Math. Program., Ser. A (2008) 115:105–149
DOI 10.1007/s10107-007-0141-5

FULL LENGTH PAPER

A strong bound on the integral of the central path curvature and its relationship with the iteration-complexity of primal-dual path-following LP algorithms

Renato D. C. Monteiro · Takashi Tsuchiya

Received: 6 October 2005 / Accepted: 10 May 2007 / Published online: 16 June 2007

Abstract The main goals of this paper are to: i) relate two iteration-complexity bounds derived for the Mizuno-Todd-Ye predictor-corrector (MTY P-C) algorithm for linear programming (LP), and; ii) study the geometrical structure of the LP central path. The first iteration-complexity bound for the MTY P-C algorithm considered in this paper is expressed in terms of the integral of a certain curvature function over the traversed portion of the central path. The second iteration-complexity bound, derived recently by the authors using the notion of crossover events introduced by Vavasis and Ye, is expressed in terms of a scale-invariant condition number associated with $m \times n$ constraint matrix of the LP. In this paper, we establish a relationship between these bounds by showing that the first one can be majorized by the second one. We also establish a geometric result about the central path which gives a rigorous justification based on the curvature of the central path of a claim made by Vavasis and Ye, in view of the behavior of their layered least squares path following LP method, that the central path consists of $\mathcal{O}(n^2)$ long but straight continuous parts while the remaining curved part is relatively "short".

R. D. C. Monteiro was supported in part by NSF Grants CCR-0203113 and CCF-0430644 and ONR grant N00014-05-1-0183. T. Tsuchiya was supported in part by Japan-US Joint Research Projects of Japan Society for the Promotion of Science "Algorithms for linear programs over symmetric cones" and the Grants-in-Aid for Scientific Research (C) 15510144 of Japan Society for the Promotion of Science.

R. D. C. Monteiro (✉)
School of Industrial and Systems Engineering, Georgia Institute of Technology,
765 Ferst St. Atlanta, GA 30332, USA
e-mail: monteiro@isye.gatech.edu

T. Tsuchiya
The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-Ku,
Tokyo, 106-8569, Japan
e-mail: tsuchiya@sun312.ism.ac.jp

Keywords Interior-point algorithms · Primal-dual algorithms · Path-following · Central path · Layered least squares steps · Condition number · Polynomial complexity · Crossover events · Scale-invariance · Predictor-corrector · Affine scaling · Strongly polynomial · Linear programming · Curvature

Mathematics Subject Classification 65K05 · 68Q25 · 90C05 · 90C51 · 90C60

1 Introduction

The main goals of this paper are to: i) relate two iteration-complexity bounds associated with the Mizuno-Todd-Ye predictor-corrector algorithm [7] for linear programming (LP), one expressed in terms of an integral of a certain curvature of the central path [11, 22] and another depending on a certain scale-invariant condition number associated with the LP constraint matrix derived recently by the authors [10] using the notion of crossover events introduced in [19], and; ii) study the geometrical structure of the central path in the context of LP using the forementioned central path curvature.

Let us consider the following dual pair of linear programs:

$$\begin{aligned} &\text{minimize}_x \; c^T x \\ &\text{subject to} \quad Ax = b, \quad x \geq 0, \end{aligned} \tag{1}$$

and

$$\begin{aligned} &\text{maximize}_{(y,s)} \; b^T y \\ &\text{subject to} \qquad A^T y + s = c, \quad s \geq 0, \end{aligned} \tag{2}$$

where $A \in \Re^{m\times n}$, $c \in \Re^n$ and $b \in \Re^m$ are given, and the vectors $x, s \in \Re^n$ and $y \in \Re^m$ are the unknown variables. The MTY P-C algorithm is a path following method which approximately follows the so called central path, i.e. a well-defined trajectory $\nu > 0 \mapsto w(\nu) = (x(\nu), y(\nu), s(\nu))$ lying within the set of strictly feasible solutions of (1) and (2) and having the property that $w^* = \lim_{\nu\to 0} w(\nu)$ exists and is a primal-dual optimal solution of (1) and (2). The derivations of the two forementioned iteration-complexity bounds share the fact that they both take into consideration the geometry of the central path by exploiting the very intuitive idea that (most) path following algorithms trace the straight parts of the central path faster than the remaining curved parts. This idea was first introduced in a paper by Karmarkar [5] which analyzes the convergence behavior of an interior-point algorithm for a very specific and simple type of LP problems using the point of view of Riemannian Geometry.

The first forementioned iteration-complexity bound involves a certain integral $I(\nu_f, \nu_i) := \int_{\nu_f}^{\nu_i} (\kappa(\nu)/\nu)\, d\nu$, which was first introduced by Sonnevend et al. [11]. The term $\kappa(\nu)$ can be interpreted as a measure of the curvature of the path at $w(\nu)$ in that the smaller its value, the more straight the path is at $w(\nu)$, and hence the larger reduction of the duality gap is obtained by a step of a path following algorithm from (points nearby) $w(\nu)$. Stoer and Zhao [22] have formally given a rigorous complexity analysis of a variant of the MTY P-C algorithm based on the forementioned curvature integral (see also [21]). More specifically, they have shown that the number of iterations for this variant of the MTY P-C algorithm to approximately traverse the path from ν_i to ν_f is bounded by $\mathcal{O}(I(\nu_f, \nu_i) + \log(\nu_i/\nu_f))$. In Sect. 2 of this paper, we provide another link between the forementioned curvature integral and the iteration-complexity of the MTY P-C algorithm. More specifically, this result shows that the number of iterations of the MTY P-C algorithm for traversing the portion $\{w(\nu) : \nu \in [\nu_f, \nu_i]\}$ of the central path multiplied by the square root of the opening $\beta > 0$ of the outer neighborhood of the central path converges to the curvature integral $I(\nu_f, \nu_i)$ as β approaches 0. Hence, the integral not only gives an upper bound but also a precise asymptotic estimate on the number of iterations performed by the MTY P-C algorithm as the opening β approaches 0.

The second forementioned iteration-complexity bound is expressed in terms of a condition number associated with the constraint matrix A, in addition to the usual quantities m, n and ν_i/ν_f. A first bound of this type was obtained in the pioneering work of Vavasis and Ye [19] in the context of a path following algorithm which performs not only ordinary Newton steps but also another type of steps called layered least squares (LLS) steps (see also [6,9] for variants of the latter algorithm). Using the notion of crossover events, they showed that their method needs to perform at most $\mathcal{O}(n^2)$ LLS steps and $O(n^{3.5}\log(\bar{\chi}_A + n))$ overall steps to obtain a primal-dual optimal solution of (1) and (2), where $\bar{\chi}_A$ is a certain condition number associated with A. As opposed to the MTY P-C algorithm, Vavasis and Ye's algorithm is not scale-invariant under the change of variables $(x, y, s) = (D\tilde{x}, \tilde{y}, D^{-1}\tilde{s})$, where D is a positive diagonal matrix. Hence, when Vavasis and Ye's algorithm is applied to the scaled pair of LP problems, the number of iterations performed by it generally changes and is now bounded by $O(n^{3.5}\log(\bar{\chi}_{AD} + n))$, as AD is the coefficient matrix for the scaled pair of LP problems. Since $\bar{\chi}_{AD}$ depends on D, the latter bound does as well. The condition number $\bar{\chi}_A$ was first introduced implicitly by Dikin [1] in the study of primal affine scaling (AS) algorithms and was later studied by several researchers including Vanderbei and Lagarias [18], Todd [13], and Stewart [12]. Properties of $\bar{\chi}_A$ are studied in [2,4,16,17].

Using the notion of crossover events, LLS steps and a few other nontrivial ideas, Monteiro and Tsuchiya [10] have shown that, for the MTY P-C algorithm, the number of iterations needed to approximately traverse the central path from ν_i to ν_f is bounded by $\mathcal{O}(n^{3.5}\log(\bar{\chi}_A^* + n) + T(\nu_i/\nu_f))$, where $\bar{\chi}_A^*$ is the infimum of $\bar{\chi}_{AD}$ as D varies over the set of positive diagonal matrices and $T(t) := \min\{n^2\log(\log t), \log t\}$ for all $t > 0$. The condition number $\bar{\chi}_A^*$ is clearly scale-invariant and the ratio $\bar{\chi}_A^*/\bar{\chi}_A$, as a function of A, can be arbitrarily small (see [10]). Hence, while the iteration-complexity obtained in [10] for the MTY P-C algorithm has the extra term $T(\nu_i/\nu_f)$, its first term can be considerably smaller than the bound obtained by Vavasis and Ye. Also note that, as ν_i/ν_f grows to ∞, the iteration-complexity bound obtained in [10] is smaller than the classical iteration-complexity bound of $\mathcal{O}(n^{1/2}\log(\nu_i/\nu_f))$ established in [7] for the MTY P-C algorithm.

In view of the above discussion, the iteration-complexity of the MTY P-C algorithm can be bounded both as $\mathcal{O}(I(\nu_f, \nu_i) + \log(\nu_i/\nu_f))$ and $\mathcal{O}(n^{3.5}\log(\bar{\chi}_A^* + n) + \log(\nu_i/\nu_f))$. These two bounds raise the natural question of whether the terms

$I(\nu_f, \nu_i)$ and $n^{3.5} \log(\bar{\chi}_A^* + n)$ are related in some way. The main and most difficult result of this paper will be to show that the total curvature integral $I(0, \infty)$ defined in a improper sense as the limit of $I(\nu_f, \nu_i)$ as $\nu_f \to 0$ and $\nu_i \to \infty$ satisfies $I(0, \infty) = \mathcal{O}(n^{3.5} \log(\bar{\chi}_A^* + n))$, thereby showing that this integral is strongly bounded by a polynomial of n and $\log \bar{\chi}_A^*$.

In their paper, Vavasis and Ye made the claim that the central path consists of $\mathcal{O}(n^2)$ long but straight continuous parts while the remaining curved portion of the path has a "logarithmic length" bounded by $\mathcal{O}(n^3 \log(\bar{\chi}_A + n))$. In this claim, the $\mathcal{O}(n^2)$ straight continuous parts of the central path correspond to the parts of the path traversed by the $\mathcal{O}(n^2)$ LLS steps while the remaining curved portion of the path corresponds to the one traversed by ordinary path following steps. Hence, their claim is based solely on the way their algorithm operates without being rigorously supported by the geometric behavior of the central path itself. Our second main result of this paper, which is obtained in the process of proving the main result mentioned above, is to formally justify this claim using solely the curvature of the central path. This result essentially shows that the points of the central path with curvature larger than a certain threshold value $\bar{\kappa} > 0$ lie in $\mathcal{O}(n^2)$ intervals, each with logarithmic length bounded by $\mathcal{O}(n \log(\bar{\chi}_A^* + n) + \log(\bar{\kappa}^{-1}))$.

