersection point of l_{z_k} with $\tilde{l}_k$ had parameter $\lambda = -x_k^*$, and the intersection point of l_{y_k} with $\tilde{l}_k$ had parameter $\lambda = \lambda_k$. Therefore $\lambda_k \leq -x_k^*$, which by (15) yields $q_k \leq 1$.

Let us prove the second part by contradiction. Suppose there exists a subset $I \subset \{1, \dots, n\}$ of $n-1$ indices such that $\sum_{k \in I} q_k < 1$. For $l \in I$, define $a_l = \frac{q_l}{1 - \sum_{k \in I} q_k} \geq 0$. We then have

$$x^* + \sum_{l \in I} a_l (1+\lambda_l) y_l = x^* + \sum_{l \in I} \frac{q_l}{1 - \sum_{k \in I} q_k}(x^* + \lambda_l e_l) = \frac{1}{1 - \sum_{k \in I} q_k} \left(x^* - \sum_{l \in I} x_l^* e_l \right),$$

where for the second equality we used (15). The leftmost side is the sum of an interior point of K and boundary points of K and is hence also an interior point of K. On the other hand, the rightmost side is proportional to $e_{\hat{k}}$, where $\hat{k}$ is the index missing in I, and is hence a boundary point of K. This leads to a contradiction, thus proving the lemma. □

Corollary 5.6. *Assume the conditions of Theorem 5.4. Then the bound given in Theorem 4.4 satisfies* $\nu_* \leq n$*. The equality* $\nu_* = n$ *holds if and only if* K *is either a simplicial cone with generators* $e_1, \dots, e_n$*, or a simplicial cone with generators* $y_1, \dots, y_n$*.*

Proof. By (22) the first assertion of the corollary is equivalent to the inequalities

$$\frac{n}{n-1} \leq \sum_k q_k \leq n.$$

These can be obtained by summing the inequalities $q_k \leq 1$, $\sum_{k\in I} q_k \geq 1$ from Lemma 5.5 over all indices k or all index sets I of cardinality $n-1$, respectively.

Assume now that $\nu_* = n$. By (22) we then have either $\sum_k q_k = \frac{n}{n-1}$ or $\sum_k q_k = n$.

Let us consider the first case. By Lemma 5.5 we have $q_k = \frac{1}{n-1}$ for all k, and hence $\lambda_k = -(n-1)x_k^*$. Note that y_k is a positive multiple of the point $x^* + \lambda_k e_k$, hence $y_k = \beta_k(x^* - (n-1)x_k^* e_k)$, $\beta_k > 0$ for all k. Let $\hat{k} \in \{1, \dots, n\}$ and $I = \{1, \dots, n\} \setminus \{\hat{k}\}$. Then we have

$$\sum_{k\in I} \beta_k^{-1} y_k = \sum_{k\in I} (x^* - (n-1)x_k^* e_k) = (n-1)x_{\hat{k}}^* e_{\hat{k}},$$

and $e_{\hat{k}}$ is contained in the relative interior of the convex cone generated by the set $\{y_k \,|\, k \in I\}$. But $e_{\hat{k}} \in \partial K$, which implies that this cone is entirely contained in ∂K. Repeating this argument for all $\hat{k}$, we see that the boundary of the simplicial cone generated by the y_k is contained in ∂K. On the other hand, $\sum_{k=1}^n \beta_k^{-1} y_k = \sum_{k=1}^n (x^* - (n-1)x_k^* e_k) = x^* \in K^o$, which implies that the points y_k are linearly independent and that K equals the cone generated by the y_k.

We now pass to the case $\sum_k q_k = n$. By Lemma 5.5 we have $q_k = 1$ for all k. This is equivalent to the relations $y_k = z_k$, and y_k is contained in the cone generated by the set $\{e_l \,|\, l \neq k\}$ for all k. Moreover, it follows that $z_k \in \partial K$ and hence $z_k \neq x^*$, which implies $x_k^* > 0$ for all k. By $\lambda_k = -x_k^*$ we then have that y_k is a positive multiple of $x^* - x_k^* e_k = \sum_{l\neq k} x_l^* e_l$, and y_k is in the relative interior of the cone generated by the set $\{e_l \,|\, l \neq k\}$. As in the previous paragraph, this whole cone must then be in ∂K for all k, and ∂K contains the boundary of the simplicial cone generated by the e_k. Thus K equals this simplicial cone. This proves one direction of the equivalence asserted in the second part of the corollary.

Let us prove the opposite direction. Assume that K is a simplicial cone generated by either the vectors e_k or the vectors y_k. By possibly exchanging the roles of e_k and y_k for all k, we by virtue of Lemma 5.1 can assume that K is generated by the e_k. The line l_k intersects ∂K in e_k and in the face opposite to e_k, which implies that the second intersection point y_k coincides with z_k. But then $\lambda_k = -x_k^*$ and $q_k = 1$ for all k, which by (22) yields the assertion to be proven. □

Theorem 5.7. *Assume the conditions of Assumption 4.1. The lower bound* ν_* *given in Theorem 4.4 cannot exceed the dimension* n *of the cone* K*. The equality* $\nu_* = n$ *holds if and only if* K *is a simplicial cone with generators* $\tilde{e}_1, \dots, \tilde{e}_n$ *such that* $\tilde{e}_k \in \{e_k, y_k\}$ *for every* $k = 1, \dots, n$*.*

Proof. The theorem is a consequence of Lemmas 5.1, 5.3, and Corollary 5.6. □

6 Relation with Nesterovs and Nemirovskis bound

In this section we deduce the lower bound of Nesterov and Nemirovski, applied to convex cones, from our result. Proposition 2.3.6 in [5] states that if $C \subset \mathbb{R}^n$ is a convex polytope and $e_n \in \partial C$ is a boundary point belonging to exactly k $(n-1)$-dimensional faces of C, such that the normal vectors to these faces are linearly independent, then k is a lower bound on the concordance parameter of any self-concordant barrier on C. We are interested in the situation when C is a cone and prove the following slightly stronger result for this case.