This paper is organized as follows. In Sect. 2, we describe the curvature of the central path, the integral based on this curvature and the MTY P-C algorithm. This section also gives another explanation of how the iteration-complexity of the MTY P-C algorithm relates to the integral and states the two main results of the paper. Finally, it also presents a simple LP instance to illustrate the main results of the paper. We remark that a reader trying to gain some insight into the results of this paper without delving with their technicalities should only read the paper up to the end of Sect. 2. Section 3 explains some basic tools used in the derivation of the main results such as crossover events and LLS steps. Section 4 derives one of the main technical results which plays a fundamental role in the derivation of the strong bound on $I(0, \infty)$. The main results are established in Sect. 5 and some concluding remarks are made in Sect. 6.

The following notation is used throughout our paper. We denote the vector of all ones by e. Its dimension is always clear from the context. The symbols $\Re^n$, $\Re^n_+$ and $\Re^n_{++}$ denote the n-dimensional Euclidean space, the nonnegative orthant of $\Re^n$ and the positive orthant of $\Re^n$, respectively. The set of all $m \times n$ matrices with real entries is denoted by $\Re^{m \times n}$. If J is a finite index set then $|J|$ denotes its cardinality, that is, the number of elements of J. For $J \subseteq \{1, \ldots, n\}$ and $w \in \Re^n$, we let w_J denote the subvector $[w_i]_{i \in J}$; moreover, if E is an $m \times n$ matrix then E_J denotes the $m \times |J|$ submatrix of E corresponding to J. For a vector $w \in \Re^n$, we let $\max(w)$ and $\min(w)$ denote the largest and the smallest component of w, respectively, $\mathrm{Diag}(w)$ denote the diagonal matrix whose i-th diagonal element is w_i for $i = 1, \ldots, n$, and w^{-1} denote the vector $[\mathrm{Diag}(w)]^{-1} e$ whenever it is well-defined. For two vectors $u, v \in \Re^n$, uv denotes their Hadamard product, i.e. the vector in $\Re^n$ whose ith component is $u_i v_i$. The Euclidean norm and the ∞-norm are denoted by $\|\cdot\|$ and $\|\cdot\|_\infty$, respectively. For a matrix E, $\mathrm{Im}(E)$ denotes the subspace generated by the columns of E and $\mathrm{Ker}(E)$ denotes the subspace orthogonal to the rows of E. The superscript ^{T} denotes transpose.

2 Relation between the MTY P-C algorithm and a curvature of the central path

In this section, we discuss the notion of a curvature of the central path and its relationship to the iteration-complexity of path-following algorithms. We then state the two main results derived in this paper, namely: i) the derivation of a strong bound on a certain (improper) integral of the forementioned curvature (Theorem 2.4); and ii) a geometric characterization of the central path in terms of its curvature (Theorem 2.5).

2.1 Curvature of the central path

In this subsection we describe the assumptions imposed on the pair of dual LP problems (1) and (2) and review the definition of the primal-dual central path and its corresponding 2-norm neighborhoods. We also introduce and motivate the notion of a certain curvature of the central path and discuss previous works which have used this curvature to derive iteration-complexity bounds of path-following algorithms which take into account the geometry of the central path. Finally, we state two major results derived in this paper, namely: 1) a strong bound on the forementioned curvature (see Theorem 2.4); and 2) a geometric characterization of the central path in terms of its curvature (see Theorem 2.5).

Given $A \in \Re^{m\times n}$, $c \in \Re^n$ and $b \in \Re^m$, consider the dual pair of linear programs (1) and (2), where $x \in \Re^n$ and $(y, s) \in \Re^m \times \Re^n$ are their respective variables. The set of strictly feasible solutions for these problems are

$$\begin{aligned} \mathcal{P}^{++} &:= \{x \in \Re^n : Ax = b,\ x > 0\}, \\ \mathcal{D}^{++} &:= \{(y, s) \in \Re^{m\times n} : A^T y + s = c,\ s > 0\}, \end{aligned}$$

respectively. Throughout the paper we make the following assumptions on the pair of problems (1) and (2).

A.1 $\mathcal{P}^{++}$ and $\mathcal{D}^{++}$ are nonempty.
A.2 The rows of A are linearly independent.

Under the above assumptions, it is well-known that for any $\nu > 0$ the system

$$xs = \nu e, \tag{3}$$
$$Ax = b, \quad x > 0, \tag{4}$$
$$A^T y + s = c, \quad s > 0, \tag{5}$$

has a unique solution $w = (x, y, s)$, which we denote by $w(\nu) = (x(\nu), y(\nu), s(\nu))$. The central path is the set consisting of all these solutions as ν varies in $(0, \infty)$. As ν converges to zero, the path $(x(\nu), y(\nu), s(\nu))$ converges to a primal-dual optimal solution (x^*, y^*, s^*) for problems (1) and (2). Given a point $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$, its duality gap and its normalized duality gap are defined as $x^T s$ and $\mu = \mu(w) := x^T s/n$, respectively.

We define the proximity measure of a point $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ with respect to the central path by

$$\phi(w) = \|xs/\mu - e\|. \tag{6}$$

Clearly, $\phi(w) = 0$ if and only if $w = w(\mu)$, or equivalently w coincides with its associated central point. The 2-norm neighborhood of the central path with opening $\beta > 0$ is defined as

$$\mathcal{N}(\beta) := \{w \in \mathcal{P}^{++} \times \mathcal{D}^{++} : \phi(w) \leq \beta\}.$$

The curvature of the central path is the function $\kappa : (0, \infty) \to [0, \infty)$ defined as

$$\kappa(\nu) := \|\nu\dot{x}(\nu)\dot{s}(\nu)\|^{1/2} \quad \forall \nu > 0, \tag{7}$$

where $(\dot{x}(\nu), \dot{y}(\nu), \dot{s}(\nu))$ denote the derivative of the central path at ν. The following result states some simple facts about this function.

Lemma 2.1 *The following statements hold*:

i) $\kappa(\nu) \leq \sqrt{n/2}$ *for all* $\nu > 0$*;*
ii) *if* $\kappa(\nu_0) = 0$ *for some* $\nu_0 > 0$ *then* $\kappa(\nu) = 0$ *for every* $\nu > 0$.

Proof Differentiating (3)–(5) with respect to ν, we conclude that for every $\nu > 0$:

$$\dot{x}(\nu)s(\nu) + x(\nu)\dot{s}(\nu) = e, \quad A\dot{x}(\nu) = 0, \quad A^T\dot{y}(\nu) + \dot{s}(\nu) = 0. \tag{8}$$

To show i), fix $\nu > 0$ and define $p := \dot{x}(\nu)s(\nu)$ and $q := x(\nu)\dot{s}(\nu)$. Using (3) and (8), we easily see that $p + q = e$ and $p^T q = 0$. These identities, the Pythagorean theorem, relation (3) and the definition of $\kappa(\nu)$ then imply that

$$2\kappa(\nu)^2 = 2\|pq\| \leq 2\|p\|\,\|q\| \leq \|p\|^2 + \|q\|^2 = \|e\|^2 = n,$$

from which i) follows. To show ii), assume that $\kappa(\nu_0) = 0$ for some $\nu_0 > 0$, and define $\tilde{w}(\nu) := w(\nu_0) + (\nu - \nu_0)\dot{w}(\nu_0)$ for every $\nu > 0$. Using the assumption that $\kappa(\nu_0) = 0$, and hence that $\dot{x}(\nu_0)\dot{s}(\nu_0) = 0$, we easily see that $\tilde{w}(\nu) = (\tilde{x}(\nu), \tilde{y}(\nu), \tilde{s}(\nu))$ satisfies system (3)–(5) for every $\nu > 0$. Since the solution of this system is unique, it follows that for every $\nu > 0$, $w(\nu) = \tilde{w}(\nu) = w(\nu_0) + (\nu - \nu_0)\dot{w}(\nu_0)$, and hence $\dot{w}(\nu) = \dot{w}(\nu_0)$. Clearly, this implies that $\kappa(\nu) = \kappa(\nu_0) = 0$ for every $\nu > 0$. □

Since the case covered by statement ii) of Lemma 2.1 is trivial, from now on we make the following assumption which discards this case.

A.3 $\kappa(\nu) > 0$ for every $\nu > 0$.

To give some intuition behind the definition (7), let $\nu > 0$ and $\beta \in (0, 1)$ be given, and consider the set

$$\mathcal{T}(\beta, \nu) := \{t \in \Re : w(\nu) - t\nu\dot{w}(\nu) \in \mathcal{N}(\beta)\}$$

It can be easily shown that

$$\mathcal{T}(\beta, \nu) = \left\{ t \in \Re : \frac{t^2}{1-t} \kappa^2(\nu) \le \beta \right\},$$

and hence that $\mathcal{T}(\beta, \nu)$ is a closed interval containing 0 whose left endpoint $t^-(\beta, \nu)$ and right endpoint $t^+(\beta, \nu)$ are given by

$$t^-(\beta, \nu) = \frac{-2}{(1 + 4\kappa(\nu)^2/\beta)^{1/2} - 1}, \quad t^+(\beta, \nu) = \frac{2}{(1 + 4\kappa(\nu)^2/\beta)^{1/2} + 1}. \tag{9}$$

Note that the point $w(\nu) - t\nu\dot{w}(\nu)$ is nothing more than the first-order estimation of the central path point $w((1-t)\nu)$. As t varies, this set of points defines a line tangent to the central path at $w(\nu)$. The scalars $t \in \Re$ in the interval $\mathcal{T}(\beta, \nu)$ correspond to points in this tangent line whose distance to the central path defined by (6) is less than or equal to β. Observe that the smaller the curvature at ν is, the larger the interval $\mathcal{T}(\beta, \nu)$ becomes, and hence the larger the region in which the first-order estimation $w(\nu) - t\nu\dot{w}(\nu)$ provides a β-approximation of the central path. Note also that the length of the interval $\mathcal{T}(\beta, \nu)$ converges to 0 as $\beta \downarrow 0$. However, for fixed $\nu > 0$ and as $\beta \downarrow 0$, the length of this interval divided by $\sqrt{\beta}$ converges to a quantity related to the $\kappa(\nu)$ as the following result states.

Proposition 2.2 *For every $\nu > 0$, we have*

$$\lim_{\beta \downarrow 0} \frac{|t^-(\beta, \nu)|}{\sqrt{\beta}} = \lim_{\beta \downarrow 0} \frac{t^+(\beta, \nu)}{\sqrt{\beta}} = \frac{1}{\kappa(\nu)}$$

Proof It follows as an immediate consequence of (9). □

Given $\nu_i \ge \nu_f \in [0, \infty]$, the following integral

$$I(\nu_f, \nu_i) := \int_{\nu_f}^{\nu_i} \frac{\kappa(\nu)}{\nu} \, d\nu \tag{10}$$

is known to play a fundamental role in the iteration-complexity analysis of primal-dual path following algorithm. This integral was first introduced by Sonnevend et al. [11] with the goal of deriving iteration-complexity bounds for path-following algorithms which take into account the geometric structure of the central path. Moreover, Zhao and Stoer [22] have shown that, when started from a suitably centered initial iterate w^i, a variant of the MTY P-C algorithm generates an iterate w^f satisfying $\mu(w^f) \le \nu_f$ in at most $\mathcal{O}\left(I(\nu_f, \nu_i) + \log(\nu_i/\nu_f)\right)$ iterations, where $\nu_i := \mu(w^i)$. Another result along the same direction is derived in the next subsection which shows that the number of iterations to reduce μ from $\mu = \nu_i$ to $\mu \le \nu_f$ asymptotically approaches $I(\nu_f, \nu_i)/\sqrt{\beta}$, as the opening β of the outer neighborhood employed by the MTY P-C algorithm tends to 0.