Theorem 6.1. *Let* $K \subset \mathbb{R}^n$*,* $n \geq 3$*, be a regular polyhedral convex cone, let* $e_{k+1} \in \partial K$ *be a nonzero boundary point belonging to exactly* k $(n-1)$*-dimensional faces of* K*,* $2 \leq k \leq n-1$*, such that the normals to these faces are linearly independent. Then* $\nu_* = k+1$ *is a lower bound on the concordance parameter of any logarithmically homogeneous self-concordant barrier on* K*.*

Proof. By an appropriate choice of coordinates, we can assume that $e_{k+1} = (1, 0, \dots, 0)^T$, the set $\{x = (x_0, \dots, x_{n-1})^T \in K \,|\, x_0 = 1\}$ is a compact section of K, and there exists a neighbourhood U of e_{k+1} such that $x \in U \cap K$ if and only if $x \in U$ and $x_j \leq 0$, $j = 1, \dots, k$. Let us further assume without restriction of generality that $k = n-1$, otherwise we replace the cone K by the $(k+1)$-dimensional intersection $K \cap L$, where $L \subset \mathbb{R}^n$ is the $(k+1)$-dimensional linear subspace determined by the equations $x_{k+1} = \dots = x_{n-1} = 0$.

Set $x^* = (1, -\varepsilon, \dots, -\varepsilon)^T$, $e_j = (1, -\varepsilon, \dots, -\varepsilon, 0, -\varepsilon, \dots, -\varepsilon)^T$, $j = 1, \dots, n-1$, where the zero is located at position j. For small enough $\varepsilon > 0$ we then have $x^* \in K^o$, $e_j \in \partial K$, $j = 1, \dots, n$. Let further l_j be the line through x^* and e_j and let y_j be the intersection point of l_j with ∂K opposite to e_j, $j = 1, \dots, n$. We then have $y_j = (1, -\varepsilon, \dots, -\varepsilon, -\alpha_j, -\varepsilon, \dots, -\varepsilon)^T$ for $j = 1, \dots, n-1$, $y_n = (1, -\alpha_n, \dots, -\alpha_n)^T$, and $\lim_{\varepsilon \to 0} \alpha_j > 0$ for all $j = 1, \dots, n$. For $i \in \{1, \dots, n\}$, denote by $\hat{L}_i \subset \mathbb{R}^n$ the $(n-1)$-dimensional linear subspace spanned by all e_j except e_i, and let z_i be the intersection point of $\hat{L}_i$ with the line l_i. It is not hard to check that $z_i = (1, -\varepsilon, \dots, -\varepsilon, -\beta_i, -\varepsilon, \dots, -\varepsilon)^T$, $i = 1, \dots, n-1$, $z_n = (1, -\beta_n, \dots, -\beta_n)^T$ with $\beta_i = \varepsilon\frac{n-2}{n-3}$, $i = 1, \dots, n-1$, and $\beta_n = \varepsilon\frac{n-2}{n-1}$. For $n = 3$ the line l_i does not intersect L_i, $i = 1, 2$, and we set $\beta_1 = \beta_2 = \infty$ in this case.

The cross-ratios (15) are then given by

$$q_j = [0, -\varepsilon; -\alpha_j, -\varepsilon\frac{n-2}{n-3}] = \frac{\alpha_j}{(n-2)(\alpha_j - \varepsilon)}, \quad j = 1, \dots, n-1;$$

$$q_n = [0, -\varepsilon; -\alpha_n, -\varepsilon\frac{n-2}{n-1}] = \frac{-\alpha_n}{(n-2)(\alpha_n - \varepsilon)}.$$

Theorem 4.4 then yields the lower bound

$$\nu_* = \frac{2}{1 - \frac{|(\sum_{j=1}^{n-1} \frac{\alpha_j}{\alpha_j - \varepsilon}) - \frac{\alpha_n}{\alpha_n - \varepsilon} - 2(n-2)|}{\sum_{j=1}^{n} \frac{\alpha_j}{\alpha_j - \varepsilon}}}$$

on the concordance parameter of any logarithmically homogeneous self-concordant barrier on K. The relation $\lim_{\varepsilon \to 0} \nu_* = n$ completes the proof. □

Corollary 6.2. *Let $n \geq 3$. The epigraph of the $||\cdot||_\infty$-norm in $\mathbb{R}^{n-1}$,*

$$K_{n,\infty} = \{(x_0, \dots, x_{n-1})^T \,|\, x_0 \geq |x_k| \;\forall\; k = 1, \dots, n-1\} \subset \mathbb{R}^n, \tag{23}$$

cannot have a logarithmically homogeneous self-concordant barrier with concordance parameter less than $\nu_* = n$. □

An optimal barrier for the cone $K_{n,\infty}$ is given by

$$F(x_0, \dots, x_{n-1}) = -\sum_{k=1}^{n-1} \log(x_0^2 - x_k^2) + (n-2)\log x_0. \tag{24}$$

For a proof see the appendix.