One of the goals of this paper is to provide a strong bound on the improper integral $I(0,\infty)$. Before presenting this bound, we introduce an important notion of a condition number associated with the constraint matrix A. Let $\mathcal{D}$ denote the set of all positive definite $n \times n$ diagonal matrices and define

$$\bar{\chi}_A := \sup \left\{ \| A^T (A\tilde{D}A^T)^{-1} A\tilde{D} \| \ : \tilde{D} \in \mathcal{D} \right\} < \infty. \tag{11}$$

The parameter $\bar{\chi}_A$ plays a fundamental role in the complexity analysis of algorithms for linear programming and layered least squares problems (see [19] and references therein). Its finiteness has been firstly established by Dikin [1]. Other authors have also given alternative derivations of the finiteness of $\bar{\chi}_A$ (see for example Stewart [12], Todd [13] and Vanderbei and Lagarias [18]).

We summarize in the next proposition a few important facts about the parameter $\bar{\chi}_A$.

Proposition 2.3 *Let $A \in \Re^{m\times n}$ with full row rank be given. Then, the following statements hold*:

a) $\bar{\chi}_{HA} = \bar{\chi}_A$ *for any nonsingular matrix* $H \in \Re^{m\times m}$;
b) $\bar{\chi}_A = \max\{\|G^{-1}A\| : G \in \mathcal{G}\}$ *where $\mathcal{G}$ denotes the set of all $m \times m$ nonsingular submatrices of A*;
c) *If the entries of A are all integers, then $\bar{\chi}_A$ is bounded by $2^{\mathcal{O}(L_A)}$, where L_A is the input bit length of A*;
d) $\bar{\chi}_A = \bar{\chi}_F$ *for any* $F \in \Re^{(n-m)\times n}$ *such that* $\mathrm{Ker}(A) = \mathrm{Im}(F^T)$.

Proof Statement a) readily follows from the definition (11). The inequality $\bar{\chi}_A \geq \max\{\|G^{-1}A\| : G \in \mathcal{G}\}$ is established in Lemma 3 of [19] while the proof of the reverse inequality is given in [13] (see also Theorem 1 of [14]). Hence, b) holds. The proof of c) can be found in Lemma 24 of [19]. A proof of d) can be found in [4]. □

We are now in a position to state the main result of this paper which gives a strong bound on $I(0,\infty)$. Its proof is quite involved and is the subject of Sects. 3, 4 and 5, with the actual proof given at the end of Sect. 5.

Theorem 2.4 *We have*

$$I(0,\infty) = \mathcal{O}(n^{3.5}\log(\bar{\chi}_A^* + n)), \tag{12}$$

where $\bar{\chi}_A^* := \inf\{\bar{\chi}_{AD} : D \in \mathcal{D}\}$.

The central path is scale-invariant in the following sense. If the change of variables $(x, y, s) = (D\tilde{x}, \tilde{y}, D^{-1}\tilde{s})$, for some $D \in \mathcal{D}$, is performed on the pair of problems (1) and (2), then the central path $(\tilde{x}(\nu), \tilde{y}(\nu), \tilde{s}(\nu))$ with respect to resulting dual pair of scaled LP problems is related to the original central path $(x(\nu), y(\nu), s(\nu))$ according to $(x(\nu), y(\nu), s(\nu)) = (D\tilde{x}(\nu), \tilde{y}(\nu), D^{-1}\tilde{s}(\nu))$, i.e. the two paths are essentially the same entity. Hence the curvature of the two central paths at a certain $\nu > 0$ coincide, that is the curvature of the central path is invariant under the above transformation.

In Sects. 3–5, we will consider only the pair of LP problems and show that $I(0,\infty) \le \mathcal{O}(n^{3.5}\log(\bar{\chi}_A + n))$ (see Theorem 5.10). It turns out that, in view of the observation in the previous paragraph, this inequality implies (12). Indeed, applying this inequality to the scaled pair of LP problems corresponding to a given $D \in \mathcal{D}$, we conclude that $I(0,\infty) \le \mathcal{O}(n^{3.5}\log(\bar{\chi}_{AD} + n))$ since $I(0,\infty)$ is invariant. Since this relation holds for every $D \in \mathcal{D}$, it follows from the definition of $\bar{\chi}_A^*$ that (12) holds.

In the process of proving the above result, we will establish the following geometric result about the central path.

Theorem 2.5 *For any constant $\bar{\kappa} \in (0, \sqrt{n/2})$, there exist $l \le n(n-1)/2$ closed intervals $I_k = [d_k, e_k]$, $k = 1, \ldots, l$, such that*:

i) $e_{k+1} \le d_k$ *for all* $k = 1, \ldots, l-1$;
ii) $\{\nu > 0 : \kappa(\nu) \ge \bar{\kappa}\} \subseteq \cup_{k=1}^{l} I_k$;
iii) $\log(e_k/d_k) = \mathcal{O}\left(n\log(\bar{\chi}_A^* + n) + n|\log\bar{\kappa}|\right)$ *for all* $k = 1, \ldots, l$.

The above result says that the set of points with large curvature, i.e. with curvature greater than or equal to $\bar{\kappa}$, is contained in at most $\mathcal{O}(n^2)$ intervals, each with length in a logarithmic scale bounded by $\mathcal{O}\left(n\log(\bar{\chi}_A^* + n) + n|\log\bar{\kappa}|\right)$. The rate of convergence of path-following algorithms along this part of the central path is slow but the result above says that this curved part is generally a small portion of the central path. On the other hand, the remaining straight part of the central path is generally much longer but path-following algorithms exhibit fast rate of convergence along it.

Finally, by reasoning exactly as in the two paragraphs following Theorem 2.4, we observe that to establish iii) of Theorem 2.5, it suffices to show that $\log(e_k/d_k) = \mathcal{O}\left(n\log(\bar{\chi}_A + n) + n|\log\bar{\kappa}|\right)$ for all $k = 1, \ldots, l$ (see Theorem 5.5).

2.2 Relation between the MTY P-C algorithm and the curvature integral

In this subsection we review the well-known MTY P-C algorithm introduced in [7] and the iteration-complexity bounds that have been obtained in [7] and [10] for it. We then establish a result showing that the number of iterations to reduce μ from $\mu = \nu_i$ to $\mu \le \nu_f$ can be asymptotically estimated as $I(\nu_f, \nu_i)/\sqrt{\beta}$ as the opening β of the central path neighborhood used by the MTY P-C algorithm approaches 0.

Each iteration of the MTY P-C algorithm consists of two steps, namely the predictor (or affine scaling) step and the corrector (or centrality) step. The search direction used by both steps at a given point $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ is the unique solution of the following linear system of equations

$$\begin{aligned} s\Delta x + x\Delta s &= \sigma\mu e - xs, \\ A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0, \end{aligned} \tag{13}$$

where $\mu = \mu(w)$ and $\sigma \in [0, 1]$ is a prespecified parameter, commonly referred to as the centrality parameter. When $\sigma = 0$, we denote the solution of (13) by Δw^{a} and refer to it as the (primal-dual) affine scaling (AS) direction at w; it is the direction used in the predictor step. When $\sigma = 1$, we denote the solution of (13) by Δw^{c} and refer to it as the centrality direction at w; it is the direction used in the corrector step.

To describe an entire iteration of the MTY P-C algorithm, suppose that a constant $\beta \in (0, 1/2]$ is given. Given a point $w = (x, y, s) \in \mathcal{N}(\beta^2)$, this algorithm generates the next point $w^+ = (x^+, y^+, s^+) \in \mathcal{N}(\beta^2)$ as follows. It first moves along the direction Δw^{a} until it hits the boundary of the larger neighborhood $\mathcal{N}(\beta)$. More specifically, it computes the point $w^{\mathrm{a}} := w + \alpha_{\mathrm{a}} \Delta w^{\mathrm{a}}$ where

$$\alpha_{\mathrm{a}} := \sup\{\alpha \in [0, 1] : w + \alpha' \Delta w^{\mathrm{a}} \in \mathcal{N}(\beta),\ \forall \alpha' \in [0, \alpha]\}. \tag{14}$$

Next, the point w^+ inside the smaller neighborhood $\mathcal{N}(\beta^2)$ is generated by taking a unit step along the centrality direction Δw^{c} at the point w^{a}, that is, $w^+ := w^{\mathrm{a}} + \Delta w^{\mathrm{c}} \in \mathcal{N}(\beta^2)$. Starting from a point $w^0 \in \mathcal{N}(\beta^2)$ and successively performing iterations as described above, the MTY P-C algorithm generates a sequence of points $\{w^k\} \subseteq \mathcal{N}(\beta^2)$ which converges to the primal-dual optimal face of problems (1) and (2).

The following iteration-complexity bound for the MTY P-C algorithm was obtained in [7] and [10].

Theorem 2.6 *Given $0 < \epsilon \le 1$ and an initial point $w^0 \in \mathcal{N}(\beta^2)$ with $\beta \in (0, 1/2]$, the MTY P-C algorithm generates an iterate $w^k \in \mathcal{N}(\beta^2)$ satisfying $\mu(w^k) \le \epsilon \mu(w^0)$ in at most*

$$\mathcal{O}\left(\min\{\sqrt{n} \log(\epsilon^{-1}),\ T(\epsilon^{-1}) + n^{3.5} \log(\bar{\chi}_A^* + n)\} \right) \tag{15}$$

iterations, where $T(t) := \min\{n^2 \log(\log t), \log t\}$ for all $t > 0$.

As mentioned in the previous subsection, the curvature integral $I(\nu_f, \nu_i)$ has been used to provide iteration-complexity bounds for primal-dual path following algorithms. Zhao and Stoer [22] have shown that, when started from a suitably centered initial iterate w^i, a variant of the MTY P-C algorithm generates an iterate w^f satisfying $\mu(w^f) \le \nu_f$ in at most $\mathcal{O}\left(I(\nu_f, \nu_i) + \log(\nu_i/\nu_f)\right)$, where $\nu_i := \mu(w^i)$.

Our next result gives another relationship between the number of iterations of the MTY P-C algorithm and the above integral.