Note that since every convex quadrangle is projectively equivalent to a square, the barrier (24) for $n = 3$ yields an optimal barrier also for an arbitrary regular polyhedral cone $K \subset \mathbb{R}^3$ which is generated by 4 extreme rays.

7 Further examples

7.1 Power cone

The power cone is a 3-dimensional regular convex cone defined by

$$K_p = \{(u, v, w)^T \,|\, u \geq 0,\; v \geq 0,\; u^{1/p} v^{1/q} \geq |w|\}, \tag{25}$$

where $p \in (2, \infty)$ is a parameter and $\frac{1}{p} + \frac{1}{q} = 1$. In [4] Nesterov proposed a logarithmically homogeneous self-concordant barrier with concordance parameter $\nu = 4$ for this cone. In [1] a logarithmically homogeneous self-concordant barrier with concordance parameter $\nu = 3$ was proposed and a logarithmically homogeneous function with homogeneity parameter $\nu = 3 - \frac{2}{p}$ was conjectured to be self-concordant. We shall now give a lower bound on the concordance parameter of any logarithmically homogeneous self-concordant barrier on K_p.

Consider the compact section $D = \{(u, v, w)^T \in K_p \mid u + v = 1\}$ of K_p. Introducing the variable $\rho = u - v \in [-1, 1]$, we can parameterize D by $(\rho, w)^T$. Inserting the relations $u = \frac{1+\rho}{2}$, $v = \frac{1-\rho}{2}$ into the inequality defining the cone K_p, we see that D is given by the set $\{(\rho, w)^T \mid |\rho| \leq 1,\ (1+\rho)^{1/p}(1-\rho)^{1/q} \geq 2|w|\}$.

Let now $\gamma \geq 0$ be the root of the transcendent equation

$$q(1 + \gamma^{-1/q}) = p(1 + \gamma^{-1/p}). \tag{26}$$

Choose $\rho_2, \rho_3 \in (-1, 1)$ such that $\rho_2 < \rho_3$ and $\gamma = \frac{(1-\rho_3)(1+\rho_2)}{(1+\rho_3)(1-\rho_2)}$. Then ρ_3 can be expressed as a function of ρ_2 by $\rho_3 = \frac{1+\rho_2-\gamma(1-\rho_2)}{1+\rho_2+\gamma(1-\rho_2)}$. Define further $w_k = \frac{1}{2}(1 + \rho_k)^{1/p}(1 - \rho_k)^{1/q}$, $k = 2, 3$.

Set $e_1 = (1, 0)^T$, $y_1 = (-1, 0)^T$, $e_2 = (\rho_2, w_2)^T$, $y_2 = (\rho_3, -w_3)^T$, $e_3 = (\rho_2, -w_2)^T$, $y_3 = (\rho_3, w_3)^T$. All these 6 points are located on the boundary of D, and the lines l_k through e_k, y_k, $k = 1, 2, 3$, intersect in the common point $x^* = (\frac{\rho_2 z_3 + \rho_3 z_2}{z_2 + z_3}, 0)^T$, which lies in the interior of the triangle formed by the e_k. The line l_1 intersects the line through e_2, e_3 in the point $z_1 = (\rho_2, 0)^T$, while l_3 intersects the line through e_1, e_2 in the point

$$z_3 = \frac{1}{(1 - \rho_2)(w_2 + w_3) + w_2(\rho_3 - \rho_2)} \begin{pmatrix} \rho_2 w_3(1 - \rho_2) + w_2(2\rho_3 - \rho_2\rho_3 - \rho_2) \\ w_2(w_2(1 - \rho_3) + w_3(1 - \rho_2)) \end{pmatrix}.$$

From this we readily compute the cross-ratios

$$q_1 = \frac{2w_2(\rho_3 - \rho_2)}{(1 - \rho_2)(w_2(1 + \rho_3) + w_3(1 + \rho_2))}, \tag{27}$$

$$q_2 = q_3 = \frac{w_2(1 - \rho_3) + w_3(1 - \rho_2)}{2w_3(1 - \rho_2)}, \tag{28}$$

and hence

$$\sum_{k=1}^{3} q_k - 1 = \frac{w_2(1 + \rho_3)(w_2(1 - \rho_3) + w_3(1 - \rho_2))}{w_3(1 - \rho_2)(w_2(1 + \rho_3) + w_3(1 + \rho_2))} = \frac{1 + \gamma^{1/p}}{1 + \gamma^{1/q}} > 1.$$

By (22) we then obtain

$$\nu_* = q_1 + q_2 + q_3 = 1 + \frac{1 + \gamma^{1/p}}{1 + \gamma^{1/q}}. \tag{29}$$

Corollary 7.1. *Let $K_p \subset \mathbb{R}^3$ be the power cone given by (25) with parameter $p \geq (2, +\infty)$. If F is a logarithmically homogeneous self-concordant barrier on K_p, then its concordance parameter satisfies the inequality $\nu \geq 1 + \frac{1+\gamma^{1/p}}{1+\gamma^{1/q}}$, where γ is given by (26) and $\frac{1}{p} + \frac{1}{q} = 1$.* □

In Fig. 1 the lower bound ν_* as a function of p is depicted along with the concordance parameters of the barriers proposed in [4] and [1].

7.2 Epigraph of the $||\cdot||_p$ norm

Next we consider the epigraph of the $||\cdot||_p$-norm in $\mathbb{R}^2$ for $1 \leq p \leq \infty$, i.e., the 3-dimensional cone

$$K_{3,p} = \{(x_0, x_1, x_2)^T \mid x_0 \geq (|x_1|^p + |x_2|^p)^{1/p}\}. \tag{30}$$

In [4] Nesterov proposed a method to construct a logarithmically homogeneous self-concordant barrier with concordance parameter $\nu = 2\tilde{\nu}$ for this cone if a logarithmically homogeneous self-concordant barrier for the corresponding power cone K_p is available which has concordance parameter $\tilde{\nu}$. In [1] the

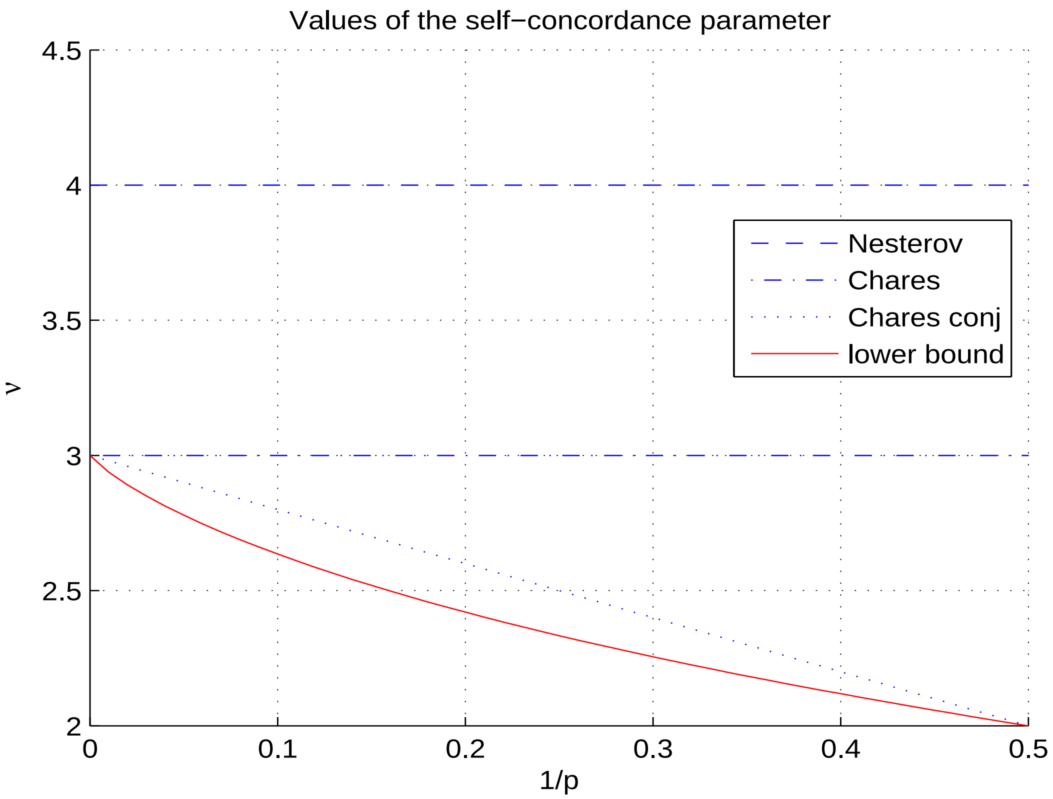

Figure 1: Concordance parameters of barriers for the power cone and lower bound

universal barrier [5, Section 2.5] for $K_{3,p}$ was computed and a concordance parameter $\nu = \frac{3p}{p+1}$ $(p \geq 2)$ for this barrier was conjectured on the basis of randomized numerical experiments. We shall now give a lower bound on the concordance parameter for any logarithmically homogeneous self-concordant barrier on $K_{3,p}$.

For $p = 2$, $K_{3,p}$ is the 3-dimensional Lorentz cone, whose optimal concordance parameter is 2. For $p = 1$ and $p = \infty$ $K_{3,p}$ is a polyhedral cone with 4 extreme rays. This case was considered in the previous section, where it was established that the optimal concordance parameter equals 3.