Theorem 2.7 *Let $\beta \in (0, 1/2]$. For given $w^0 \in \mathcal{N}(\beta)$ and $0 < \nu_f < \mu(w^0)$, denote by $\#(w^0, \nu_f, \beta)$ the number of iterations of the MTY P-C algorithm with $\beta \in (0, 1/2]$ needed to reduce the duality gap from $\nu_i := \mu(w^0)$ to ν_f. Then,*

$$\lim_{\beta \to 0} \frac{I(\nu_f, \nu_i)/\sqrt{\beta}}{\#(w^0, \nu_f, \beta)} = 1. \tag{16}$$

Before giving the proof of the above result, we will first state a few easy but important technical results. Given $r \in \Re^n_{++}$ and $\nu > 0$, the system of equations

$$xs = \nu r, \quad Ax = b, \quad A^T y + s = c, \tag{17}$$

has a unique solution $w = (x, y, s) \in \Re^n_{++} \times \Re^m \times \Re^n_{++}$, which we denote by $w(\nu, r) = (x(\nu, r), y(\nu, r), s(\nu, r))$. It is well-known that the map $(\nu, r) \in \Re_{++} \times$

$\Re^n_{++} \mapsto w(\nu, r)$ is analytic. Clearly, $w(\nu, e) = w(\nu)$ for every $\nu > 0$. We will denote the partial derivative of $w(\nu, r)$ with respect to ν simply by $\dot{w}(\nu, r)$. The following result relates the affine scaling direction $\Delta w^a(w)$ at the point $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ with this partial derivative.

Lemma 2.8 *For every $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$, we have $\Delta w^a(w) = -\mu \dot{w}(\mu, r)$ where $\mu = \mu(w)$ and $r = r(w) := xs/\mu$.*

Proof Differentiating (17) with respect to ν and using the first relation in (17), we easily see that $-\mu \dot{w}(\mu, r)$ is a solution of the linear system (13) with $\sigma = 0$, and hence that $\Delta w^a(w) = -\mu \dot{w}(\mu, r)$. □

We are now ready to prove Theorem 2.7.

Proof of Theorem 2.7 Define $\psi(\nu, r) := \nu \dot{x}(\nu, r)\dot{s}(\nu, r)$ for every $\nu > 0$. Clearly, $\psi(\nu, r)$ is analytic over the set $\Re_{++} \times \Re^n_{++}$. It is easy to see that this implies that there exists a constant $C = C(\nu_f, \nu_i) > 0$ such that

$$\max \left\{ \|\psi(\nu, r) - \psi(\nu, e)\| : \nu \in [\nu_f, \nu_i] \right\} \le C \|r - e\|, \tag{18}$$

for every r lying in the closed ball $\{r \in \Re^n : \|r - e\| \le 1/2\}$. Let $w^0, \ldots, w^K$ be a (finite) sequence generated by the MTY P-C algorithm with opening β satisfying the conditions that $\mu_0 = \nu_i$ and $\mu_K \le \nu_f < \mu_{K-1}$, where $\mu_k := \mu(w^k)$ for all $k = 0, \ldots, K$. By Lemma 2.8, we have that $\Delta w^a(w^k) = -\mu_k \dot{w}(\mu_k, r^k)$, where $r^k := x^k s^k / \mu_k$, for every $k = 0, \ldots, K$. It is easy to see that the stepsize α^a_k defined in (14) at iterate w^k satisfies

$$\begin{aligned}(1 - \alpha^a_k)\beta &= \left\| (1 - \alpha^a_k)(r^k - e) + (\alpha^a_k)^2 \frac{\Delta x^a(w^k)\Delta s^a(w^k)}{\mu_k} \right\| \\ &= \left\| (1 - \alpha^a_k)(r^k - e) + (\alpha^a_k)^2 \psi(\mu_k, r^k) \right\| = \left\| q^k + (\alpha^a_k)^2 \psi(\mu_k, e) \right\|\end{aligned} \tag{19}$$

where the second last equality follows from the definition $\psi(\cdot, \cdot)$ and Lemma 2.8 and the vector q^k in the last expression is defined as

$$q^k := (1 - \alpha^a_k)(r^k - e) + (\alpha^a_k)^2 [\psi(\mu_k, r^k) - \psi(\mu_k, e)]$$

Since $\alpha^a_k \in [0, 1]$ and $\|r^k - e\| \le \beta^2 \le 1/4$, it follows from the above relation and (18) that $\|q^k\| \le (C + 1)\|r^k - e\| = \mathcal{O}(\beta^2)$. This estimate together with (7) and (19) then imply that

$$(\alpha^a_k)^2 \kappa(\mu_k)^2 = (\alpha^a_k)^2 \|\psi(\mu_k, e)\| = \mathcal{O}(\beta). \tag{20}$$

Noting that the quantity $\kappa(\nu)$ is bounded below by a positive constant over the (compact) interval $[\nu_f, \nu_i]$ in view of Assumption A.3, it follows from (20) that

$\alpha_k^a = \mathcal{O}(\beta^{1/2})$. This conclusion together with (19) again then imply that

$$(\alpha_k^a)^2 \kappa(\mu_k)^2 = (\alpha_k^a)^2 \|\psi(\mu_k, e)\| = \beta \left(1 + \mathcal{O}(\beta^{1/2})\right). \tag{21}$$

Using the identity $\mu_{k+1} = (1 - \alpha_k^a)\mu_k$, relation (21), the mean value theorem and the compactness of the interval $[\nu_f, \nu_i]$, it is easy to see that

$$\begin{aligned}
\int_{\mu_{k+1}}^{\mu_k} \frac{\kappa(\nu)}{\nu}\, d\nu &= \frac{\kappa(\mu_k)}{\mu_k}(\mu_k - \mu_{k+1}) + \mathcal{O}\left((\mu_k - \mu_{k+1})^2\right) \\
&= \alpha_k^a \kappa(\mu_k) + \mathcal{O}\left((\alpha_k^a)^2 \mu_k^2\right) = \alpha_k^a \kappa(\mu_k) + \mathcal{O}(\beta) \\
&= \beta^{1/2}\left(1 + \mathcal{O}(\beta^{1/2})\right)^{1/2} + \mathcal{O}(\beta),
\end{aligned}$$

where the second last equality follows from the fact that $\alpha_k^a = \mathcal{O}(\beta^{1/2})$ and $\mu_k \le \nu_i$ for all k. The last relation and (21) then imply

$$\begin{aligned}
\frac{I(\nu_f, \nu_i)/\sqrt{\beta}}{\#(w^0, \nu_f, \beta)} &= \frac{1}{K\beta^{1/2}} \int_{\nu_f}^{\nu_i} \frac{\kappa(\nu)}{\nu}\, d\nu \\
&= \frac{1}{K\beta^{1/2}} \left(\sum_{k=0}^{K-2} \int_{\mu_{k+1}}^{\mu_k} \frac{\kappa(\nu)}{\nu}\, d\nu + \int_{\nu_f}^{\mu_{K-1}} \frac{\kappa(\nu)}{\nu}\, d\nu \right) \\
&= \frac{1}{K\beta^{1/2}} \left((K-1)\beta^{1/2}[1+\mathcal{O}(\beta^{1/2})]^{1/2} + (K-1)\mathcal{O}(\beta) + \mathcal{O}(\beta^{1/2}) \right).
\end{aligned}$$

Since $\alpha_k^a = \mathcal{O}(\beta^{1/2})$, we have that α_k^a converges to 0, as β tends to 0, for every $k = 0, \ldots, K-1$. This observation implies that $K = K(\beta)$ must converge to ∞ as β tends to 0. Using this observation in the previous relation, we conclude that (16) holds. □

Before ending this subsection, we introduce a new *canonical* parametrization of the central path based on the function $\eta(\nu) := I(0, \nu)$. Note that this function is well-defined and strictly increasing in view of Assumption A.3 and Theorem 2.4. Hence, it has an inverse $\nu(\eta)$ whose domain is the open interval $(0, \bar{\eta})$, where $\bar{\eta} := I(0, \infty)$. Using this function, we can reparametrize the central path $w(\nu)$ in terms of η as $\tilde{w}(\eta) = w(\nu(\eta))$. This parametrization is quite natural in the sense that the iteration-complexity of the MTY P-C algorithm to trace the reparametrized central path from $\eta = \eta_i$ to $\eta = \eta_f \le \eta_i$ can be estimated as $(\eta_i - \eta_f)/\sqrt{\beta}$. It would be worth noting that the MTY P-C algorithm admits an interpretation as a predictor-corrector type numerical integration formula of an ordinary differential equation with respect to this parametrization η.

2.3 An illustrative example

In this subsection, we consider a simple LP instance to illustrate the results described in Sects. 2.1 and 2.2. The first goal of this example is to illustrate the result of Theorem 2.7 which asymptotically relates the number of iterations performed by the MTY P-C algorithm with the curvature integral. The second goal is to illustrate how the straight parts of the primal-dual central path actually look as being the curved parts of the dual-only central path, thereby showing that examination of only one of the primal or dual components of the central path might give you the wrong indication of what pieces of the (primal-dual) central path constitute its straight parts.

Consider the following simple LP instance

$$\begin{aligned} &\max\ b^T y \\ &\text{s.t.} \quad c - A^T y \geq 0, \end{aligned}$$

where

$$A = \begin{pmatrix} 0 & \frac{\sqrt{6}}{3} & -\frac{\sqrt{6}}{3} & 0 \\ 0 & \frac{\sqrt{2}}{3} & \frac{\sqrt{2}}{3} & -\frac{2\sqrt{2}}{3} \\ -1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}, \quad b = \begin{pmatrix} -10^{-9} \\ -10^{-5} \\ -1 \end{pmatrix}, \quad c = \begin{pmatrix} 0 \\ \frac{2\sqrt{6}}{3} \\ 0 \\ 0 \end{pmatrix}.$$

The feasible region of this LP instance (which we refer to as the dual problem) is a tetrahedron which is drawn together with the (dual) central path in Fig. 1. By choosing $|b_3| \gg |b_2| \gg |b_1|$ forces the dual central path to make two sharp turns as the optimal solution, i.e. the origin, is approached, namely: they occur as the facet and edge of the tetrahedron lying in the (y_1, y_2)-plane and along the y_1-axis, respectively, are approached.

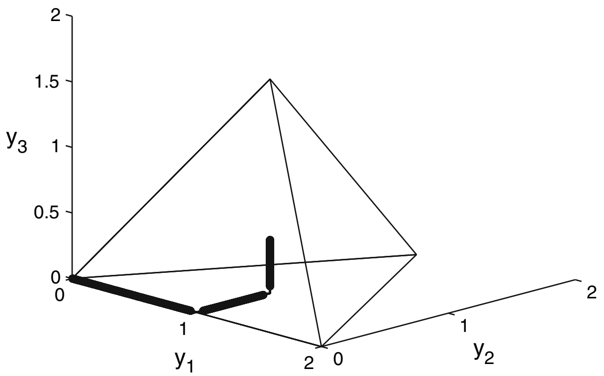

Fig. 1 Figure for the LP instance

Figure 2 plots the number of iterations performed by the MTY P-C algorithm versus the logarithm of the current duality gap for different values of the opening β, namely when $\sqrt{\beta}$ is equal to 0.0025, 0.005, 0.01 and 0.02, respectively. (The points in the graph is plotted every ten iterations of the MTY P-C algorithm.) Observe that the points corresponding to the same value of β appear to form a smooth curve and that the different "curves", except for the different scalings along the vertical axis, have the same shape. In fact, Theorem 2.7 claims that by changing the scaling along the vertical axis, i.e. by multiplying the number of iterations by $\sqrt{\beta}$, the curves converge to a well-defined smooth curve (see Fig. 3) defined by the set of points

$$\left\{ \left(\log \mu \, , \int_{\mu}^{\nu_i} \frac{\kappa(\nu)}{\nu} \, d\nu \right) : 0 < \mu \leq \nu_i \right\},$$

where ν_i denotes the duality gap at the initial iterate.