Let us consider the case $2 < p < \infty$. Let $\gamma \in (0,1)$ be the unique solution of the equation $(1-\frac{2}{p})\gamma^{1-1/p} + (1-\frac{1}{p})\gamma^{1-2/p} - \frac{1}{p} = 0$. Note that this definition coincides with (26). For $\delta \in (0,1)$, set

$$\rho_2 = 1-\delta, \quad \rho_3 = 1-\gamma\delta, \quad w_2 = (1-\rho_2^p)^{1/p}, \quad w_3 = (1-\rho_3^p)^{1/p}. \tag{31}$$

Consider the compact section $D = \{(\rho,w)^T \,|\, (1,\rho,w)^T \in K_{3,p}\}$ of the cone $K_{3,p}$ and let $e_1, e_2, e_3 \in \partial D$, $x^* \in D^o$ be as in the previous subsection. Then the cross-ratios q_k will be given by the same formulas (27), and

$$q_1 + q_2 + q_3 = 2 + \frac{w_2^2(1-\rho_3^2) - w_3^2(1-\rho_2^2)}{w_3(1-\rho_2)(w_3(1+\rho_2) + w_2(1+\rho_3))}. \tag{32}$$

Now note that $\frac{1-x^2}{(1-x^p)^{2/p}}$ is a strictly monotonely decreasing function for $x \in (0,1)$, hence $\frac{1-\rho_3^2}{w_3^2} < \frac{1-\rho_2^2}{w_2^2}$ and $q_1 + q_2 + q_3 < 2$. By (22) we then get the lower bound

$$\nu_* = \frac{q_1+q_2+q_3}{q_1+q_2+q_3-1} = 2 + \frac{w_3^2(1-\rho_2^2) - w_2^2(1-\rho_3^2)}{w_2(1+\rho_3)[w_3(1-\rho_2) + w_2(1-\rho_3)]} = 2 + \frac{(w_3/w_2)^2(2-\delta) - \gamma(2-\gamma\delta)}{(2-\gamma\delta)(w_3/w_2+\gamma)}.$$

We have

$$\lim_{\delta\to 0} \frac{w_3}{w_2} = \left(\lim_{\delta\to 0} \frac{1-(1-\gamma\delta)^p}{1-(1-\delta)^p}\right)^{1/p} = \left(\lim_{\delta\to 0} \frac{\gamma(1-\gamma\delta)^{p-1}}{(1-\delta)^{p-1}}\right)^{1/p} = \gamma^{1/p}, \tag{33}$$

and therefore

$$\lim_{\delta\to 0}\nu_* = 2 + \frac{\gamma^{2/p}-\gamma}{\gamma^{1/p}+\gamma} = 2 + \frac{\gamma^{1/p}-\gamma^{1-1/p}}{1+\gamma^{1-1/p}},$$

which, remarkably, coincides with (29).

Let us now consider the case $1 < p < 2$. Let $\gamma \in (0,1)$ be the unique solution of the equation $(\frac{2}{p}-1)\gamma^{1/p} + \frac{1}{p}\gamma^{2/p-1} - (1-\frac{1}{p}) = 0$. For $\delta \in (0,1)$, define ρ_2, ρ_3, w_2, w_3 by (31) and let $e_1, e_2, e_3 \in \partial D$, $x^* \in D^o$ be again as in the previous subsection. Then we again obtain (32), but $\frac{1-x^2}{(1-x^p)^{2/p}}$ is now a strictly monotonely increasing function for $x \in (0,1)$. Hence $\frac{1-\rho_3^2}{w_3^2} > \frac{1-\rho_2^2}{w_2^2}$ and $q_1+q_2+q_3 > 2$. By (22) we then get the lower bound

$$\nu_* = q_1+q_2+q_3 = 2 + \frac{(w_2/w_3)^2(2-\gamma\delta)\gamma - (2-\delta)}{(2-\delta)+w_2/w_3(2-\gamma\delta)}.$$

As in (33) we get $\lim_{\delta\to 0}\frac{w_2}{w_3} = \gamma^{-1/p}$, which yields

$$\lim_{\delta\to 0}\nu_* = 2 + \frac{\gamma^{1-2/p}-1}{1+\gamma^{-1/p}} = 2 + \frac{\gamma^{1-1/p}-\gamma^{1/p}}{1+\gamma^{1/p}}.$$

Note that this again coincides with (29), but with p and q interchanged.

Combining these results, we get the following lower bound.

Corollary 7.2. *Let $K_{3,p} \subset \mathbb{R}^3$ be the epigraph of the $||\cdot||_p$-norm given by (30) with parameter $p \in [1,\infty]$. Set $c = \min(\frac{1}{p}, 1-\frac{1}{p}) \in [0,\frac{1}{2}]$ and let $\gamma \in [0,1)$ be a solution of the equation $(1-2c)\gamma^{1-c} + (1-c)\gamma^{1-2c} - c = 0$. If F is a logarithmically homogeneous self-concordant barrier on $K_{3,p}$, then its concordance parameter satisfies the inequality $\nu \geq 1 + \frac{1+\gamma^c}{1+\gamma^{1-c}}$.* □

As already noted, as functions of p the lower bounds given in Corollaries 7.1 and 7.2 coincide.

References

[1] Robert Chares. *Cones and Interior-Point Algorithms for Structured Convex Optimization involving Powers and Exponentials.* PhD thesis, Université Catholique de Louvain, Louvain-la-Neuve, 2008.

[2] Osman Güler and Levent Tunçel. Characterization of the barrier parameter of homogeneous convex cones. *Math. Program.*, 81(1):55–76, 1998.

[3] Roland Hildebrand. Barriers on projective convex sets, 2011. To appear in AIMS Proceedings, 2011.

[4] Yuri Nesterov. Towards nonsymmetric conic optimization. Discussion paper 2006/28, CORE, Louvain-la-Neuve, 2006.

[5] Yurii Nesterov and Arkadii Nemirovskii. *Interior-point polynomial algorithms in convex programming*, volume 13 of *SIAM Stud. Appl. Math.* SIAM, Philadelphia, 1994.

A Self-concordance of function (24)

In this section we prove that the function F given by (24) is a self-concordant barrier for the cone $K_{n,\infty}$ defined in (23). Clearly F is defined and smooth on $K_{n,\infty}^o$ and satisfies (2) and (3) with homogeneity parameter $\nu = n$. It rests to show that F has a positive definite Hessian and satisfies condition (1). This will be accomplished in a number of steps.

From [3] we have the following result.

Lemma A.1. *Let $K \subset \mathbb{R}^n$ be a regular convex cone such that $e_0 = (1,0,\dots,0)^T \in (K^*)^o$. Let $F : K^o \to \mathbb{R}$ be a smooth function satisfying (3). On the interior D^o of the compact section $D = \{x = (x_0,\dots,x_{n-1})^T \in K \,|\, x_0 = 1\}$ of K, define the function $f(x_1,\dots,x_{n-1}) = \nu^{-1}F(1,x_1,\dots,x_{n-1})$. Then $F'' \succ 0$ on K^o if and only if $f'' - f'(f')^T \succ 0$ (the gradient f' is seen as a column vector) and F satisfies (1) if and only if $|f'''[h,h,h] - 6f''[h,h]f'[h] + 4(f'[h])^3| \le 2\gamma(f''[h,h] - (f'[h])^2)^{3/2}$ for all tangent vectors $h \in TD^o$, where $\gamma = \frac{\nu-2}{\sqrt{\nu-1}}$.* □

In the case $K = K_{n,\infty}$, $D \subset \mathbb{R}^{n-1}$ is the unit cube given by $|x_k| \le 1$ for all $k = 1,\dots,n-1$, $\gamma = \frac{n-2}{\sqrt{n-1}}$, and

$$f(x_1,\dots,x_{n-1}) = -\frac{1}{n}\sum_{k=1}^{n-1}\log(1-x_k^2).$$

For a tangent vector $h = (h_1,\dots,h_{n-1})^T$ we obtain

$$\begin{aligned} f'[h] &= \frac{1}{n}\sum_{k=1}^{n-1}\left(\frac{h_k}{1-x_k} - \frac{h_k}{1+x_k}\right), \\ f''[h,h] &= \frac{1}{n}\sum_{k=1}^{n-1}\left(\left(\frac{h_k}{1-x_k}\right)^2 + \left(\frac{h_k}{1+x_k}\right)^2\right), \\ f'''[h,h,h] &= \frac{2}{n}\sum_{k=1}^{n-1}\left(\left(\frac{h_k}{1-x_k}\right)^3 - \left(\frac{h_k}{1+x_k}\right)^3\right). \end{aligned}$$

Introduce the variables $a_k^- = \frac{h_k}{1-x_k}$, $a_k^+ = \frac{h_k}{1+x_k}$, $k = 1,\dots,n-1$. We then have $a_k^-a_k^+ = \frac{h_k^2}{1-x_k^2} \ge 0$, with $a_k^-a_k^+ = 0$ if and only if $h_k = a_k^- = a_k^+ = 0$. We obtain

$$f''[h,h] - (f'[h])^2 = \left[\frac{1}{n}\sum_{k=1}^{n-1}(a_k^- - a_k^+)^2 - \frac{1}{n^2}\left(\sum_{k=1}^{n-1}(a_k^- - a_k^+)\right)^2\right] + \frac{2}{n}\sum_{k=1}^{n-1}a_k^-a_k^+. \tag{34}$$

The term in brackets is a positive definite quadratic form in the variables $a_k^- - a_k^+$, while the remainder is positive for every $h \ne 0$. It follows that $f'' - f'(f')^T \succ 0$ and hence $F'' \succ 0$.

It rests to show that

$$|f'''[h,h,h] - 6f''[h,h]f'[h] + 4(f'[h])^3| \le \frac{2(n-2)}{\sqrt{n-1}}(f''[h,h] - (f'[h])^2)^{3/2}. \tag{35}$$

This will be accomplished by virtue of the following auxiliary result.

Lemma A.2. *Let $b_k^-, b_k^+ \in \mathbb{R}$, $k = 1,\dots,n-1$, $n \ge 2$. Then the value of the optimization problem*

$$\min\left[\frac{1}{n}\sum_{k=1}^{n-1}\left((b_k^-)^3 - (b_k^+)^3\right) - \frac{3}{n^2}\sum_{k,l=1}^{n-1}\left((b_k^-)^2 + (b_k^+)^2\right)(b_l^- - b_l^+) + \frac{2}{n^3}\left(\sum_{k=1}^{n-1}(b_k^- - b_k^+)\right)^3\right]:$$

$$b_k^-, b_k^+ \ge 0 \;\forall\; k = 1,\dots,n-1, \quad \frac{1}{n}\sum_{k=1}^{n-1}\left((b_k^-)^2 + (b_k^+)^2\right) - \frac{1}{n^2}\left(\sum_{k=1}^{n-1}(b_k^- - b_k^+)\right)^2 = 1$$

is bounded from below by $-\frac{n-2}{\sqrt{n-1}}$.