In Fig. 4, the forementioned limiting curve is plotted again. It is easy to see that the absolute value of the slope of this curve at $\log \mu$ is exactly the curvature $\kappa(\mu)$. Note that three continuous parts of this curve are drawn using circles. They correspond to the parts of the central path where the duality gap reduces slowly, i.e. the curvature is large. It is interesting to note that these three pieces of the curve correspond to the three continuous pieces (drawn in circles) of the dual-only central path plotted in Fig. 1. Note that these curved pieces actually look quite straight in the dual space.

On the other hand, in Fig. 4, the two continuous parts (drawn using normal dots) between the three continuous parts mentioned above correspond to the parts of the central path where the curvature is small, i.e., the duality gap reduces quickly. It is interesting to note that these two pieces of the curve correspond to the two continuous pieces (drawn in normal dots) of the dual-only central path plotted in Fig. 1. Note that

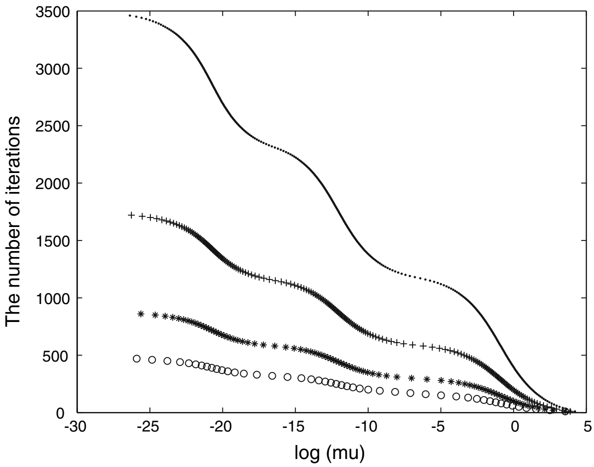

Fig. 2 $\log \mu$ versus $\#(w^0, \mu, \beta)$ with $\mu(w^0) = 10^2$ ($\cdot : \sqrt{\beta} = 0.0025$; $+ : \sqrt{\beta} = 0.005$; $* : \sqrt{\beta} = 0.01$; $\circ : \sqrt{\beta} = 0.02$)

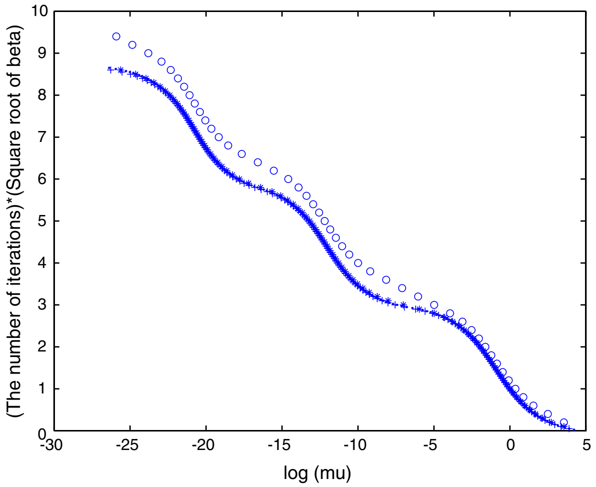

Fig. 3 $\log \mu$ versus $\sqrt{\beta} \cdot \#(w^0, \mu, \beta)$ with $\mu(w^0) = 10^2$ ($\cdot : \sqrt{\beta} = 0.0025$; $+ : \sqrt{\beta} = 0.005$; $* : \sqrt{\beta} = 0.01$; $\circ : \sqrt{\beta} = 0.02$)

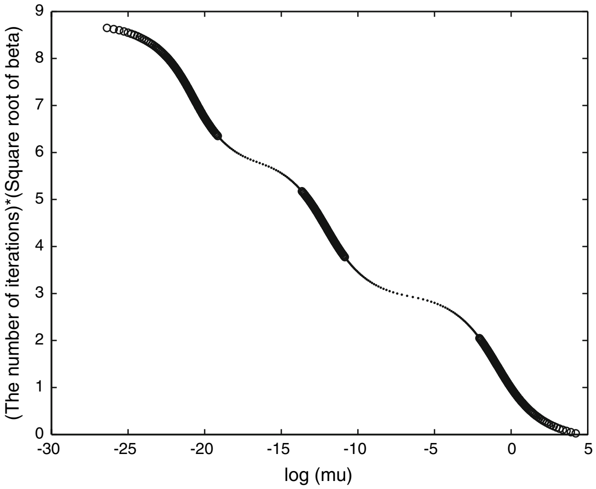

Fig. 4 log μ versus $\sqrt{\beta} \cdot \#(10^2, \mu, \beta)$ (the *solid dots* and *circles* corresponding to the ones in Fig. 1)

these pieces actually look like points making very sharp turns in the dual space in spite of the fact that the curvature is small.

Hence, we conclude that examination of the dual central path only may give us the wrong perception of what constitute the curved parts of the primal-dual central path. In this example, the two straight parts of the primal-dual central path actually correspond to the two sharp turns of the dual central path drawn in Fig. 1.

Finally, we plot in Fig. 5 the canonical parametrization $\eta(\mu) = I(0, \mu)$ versus $\log \mu$. Note that the boundedness of the integral can be observed in this figure. The limit $I(0, \infty)$ of the parametrization $\eta(\mu)$ as $\mu \to \infty$ depends on b and c, but it is strongly bounded by $\mathcal{O}(n^{3.5} \log(\bar{\chi}_A^* + n))$.

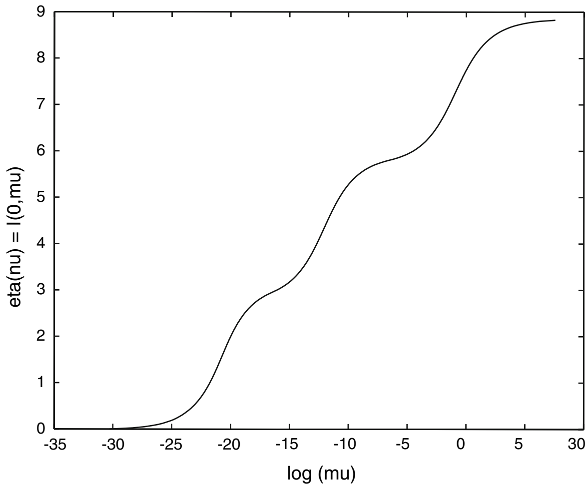

Fig. 5 $\log \mu$ versus $\eta(\mu) = I(0, \mu)$

3 Basic tools

In this section we introduce the basic tools that will be used in the proof of Theorem 2.4. The analysis heavily relies on the notion of crossover events due to Vavasis and Ye [19]. In Sect. 3.1, we give a definition of crossover event which is slightly different than the one introduced in [19] and then discuss some of its properties. In Sect. 3.2, we describe the notion of an LLS direction introduced in [19] and then state a proximity result that gives sufficient conditions under which the AS direction can be well approximated by an LLS direction. We also review from a different perspective an important result from [19], namely Lemma 17 of [19], that essentially guarantees the occurrence of crossover events. Since this result is stated in terms of the residual of an LLS step, the use of the proximity result of Sect. 3.2 allows us to obtain a similar result stated in terms of the residual of the AS direction. In Sect. 3.3, we introduce two ordered partitions of the set of variables which play an important role in our analysis.

3.1 Crossover events

In this subsection we discuss the important notion of a crossover event developed by Vavasis and Ye [19].

Definition For indices $i \neq j \in \{1, \ldots, n\}$, scalars $0 < \nu_1 < \nu_0$, and a constant $\mathcal{C} \geq 1$, a $\mathcal{C}$-*crossover event* for the pair (i, j) is said to occur on the interval $(\nu_1, \nu_0]$ if

$$\begin{aligned} &\text{there exists } \nu \in (\nu_1, \nu_0] \text{ such that } \frac{s_j(\nu)}{s_i(\nu)} = \frac{x_i(\nu)}{x_j(\nu)} \leq \mathcal{C}, \\ &\text{and, } \frac{s_j(\nu')}{s_i(\nu')} = \frac{x_i(\nu')}{x_j(\nu')} > \mathcal{C} \text{ for all } \nu' \leq \nu_1. \end{aligned} \tag{22}$$

Moreover, the interval $(\nu_1, \nu_0]$ is said to contain a $\mathcal{C}$-crossover event if (22) holds for some indices $i \neq j \in \{1, \ldots, n\}$.

Note that the notion of a crossover event is independent of any algorithm and is a property of the central path only. We have the following simple but crucial result about crossover events.

Proposition 3.1 *Let $\mathcal{C} \geq 1$ be a given constant. There can be at most $n(n-1)/2$ disjoint intervals of the form $(\nu_1, \nu_0]$ containing $\mathcal{C}$-crossover events.*

3.2 LLS directions and their relationship with the AS direction

In this subsection we describe another type of direction which plays an important role on a criterion which guarantees the occurrence of crossover events (see Lemma 3.3), namely the layered least squares (LLS) direction, which was first introduced by Vavasis and Ye in [19]. We also state a proximity result which describes how the AS direction can be well-approximated by suitable LLS directions.