Proof. Since the quadratic form defined by the right-hand side of (34) is positive definite on the positive orthant, the feasible set of the problem is compact and the minimum exists. Hence there exist Lagrange multipliers ρ, $\lambda \ge 0$, $\eta_k \ge 0$, $\mu_k \ge 0$, $k = 1,\dots,n-1$, such that the Lagrange function

$$\begin{aligned} \mathcal{L} &= \lambda\left(\frac{1}{n}\sum_{k=1}^{n-1}\left((b_k^-)^3 - (b_k^+)^3\right) - \frac{3}{n^2}\sum_{k,l=1}^{n-1}\left((b_k^-)^2 + (b_k^+)^2\right)(b_l^- - b_l^+) + \frac{2}{n^3}\left(\sum_{k=1}^{n-1}(b_k^- - b_k^+)\right)^3\right) \\ &\quad -3\sum_{k=1}^{n-1}(\eta_k b_k^- + \mu_k b_k^+) - \frac{3}{2}\rho\left(\frac{1}{n}\sum_{k=1}^{n-1}\left((b_k^-)^2 + (b_k^+)^2\right) - \frac{1}{n^2}\left(\sum_{k=1}^{n-1}(b_k^- - b_k^+)\right)^2 - 1\right) \end{aligned}$$

has an unconditional minimum with respect to the variables $b_k^\pm$ at the optimal solution. We then have at the optimum

$$\begin{aligned}\frac{1}{3}\frac{\partial L}{\partial b_k^-} &= \lambda\left(\frac{(b_k^-)^2}{n} - \frac{2b_k^-}{n^2}\sum_{l=1}^{n-1}(b_l^- - b_l^+) - \frac{1}{n^2}\sum_{l=1}^{n-1}\left((b_l^-)^2 + (b_l^+)^2\right) + \frac{2}{n^3}\left(\sum_{l=1}^{n-1}(b_l^- - b_l^+)\right)^2\right)\\ &\quad -\eta_k - \rho\left(\frac{b_k^-}{n} - \frac{1}{n^2}\sum_{l=1}^{n-1}(b_l^- - b_l^+)\right) = 0,\\ \frac{1}{3}\frac{\partial L}{\partial b_k^+} &= \lambda\left(-\frac{(b_k^+)^2}{n} - \frac{2b_k^+}{n^2}\sum_{l=1}^{n-1}(b_l^- - b_l^+) + \frac{1}{n^2}\sum_{l=1}^{n-1}\left((b_l^-)^2 + (b_l^+)^2\right) - \frac{2}{n^3}\left(\sum_{l=1}^{n-1}(b_l^- - b_l^+)\right)^2\right)\\ &\quad -\mu_k - \rho\left(\frac{b_k^+}{n} + \frac{1}{n^2}\sum_{l=1}^{n-1}(b_l^- - b_l^+)\right) = 0,\end{aligned}$$

and $b_k^-\eta_k = b_k^+\mu_k = 0$ for all $k = 1,\dots,n-1$. Noting $\sum_{l=1}^{n-1}(b_l^- - b_l^+) = \sigma$ and using that $\frac{1}{n}\sum_{l=1}^{n-1}\left((b_l^-)^2 + (b_l^+)^2\right) - \frac{\sigma^2}{n^2} = 1$ we then get

$$\begin{aligned}\eta_k &= \lambda\left(\frac{(b_k^-)^2}{n} - \frac{2b_k^-}{n^2}\sigma + \frac{\sigma^2}{n^3} - \frac{1}{n}\right) - \rho\left(\frac{b_k^-}{n} - \frac{\sigma}{n^2}\right),\\ \mu_k &= \lambda\left(-\frac{(b_k^+)^2}{n} - \frac{2b_k^+}{n^2}\sigma - \frac{\sigma^2}{n^3} + \frac{1}{n}\right) - \rho\left(\frac{b_k^+}{n} + \frac{\sigma}{n^2}\right).\end{aligned}$$

From this we obtain

$$\sum_{k=1}^{n-1}(b_k^-\eta_k + b_k^+\mu_k) = \lambda\left(\frac{1}{n}\sum_{k=1}^{n-1}\left((b_k^-)^3 - (b_k^+)^3\right) - \frac{3\sigma}{n} - \frac{\sigma^3}{n^3}\right) - \rho = 0.$$

If $\lambda = 0$, then $\rho = 0$ and hence also $\eta_k = \mu_k = 0$ for all k, which contradicts the condition that the Lagrange multipliers cannot be simultaneously zero. Hence $\lambda > 0$ and we can impose the normalization condition $\lambda = 1$. It follows that $\rho = \frac{1}{n}\sum_{k=1}^{n-1}\left((b_k^-)^3 - (b_k^+)^3\right) - \frac{3\sigma}{n} - \frac{\sigma^3}{n^3}$ equals the value of the optimization problem.