For any point $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$, we define

$$\delta(w) := s^{1/2} x^{-1/2} \in \Re^n. \tag{23}$$

Moreover, for any search direction $\Delta w = (\Delta x, \Delta y, \Delta s)$ at w, we refer to the quantity

$$\begin{aligned}(Rx, Rs) &:= \left(\frac{\delta(x + \Delta x)}{\sqrt{\mu}}, \frac{\delta^{-1}(s + \Delta s)}{\sqrt{\mu}} \right) \\ &= \left(\frac{x^{1/2} s^{1/2} + \delta \Delta x}{\sqrt{\mu}}, \frac{x^{1/2} s^{1/2} + \delta^{-1} \Delta s}{\sqrt{\mu}} \right),\end{aligned} \tag{24}$$

where $\delta := \delta(w)$, to as the *residual* of Δw (at w). We will denote the residual of the affine scaling direction $\Delta w^{\mathrm{a}} = (\Delta x^{\mathrm{a}}, \Delta y^{\mathrm{a}}, \Delta s^{\mathrm{a}})$ at w as $(Rx^{\mathrm{a}}(w), Rs^{\mathrm{a}}(w))$. Note that if $(Rx^{\mathrm{a}}, Rs^{\mathrm{a}}) := (Rx^{\mathrm{a}}(w), Rs^{\mathrm{a}}(w))$ and $\delta := \delta(w)$, then

$$Rx^{\mathrm{a}} = -\frac{1}{\sqrt{\mu}} \delta^{-1} \Delta s^{\mathrm{a}}, \quad Rs^{\mathrm{a}} = -\frac{1}{\sqrt{\mu}} \delta \Delta x^{\mathrm{a}}, \tag{25}$$

and

$$Rx^{\mathrm{a}} + Rs^{\mathrm{a}} = \frac{x^{1/2} s^{1/2}}{\sqrt{\mu}}, \tag{26}$$

due to the fact that $(\Delta x^{\mathrm{a}}, \Delta y^{\mathrm{a}}, \Delta s^{\mathrm{a}})$ satisfies the first equation in (13) with $\sigma = 0$. The following quantity plays an important role in our analysis:

$$\varepsilon_{\infty}^{\mathrm{a}}(w) := \max_i \left\{ \min \left\{ \left| Rx_i^{\mathrm{a}}(w) \right|, \left| Rs_i^{\mathrm{a}}(w) \right| \right\} \right\}. \tag{27}$$

We will now give the definition of the LLS direction. Let $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ and a partition $J = (J_1, \ldots, J_p)$ of the index set $\{1, \ldots, n\}$ be given and let $\delta := \delta(w)$. The primal LLS direction $\Delta x^{\rm ll} = (\Delta x^{\rm ll}_{J_1}, \ldots, \Delta x^{\rm ll}_{J_p})$ at w with respect to J is defined recursively according to the order $\Delta x^{\rm ll}_{J_p}, \ldots, \Delta x^{\rm ll}_{J_1}$ as follows. Assume that the components $\Delta x^{\rm ll}_{J_p}, \ldots, \Delta x^{\rm ll}_{J_{k+1}}$ have been determined. Let $\Pi_{J_k} : \Re^n \to \Re^{J_k}$ denote the projection map defined as $\Pi_{J_k}(u) = u_{J_k}$ for all $u \in \Re^n$. Then $\Delta x^{\rm ll}_{J_k} := \Pi_{J_k}(L^x_k)$ where L^x_k is given by

$$\begin{aligned} L^x_k &:= \mathrm{Argmin}_{\Delta x \in \Re^n} \left\{ \|\delta_{J_k}(x_{J_k} + \Delta x_{J_k})\|^2 : \Delta x \in L^x_{k+1} \right\} \\ &= \mathrm{Argmin}_{\Delta x \in \Re^n} \left\{ \|\delta_{J_k}(x_{J_k} + \Delta x_{J_k})\|^2 : \Delta x \in \mathrm{Ker}(A),\ \Delta x_{J_i} \right. \\ &\qquad\qquad \left. = \Delta x^{\rm ll}_{J_i} \ \text{ for all } i = k+1, \ldots, p \right\}, \end{aligned} \tag{28}$$

with the convention that $L^x_{p+1} := \mathrm{Ker}(A)$. The slack component $\Delta s^{\rm ll} = (\Delta s^{\rm ll}_{J_1}, \ldots, \Delta s^{\rm ll}_{J_p})$ of the dual LLS direction $(\Delta y^{\rm ll}, \Delta s^{\rm ll})$ at w with respect to J is defined recursively as follows. Assume that the components $\Delta s^{\rm ll}_{J_1}, \ldots, \Delta s^{\rm ll}_{J_{k-1}}$ have been determined. Then $\Delta s^{\rm ll}_{J_k} := \Pi_{J_k}(L^s_k)$ where L^s_k is given by

$$\begin{aligned} L^s_k &:= \mathrm{Argmin}_{\Delta s \in \Re^n} \left\{ \|\delta^{-1}_{J_k}(s_{J_k} + \Delta s_{J_k})\|^2 : \Delta s \in L^s_{k-1} \right\} \\ &= \mathrm{Argmin}_{\Delta s \in \Re^n} \left\{ \|\delta^{-1}_{J_k}(s_{J_k} + \Delta s_{J_k})\|^2 : \Delta s \in \mathrm{Im}(A^T),\ \Delta s_{J_i} = \Delta s^{\rm ll}_{J_i} \right. \\ &\qquad\qquad \left. \text{for all } i = 1, \ldots, k-1 \right\}, \end{aligned} \tag{29}$$

with the convention that $L^s_0 := \mathrm{Im}(A^T)$. Finally, once $\Delta s^{\rm ll}$ has been determined, the component $\Delta y^{\rm ll}$ is determined from the relation $A^T \Delta y^{\rm ll} + \Delta s^{\rm ll} = 0$.

It is easy to verify that the AS direction is a special LLS direction, namely the one with respect to the only partition in which $p = 1$ (see Sect. 5 of [19]). Clearly, the LLS direction at a given $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ depends on the partition $J = (J_1, \ldots, J_p)$ used.

We next introduce some more definitions and notation which will be used throughout the paper. Given a partition $J = (J_1, \ldots, J_p)$ of $\{1, \ldots, n\}$ and a point $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$, we define

$$\begin{aligned} \mathrm{gap}\,(w, J) &:= \min \left\{ \frac{\delta_j(w)}{\delta_i(w)} : i \in J_k \text{ and } j \in J_l \text{ for some } 1 \le k < l \le p \right\}, \\ &= \min_{1 \le k \le p-1} \frac{\min(\delta_{J_{k+1}}(w), \ldots, \delta_{J_p}(w))}{\max(\delta_{J_1}(w), \ldots, \delta_{J_k}(w))}. \end{aligned} \tag{30}$$

(By convention, we let $\mathrm{gap}\,(w, J) := \infty$ if $p = 1$.) Moreover, the partition J is said to be *ordered* at w if $\mathrm{gap}\,(w, J) \ge 1$, or equivalently, if for any $1 \le k < l \le p$, we have $\delta_i(w) \le \delta_j(w)$ for any $i \in J_k$ and $j \in J_l$. Note that the order of the indices within a layer J_k of an ordered partition J at w is irrelevant, which is the reason for viewing J_k as a set instead of an ordered tuple.

Note that when J is ordered at w, we have:

$$\begin{aligned} \operatorname{gap}(w, J) &= \min \left\{ \frac{\delta_j(w)}{\delta_i(w)} : i \in J_k \text{ and } j \in J_{k+1} \text{ for some } 1 \le k < p \right\} \\ &= \min_{1 \le k \le p-1} \frac{\min(\delta_{J_{k+1}}(w))}{\max(\delta_{J_k}(w))}. \end{aligned} \tag{31}$$

Note also that for the special where $w = w(\nu)$, we have

$$\operatorname{gap}(w(\nu), J) = \min \left\{ \frac{x_i(\nu)}{x_j(\nu)} : i \in J_k \text{ and } j \in J_l \text{ for some } 1 \le k < l \le p \right\}. \tag{32}$$

The following result, whose proof can be found in [9], gives an upper bound on the distance between the AS direction at a given $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ and a LLS direction at the same point in terms of n, $\bar{\chi}_A$ and $\operatorname{gap}(w, J)$, where J is the partition corresponding to the LLS direction. It essentially states that the larger $\operatorname{gap}(w, J)$ is, the closer these two directions will be to one another.

Proposition 3.2 *Let $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ and a partition $J = (J_1, \ldots, J_p)$ of $\{1, \ldots, n\}$ be given, and let $(Rx^{\mathrm{a}}, Rs^{\mathrm{a}})$ and $(Rx^{\mathrm{ll}}, Rs^{\mathrm{ll}})$ denote the residuals of the AS direction at w and of the LLS direction at w with respect to J, respectively. If* $\operatorname{gap}(w, J) \ge 4\, p\, \bar{\chi}_A$, *then*

$$\max \left\{ \left\| Rx^{\mathrm{a}} - Rx^{\mathrm{ll}} \right\|_\infty, \left\| Rs^{\mathrm{a}} - Rs^{\mathrm{ll}} \right\|_\infty \right\} \le \frac{12\sqrt{n}\, \bar{\chi}_A}{\operatorname{gap}(w, J)}.$$

In view of the above result, the AS direction can be well approximated by LLS directions with respect to partitions J which have large gaps. Obviously, the LLS direction with $p = 1$, which is the AS direction, provides the perfect approximation to the AS direction itself. However, this kind of trivial approximation is not useful for us due to the need of keeping the “spread” of some layers J_k under control (see Lemma 3.3 below). For a point $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ and a partition $J = (J_1, \ldots, J_p)$ of $\{1, \ldots, n\}$, the spread of the layer J_k at w, denoted by $\operatorname{spr}(w, J_k)$, is defined as

$$\operatorname{spr}(w, J_k) := \frac{\max(\delta_{J_k}(w))}{\min(\delta_{J_k}(w))}, \quad \forall k = 1, \ldots, p.$$

We will now state a result due to Vavasis and Ye which provides a sufficient condition for the occurrence of a crossover event on a interval $(\nu', \nu]$. The result is an immediate consequence of Lemma 17 of [19] and the proof of a slightly more general version of this result can be found in Lemma 3.4 of [9].

Lemma 3.3 *Let $J = (J_1, \ldots, J_p)$ be an ordered partition at $w(\nu)$ for some $\nu > 0$. Let $(Rx^{\mathrm{ll}}, Rs^{\mathrm{ll}})$ denote the residual of the LLS direction $(\Delta x^{\mathrm{ll}}, \Delta y^{\mathrm{ll}}, \Delta s^{\mathrm{ll}})$ at $w(\nu)$ with respect to J. Then, for any index $q \in \{1, \ldots, p\}$, any constant $\mathcal{C}_q \ge$* $\operatorname{spr}(w(\nu), J_q)$*,* *and any* $\nu' \in (0, \nu)$ *such that*

$$\frac{\nu'}{\nu} \le \frac{\|Rx^{\rm ll}_{J_q}\|_\infty \|Rs^{\rm ll}_{J_q}\|_\infty}{n^3 \mathcal{C}_q^2 \bar{\chi}_A^2}, \tag{33}$$

the interval $(\nu', \nu]$ *contains a* $\mathcal{C}_q$*-crossover event.*

3.3 Two important ordered partitions

In this subsection we describe two ordered partitions which play an important role in our analysis.

The first ordered partition is due to Vavasis and Ye [19]. Given a point $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ and a parameter $\bar{g} \ge 1$, this partition, which we refer to as the *VY* $\bar{g}$*-partition* at w, is defined as follows. Let $(i_1, \ldots, i_n)$ be an ordering of $\{1, \ldots, n\}$ such that $\delta_{i_1} \le \ldots \le \delta_{i_n}$, where $\delta = \delta(w)$. For $k = 2, \ldots, n$, let $r_k := \delta_{i_k}/\delta_{i_{k-1}}$ and define $r_1 := \infty$. Let $1 = k_1 < \ldots < k_p$ be all the indices k such that $r_k > \bar{g}$. The VY $\bar{g}$-partition J is then defined as $J = (J_1, \ldots, J_p)$, where $J_q := \{i_{k_q}, i_{k_q+1}, \ldots, i_{k_{q+1}-1}\}$ for all $q = 1, \ldots, p$. (Here, by convention, $k_{p+1} := n + 1$.) More generally, given a subset $I \subseteq \{1, \ldots, n\}$, we can similarly define the *VY* $\bar{g}$*-partition* of I at w by taking an ordering $(i_1, \ldots, i_m)$ of I satisfying $\delta_{i_1} \le \cdots \le \delta_{i_m}$ where $m = |I|$, defining the ratios $r_1, \cdots, r_m$ as above, and proceeding exactly as in the construction above to obtain an ordered partition $J = (J_1, \ldots, J_p)$ of I.