Let us introduce the index sets $I_- = \{k \,|\, b_k^- > 0\}$, $I_+ = \{k \,|\, b_k^+ > 0\}$, with cardinalities m_-, m_+, respectively. We then have $\eta_k = 0$ for all $k \in I_-$, $\mu_k = 0$ for all $k \in I_+$, and hence

$$(b_k^-)^2 - \left(\frac{2\sigma}{n} + \rho\right)b_k^- + \frac{\sigma^2}{n^2} - 1 + \rho\frac{\sigma}{n} = 0 \quad \forall\, k \in I_-, \tag{36}$$

$$(b_k^+)^2 + \left(\frac{2\sigma}{n} + \rho\right)b_k^+ + \frac{\sigma^2}{n^2} - 1 + \rho\frac{\sigma}{n} = 0 \quad \forall\, k \in I_+. \tag{37}$$

Summing these equations, we get

$$n\left(\frac{\sigma^2}{n^2} + 1\right) - \sigma\left(\frac{2\sigma}{n} + \rho\right) + (m_- + m_+)\left(\frac{\sigma^2}{n^2} - 1 + \rho\frac{\sigma}{n}\right) = (-n + m_- + m_+)\left(\frac{\sigma^2}{n^2} - 1 + \rho\frac{\sigma}{n}\right) = 0.$$

We then have two possibilities.

1. $\frac{\sigma^2}{n^2} - 1 + \rho\frac{\sigma}{n} = 0$.

First note that then $\sigma \neq 0$ and hence $m_- + m_+ \geq 1$. Moreover, from (36),(37) we have that $b_k^- = \frac{2\sigma}{n} + \rho > 0$ for all $k \in I_-$, $b_k^+ = -\frac{2\sigma}{n} - \rho > 0$ for all $k \in I_+$. It follows that either I_- or I_+ is the empty set, and $m_- + m_+ \leq n - 1$. Further, by definition of σ we have $\sigma = (m_- + m_+)\left(\frac{2\sigma}{n} + \rho\right)$, which gives $\rho = \sigma\left(\frac{1}{m_- + m_+} - \frac{2}{n}\right)$. Inserting this into condition 1, we get $\frac{\sigma^2}{n}\left(\frac{1}{m_- + m_+} - \frac{1}{n}\right) = 1$, which finally gives

$$\rho^2 = \frac{n}{\frac{1}{m_- + m_+} - \frac{1}{n}}\left(\frac{1}{m_- + m_+} - \frac{2}{n}\right)^2.$$

It is not hard to check that $\max_{1\leq m\leq n-1}\frac{n}{\frac{1}{m}-\frac{1}{n}}\left(\frac{1}{m}-\frac{2}{n}\right)^2=\frac{(n-2)^2}{n-1}$, with the maximum attained at $m=1$, $m=n-1$. It follows that $\rho\geq-\frac{n-2}{\sqrt{n-1}}$.

2. $\frac{\sigma^2}{n^2}-1+\rho\frac{\sigma}{n}\neq 0$ and $m_-+m_+=n$.

Since $b_k^{\pm}=0$ for $k\notin I_{\pm}$, we have

$$\begin{aligned}\eta_k &= \frac{1}{n}\left(\frac{\sigma^2}{n^2}-1+\rho\frac{\sigma}{n}\right)\geq 0 \quad \forall\, k\notin I_-, \\ \mu_k &= -\frac{1}{n}\left(\frac{\sigma^2}{n^2}-1+\rho\frac{\sigma}{n}\right)\geq 0 \quad \forall\, k\notin I_+,\end{aligned} \tag{38}$$

and either I_- or I_+ equals the full index set $\{1,\dots,n-1\}$.

From (36),(37) we obtain

$$\begin{aligned}b_k^- &= \frac{\sigma}{n}+\frac{\rho}{2}\pm\sqrt{1+\frac{\rho^2}{4}} \quad \forall\, k\in I_-, \\ b_k^+ &= -\frac{\sigma}{n}-\frac{\rho}{2}\pm\sqrt{1+\frac{\rho^2}{4}} \quad \forall\, k\in I_+.\end{aligned}$$

Since both I_- and I_+ are nonempty and $b_k^{\pm}>0$ for $k\in I_{\pm}$, the sign at the root must be positive in both cases and $\frac{\sigma^2}{n^2}-1+\rho\frac{\sigma}{n}<0$. From (38) it then follows that $I_-=\{1,\dots,n-1\}$.

This gives

$$\sigma=(n-1)\left(\frac{\sigma}{n}+\frac{\rho}{2}+\sqrt{1+\frac{\rho^2}{4}}\right)-\left(-\frac{\sigma}{n}-\frac{\rho}{2}+\sqrt{1+\frac{\rho^2}{4}}\right)=\sigma+\frac{\rho n}{2}+(n-2)\sqrt{1+\frac{\rho^2}{4}}$$

and further $\rho=-\frac{n-2}{\sqrt{n-1}}$.

This completes the proof. □

Let us now prove (35). To this end we set

$$b_k^-=\begin{cases} a_k^-, & \text{if } a_k^-\geq 0,\\ -a_k^+, & \text{if } a_k^-<0,\end{cases} \qquad b_k^+=\begin{cases} a_k^+, & \text{if } a_k^-\geq 0,\\ -a_k^-, & \text{if } a_k^-<0,\end{cases}.$$

Then $b_k^{\pm}\geq 0$ for all $k=1,\dots,n-1$, but on the other hand $a_k^- - a_k^+ = b_k^- - b_k^+$, $(a_k^-)^2+(a_k^+)^2=(b_k^-)^2+(b_k^+)^2$, $(a_k^-)^3-(a_k^+)^3=(b_k^-)^3-(b_k^+)^3$ for all $k=1,\dots,n-1$. Application of Lemma A.2 now yields the desired result.