It is easy to see that the following result holds for the partition J described in the previous paragraph (see Sect. 5 of [19]).

Proposition 3.4 *Given a subset* $I \subseteq \{1, \ldots, n\}$*, a point* $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ *and a constant* $\bar{g} \ge 1$*, the VY* $\bar{g}$*-partition* $J = (J_1, \ldots, J_p)$ *of* I *at* w *satisfies* $\mathrm{gap}\,(w, J) > \bar{g}$ *and* $\mathrm{spr}(J_q) \le \bar{g}^{|J_q|} \le \bar{g}^n$ *for all* $q = 1, \ldots, p$.

The second ordered partition at a given point $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ which is heavily used in our analysis is obtained as follows. First, we compute the AS-bipartition $(B, N) = (B(w), N(w))$ at w as

$$B(w) := \{i : |Rs^{\rm a}_i(w)| \le |Rx^{\rm a}_i(w)|\}, \quad N(w) := \{i : |Rs^{\rm a}_i(w)| > |Rx^{\rm a}_i(w)|\}. \tag{34}$$

Next, an order $(i_1, \ldots, i_n)$ of the index variables is chosen so that $\delta_{i_1} \le \ldots \le \delta_{i_n}$. Then, the first block of consecutive indices in the n-tuple $(i_1, \ldots, i_n)$ lying in the same set B or N are placed in the first layer $\mathcal{J}_1$, the next block of consecutive indices lying in the other set is placed in $\mathcal{J}_2$, and so on. As an example assume that $(i_1, i_2, i_3, i_4, i_5, i_6, i_7) \in B \times B \times N \times B \times B \times N \times N$. In this case, we have $\mathcal{J}_1 = \{i_1, i_2\}$, $\mathcal{J}_2 = \{i_3\}$, $\mathcal{J}_3 = \{i_4, i_5\}$ and $\mathcal{J}_4 = \{i_6, i_7\}$. A partition obtained according to the above construction is clearly ordered at w. We refer to it as an *ordered AS–partition* at w, and denote it by $\mathcal{J}(w)$. Finally, for a point $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$, we define

$$\mathrm{gap}\,(w) = \mathrm{gap}\,(w, \mathcal{J}(w)). \tag{35}$$

Clearly, $\mathrm{gap}\,(w) \ge 1$ for all $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$.

Another equivalent way of defining an ordered AS–partition is as a partition $\mathcal{J} = (\mathcal{J}_1, \ldots, \mathcal{J}_l)$ of $\{1, \ldots, n\}$ which is ordered at w (i.e., $\mathrm{gap}\,(w, \mathcal{J}) \geq 1$) and has the property that $\mathcal{J}_k \subseteq B(w)$ (resp., $\mathcal{J}_k \subseteq N(w)$) implies $\mathcal{J}_{k+1} \subseteq N(w)$ (resp., $\mathcal{J}_{k+1} \subseteq B(w)$), for every $k = 1, \ldots, l-1$.

Note that an ordered AS–partition at w is not uniquely determined since there can be more than one n-tuple $(i_1, \ldots, i_n)$ satisfying $\delta_{i_1} \leq \ldots \leq \delta_{i_n}$. This situation happens exactly when there are two or more indices i with the same value for δ_i. It can be easily seen that there exists a unique ordered AS–partition at w if and only if there do not exist $i \in B(w)$ and $j \in N(w)$ such that $\delta_i = \delta_j$, or equivalently $\mathrm{gap}\,(w) = 1$.

The following result will be constantly invoked throughout our presentation and follows as an immediate consequence of the second definition of an ordered AS–partition.

Proposition 3.5 *Let $w, \hat{w}$ be points in $\mathcal{P}^{++} \times \mathcal{D}^{++}$ satisfying $(B(\hat{w}), N(\hat{w})) = (B(w), N(w))$ and let $\mathcal{J} = (\mathcal{J}_1, \ldots, \mathcal{J}_l)$ be an ordered AS–partition at w. If* $\mathrm{gap}\,(\hat{w}, \mathcal{J}) \geq 1$*, then $\mathcal{J}$ is also an ordered AS-partition at $\hat{w}$, and hence* $\mathrm{gap}\,(\hat{w}) = \mathrm{gap}\,(\hat{w}, \mathcal{J})$.

For a point $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$, note that (27) and (34) imply that

$$\varepsilon_\infty^{\mathrm{a}}(w) = \max \left\{ \|Rx_N^{\mathrm{a}}(w)\|_\infty,\ \|Rs_B^{\mathrm{a}}(w)\|_\infty \right\}, \tag{36}$$

where $B := B(w)$ and $N := N(w)$.

Finally, throughout this paper, if $f(\cdot)$ is a function defined in $\mathcal{P}_{++} \times \mathcal{D}_{++}$, we will sometimes denote the value $f(w(\nu))$ simply by $f(\nu)$. For example, $B(\nu)$, $N(\nu)$, $\mathcal{J}(\nu)$, $Rx^{\mathrm{a}}(\nu)$, $Rs^{\mathrm{a}}(\nu)$, $\mathrm{gap}\,(\nu)$ and $\varepsilon_\infty^{\mathrm{a}}(\nu)$ will denote the quantities $B(w(\nu))$, $N(w(\nu))$, $\mathcal{J}(w(\nu))$, $Rx^{\mathrm{a}}(w(\nu))$, $Rs^{\mathrm{a}}(w(\nu))$, $\mathrm{gap}\,(w(\nu))$ and $\varepsilon_\infty^{\mathrm{a}}(w(\nu))$, respectively.

4 Technical results

In this section, we prove a technical result, namely Lemma 4.11, which plays a fundamental role in the proof of Theorem 2.4. This result gives an upper bound on the integral $I(\nu_1, \nu_0)$ in terms of the quantities $\varepsilon_\infty^{\mathrm{a}}(\nu_0)$, $\mathrm{gap}\,(\nu_0)$ and the ratio ν_0/ν_1, whenever the latter is not too small. With the exception of Lemmas 4.1 and 4.11, this section may be skipped on a first reading without any loss of continuity.

Lemma 4.1 *Let $\nu > 0$ be given and let (B, N) denote the AS-partition at $w(\nu)$. Consider the vectors $u, v \in \Re^n$ defined as*

$$u_i = \begin{cases} Rs_i^{\mathrm{a}}(\nu), & \text{if } i \in B; \\ Rx_i^{\mathrm{a}}(\nu), & \text{if } i \in N, \end{cases} \qquad v_i = \begin{cases} Rx_i^{\mathrm{a}}(\nu), & \text{if } i \in B; \\ Rs_i^{\mathrm{a}}(\nu), & \text{if } i \in N, \end{cases}$$

for every $i = 1, \ldots, n$. Then, the following statements hold:

i) $v_i \geq \max\{1/2, |u_i|\}$ *for all $i = 1, \ldots, n$ (and hence, $\|v\| \geq \sqrt{n}/2$)*;
ii) $\varepsilon_\infty^{\mathrm{a}}(\nu) = \|u\|_\infty \leq \|u\| \leq \|v\| \leq \sqrt{n}$;
iii) $\varepsilon_\infty^{\mathrm{a}}(\nu)/2 \leq \kappa(\nu)^2 \leq \sqrt{n}\,\varepsilon_\infty^{\mathrm{a}}(\nu)$.

Proof By the definition of the vectors u and v and the partition (B, N), we have that $|u_i| = \min\{|Rx_i^{\mathrm{a}}(\nu)|, |Rs_i^{\mathrm{a}}(\nu)|\} \le \max\{|Rx_i^{\mathrm{a}}(\nu)|, |Rs_i^{\mathrm{a}}(\nu)|\} = |v_i|$ for all $i = 1, \ldots, n$. Since $Rx^{\mathrm{a}}(\nu)$ and $Rs^{\mathrm{a}}(\nu)$ are orthogonal and $Rx^{\mathrm{a}}(\nu) + Rs^{\mathrm{a}}(\nu) = e$, we have that

$$u + v = e, \quad u^T v = 0. \tag{37}$$

If $v_i < 1/2$ for some i then, we would have $u_i = 1 - v_i > 1/2 > v_i$, which is not possible due to the fact that $|u_i| \le |v_i|$. Hence, i) follows. The equality and the first two inequalities of ii) are immediate. The last inequality of ii) follows from (37). To show iii), first note that

$$\kappa(\nu)^2 = \|\nu\dot{x}(\nu)\dot{s}(\nu)\| = \frac{\|\Delta x(\nu)\Delta s(\nu)\|}{\nu} = \|Rx^{\mathrm{a}}(\nu)Rs^{\mathrm{a}}(\nu)\| = \|uv\|.$$

Statement iii) now follows by noting that i) and ii) imply that

$$\frac{\varepsilon_\infty^{\mathrm{a}}(\nu)}{2} \le \frac{\|u\|_\infty}{2} \le \frac{\|u\|}{2} \le \min(v)\|u\| \le \|uv\| \le \|u\|_\infty\|v\| \le \varepsilon_\infty^{\mathrm{a}}(\nu)\sqrt{n}.$$

□

Lemma 4.1(iii) shows that the integrand of the integral $I(\nu_1, \nu_0)$ is majorized by the term $\sqrt{n}\,\varepsilon_\infty^{\mathrm{a}}(\nu)$. Our main task now will be to majorize this quantity by an expression involving $\varepsilon_\infty^{\mathrm{a}}(\nu_0)$, gap $(\nu_0)^{-1}$ and the ratio ν_0/ν. Since $\varepsilon_\infty^{\mathrm{a}}(\nu)$ is given by (36) and the quantities $Rx^{\mathrm{a}}(\nu)$ and $Rs^{\mathrm{a}}(\nu)$ are projections onto diagonally scaled subspaces, it is important to understand how the sizes of these projections vary as the diagonal scalings defining the subspaces change. The next result addresses this issue. For two vectors $u, v \in \Re^k_{++}$, define

$$\Delta(v; u) := \min\left\{\left\|\alpha\frac{v}{u} - e\right\|_\infty : \alpha > 0\right\}.$$

Lemma 4.2 *For i*$= 1, 2$*, let* $h_i \in \Re^n$ *and* $d_i \in \Re^n_{++}$ *be given and define*

$$\begin{aligned} p_i :&= \operatorname{argmin}_{p\in\Re^n}\{\|h_i - p\| : D_i p \in L\} \\ &= \operatorname{argmin}_{p\in\Re^n}\{\|p\| : D_i^{-1}(h_i - p) \in L^\perp\}, \end{aligned} \tag{38}$$

where $D_i := \mathrm{Diag}(d_i)$*,* L *is a given subspace of* $\Re^n$ *and* $L^\perp$ *denotes its orthogonal complement. Then,*

$$\|p_2\| \le \|h_2 - h_1\| + \Delta_{21}\|h_1\| + (1 + \Delta_{21})\|p_1\|, \tag{39}$$

$$\|h_2 - p_2\| \le \|h_2 - h_1\| + \Delta_{12}\|h_1\| + (1 + \Delta_{12})\|h_1 - p_1\|, \tag{40}$$

where $\Delta_{12} := \Delta(d_1; d_2)$ *and* $\Delta_{21} := \Delta(d_2; d_1)$. *In particular, if* $h_1 = h_2 = e$, *then*

$$\|p_2\| \le \Delta_{21}\sqrt{n} + (1 + \Delta_{21})\|p_1\|, \tag{41}$$
$$\|h_2 - p_2\| \le \Delta_{12}\sqrt{n} + (1 + \Delta_{12})\|h_1 - p_1\|. \tag{42}$$

Proof Let $\hat{\alpha} > 0$ be such that $\Delta_{21} = \|\hat{\alpha} d_2 d_1^{-1} - e\|_\infty$. Define $\hat{p}_2 := h_2 - \hat{\alpha} D_2 D_1^{-1}(h_1 - p_1)$ and note that $D_2^{-1}(h_2 - \hat{p}_2) = \hat{\alpha} D_1^{-1}(h_1 - p_1) \in L^\perp$, and hence that $\hat{p}_2$ is feasible for the second problem in (38) with $i = 2$. This observation and the fact that $\hat{\alpha}\|d_2 d_1^{-1}\|_\infty \le 1 + \Delta_{21}$ then imply

$$\begin{aligned}
\|p_2\| \le \|\hat{p}_2\| &\le \|h_2 - \hat{\alpha} D_2 D_1^{-1}(h_1 - p_1)\| \\
&\le \|h_2 - h_1 + (I - \hat{\alpha} D_2 D_1^{-1})h_1 + \hat{\alpha} D_2 D_1^{-1} p_1\| \\
&\le \|h_2 - h_1\| + \|e - \hat{\alpha} d_2 d_1^{-1}\|_\infty \|h_1\| + \|\hat{\alpha} d_2 d_1^{-1}\|_\infty \|p_1\| \\
&\le \|h_2 - h_1\| + \Delta_{21}\|h_1\| + (1 + \Delta_{21})\|p_1\|,
\end{aligned}$$

which shows (39). The proof of (40) is based on similar arguments. The last conclusion of the lemma follows as an immediate consequence of (39) and (40) and the assumption that $h_1 = h_2 = e$. □

Given $A \in \Re^{m\times n}$, $h \in \Re^m$ and $z \in \Re^n_{++}$, consider the projection $p^0 = p(z; h, A) \in \Re^n$ given by

$$p^0 := \operatorname{argmin}_{p\in\Re^n}\{\|h - p\|^2 : AZp = 0\}, \tag{43}$$

where $Z := \operatorname{Diag}(z)$. It can be easily shown that for every $\nu > 0$:

$$Rx^{\mathrm{a}}(\nu) = e - p(\nu), \quad Rs^{\mathrm{a}}(\nu) = p(\nu), \tag{44}$$

where $p(\nu) := p(x(\nu); e, A)$. Using Lemma 4.2, it is possible to relate the size of $p(\nu)$ (resp., $e - p(\nu)$) with that of $p(\nu_0)$ (resp., $e - p(\nu_0)$), but the relation will depend on the quantity $\Delta(x(\nu), x(\nu_0))$, which can be very large due to the substantial variations that the different components of $x(\nu)/x(\nu_0)$ undergo as ν changes. However, the two similar quantities $\Delta(x_B(\nu), x_B(\nu_0))$ and $\Delta(x_N(\nu), x_N(\nu_0))$ individually can be shown to be small enough for our purposes (see Lemma 4.8). In order to take advantage of this observation, it is necessary to work with a set of projections approximating the projection $p(z; h, A)$, where each projection is over a subspace defined by a subset of variables from either x_B or x_N.

For a given scaling vector $z \in \Re^n_{++}$ and partition J, the next lemma, whose proof is given in the Appendix, shows that if gap (z, J) is large then the projection matrix onto $\operatorname{Ker}(A \operatorname{Diag}(z))$ can be well approximated by a block diagonal matrix where each block is a projection matrix associated with a layer of J.

Lemma 4.3 *Let* $A \in \Re^{m\times n}$, $h \in \Re^m$, $z \in \Re^n_{++}$ *and a partition* $J = (J_1, \ldots, J_l)$ *of* $\{1, \ldots, n\}$ *be given. Define* $p^0 \in \Re^n$ *as* $p^0 := p(z; h, A)$ *and* $\tilde{p}^0 \in \Re^n$ *as*

$$\tilde{p}^0_{J_k} := \operatorname{argmin}_{\tilde{p}_{J_k}\in\Re^{J_k}}\{\,\|\tilde{p}_{J_k} - h_{J_k}\|^2 : A_{J_k} Z_{J_k} \tilde{p}_{J_k} \in \operatorname{Im}(A_{\bar{J}_k})\,\}, \tag{45}$$

for every $k = 1, \ldots, l$, *where* $\bar{J}_k := J_{k+1} \cup \ldots \cup J_l$. *Then, for every* $k = 1, \ldots, l$, *we have*

$$\|p^0_{J_k} - \tilde{p}^0_{J_k}\| \le K(3 + 2K)\|h\|, \tag{46}$$

where $K = K(z, J; A) := \bar{\chi}_A/\mathrm{gap}\,(z, J)$.

We are ready to establish a preliminary bound on $\varepsilon^{\mathrm{a}}_\infty(\nu)$ based on the previous two lemmas.

For the proof of the next result, recall that $p(\nu) = p(x(\nu); e, A)$ and let $\tilde{p}(\nu)$ denote the optimal solution of (45) with $h = e$ and $z = x(\nu)$, for every $\nu > 0$.

Lemma 4.4 *Let scalars* $0 < \nu < \nu_0$ *be given and let* $\mathcal{J}^0 = (\mathcal{J}^0_1, \ldots, \mathcal{J}^0_l)$ *denote the ordered AS-partition at* $w(\nu_0)$. *Then,*

$$\varepsilon^{\mathrm{a}}_\infty(\nu) \le \sqrt{n}\left\{\psi + (1 + \psi)\varepsilon^{\mathrm{a}}_\infty(\nu_0) + (2 + \psi)K(3 + 2K)\right\}$$

where

$$\psi = \psi(\nu, \nu_0) := \max\{\Delta(x_B(\nu); x_B(\nu_0))\,,\ \Delta(s_N(\nu); s_N(\nu_0))\}, \tag{47}$$

$$K = K(\nu, \nu_0) := \frac{\bar{\chi}_A}{\min\{\mathrm{gap}\,(w(\nu), \mathcal{J}^0), \mathrm{gap}\,(w(\nu_0), \mathcal{J}^0)\}}. \tag{48}$$

Proof First, recall that $p(\nu) = p(x(\nu); e, A)$ is related to $(Rx^{\mathrm{a}}(\nu), Rs^{\mathrm{a}}(\nu))$ according to (44) and let $\tilde{p}(\nu)$ denote the optimal solution of (45) with $h = e$, $z = x(\nu)$ and $J = \mathcal{J}^0$. Let $\mathcal{J}^0_i$ be a layer of $\mathcal{J}^0$ such that $\mathcal{J}^0_i \subseteq B$ and note that $\|p_{\mathcal{J}^0_i}(\nu_0)\|_\infty \le \varepsilon^{\mathrm{a}}_\infty(\nu_0)$ in view of (36) and (44). Using this fact, the fact that $\Delta(x_{\mathcal{J}^0_i}(\nu); x_{\mathcal{J}^0_i}(\nu_0)) \le \Delta(x_B(\nu); x_B(\nu_0)) \le \psi$, inequality (41) of Lemma 4.2 with $p_1 = \tilde{p}_{\mathcal{J}^0_i}(\nu_0)$ and $p_2 = \tilde{p}_{\mathcal{J}^0_i}(\nu)$, and Theorem 4.3 twice, we obtain

$$\begin{aligned}
\|p_{\mathcal{J}^0_i}(\nu)\| &\le \|\tilde{p}_{\mathcal{J}^0_i}(\nu) - p_{\mathcal{J}^0_i}(\nu)\| + \|\tilde{p}_{\mathcal{J}^0_i}(\nu)\| \le K(3 + 2K)\sqrt{n} + \|\tilde{p}_{\mathcal{J}^0_i}(\nu)\| \\
&\le K(3 + 2K)\sqrt{n} + \sqrt{|\mathcal{J}^0_i|}\psi + (1 + \psi)\|\tilde{p}_{\mathcal{J}^0_i}(\nu_0)\| \\
&\le K(3{+}2K)\sqrt{n}{+}\sqrt{n}\psi{+}(1{+}\psi)\left\{\|p_{\mathcal{J}^0_i}(\nu_0)\|{+}\|\tilde{p}_{\mathcal{J}^0_i}(\nu_0){-}p_{\mathcal{J}^0_i}(\nu_0)\|\right\} \\
&\le K(3{+}2K)\sqrt{n}{+}\sqrt{n}\psi{+}(1{+}\psi)\left\{\sqrt{|\mathcal{J}^0_i|}\,\|p_{\mathcal{J}^0_i}(\nu_0)\|_\infty{+}K(3{+}2K)\sqrt{n}\right\} \\
&\le \sqrt{n}\left\{\psi + (1 + \psi)\varepsilon^{\mathrm{a}}_\infty(\nu_0) + (2 + \psi)K(3 + 2K)\right\}.
\end{aligned}$$

Now, let $\mathcal{J}^0_i$ be a layer of $\mathcal{J}^0$ such that $\mathcal{J}^0_i \subseteq N$ and note that $\|e - p_{\mathcal{J}^0_i}(\nu_0)\|_\infty \le \varepsilon^{\mathrm{a}}_\infty(\nu_0)$ in view of (36) and (44). Using this fact, the fact that

$$\Delta(x_{\mathcal{J}^0_i}(\nu_0); x_{\mathcal{J}^0_i}(\nu)) = \Delta((s_{\mathcal{J}^0_i}(\nu); s_{\mathcal{J}^0_i}(\nu_0)) \le \Delta(s_N(\nu); s_N(\nu_0)) \le \psi,$$

inequality (42) of Lemma 4.2 and Theorem 4.3 twice, we can similarly show that

$$\|e - \tilde{p}_{\mathcal{J}_i^0}(\nu)\| \leq \sqrt{n}\left\{\psi + (1+\psi)\varepsilon_\infty^{\mathrm{a}}(\nu_0) + (2+\psi)K(3+2K)\right\}.$$

This inequality together with the previous one then imply that

$$\begin{aligned}
\varepsilon_\infty^{\mathrm{a}}(\nu) &\leq \max_{j=1,\ldots,n}\{\min(|p_j(\nu)|\,,\ |1-p_j(\nu)|)\}\\
&\leq \max_{i=1,\ldots,l}\{\min(\|p_{\mathcal{J}_i^0}(\nu)\|_\infty\,,\ \|e-p_{\mathcal{J}_i^0}(\nu)\|_\infty)\}\\
&\leq \sqrt{n}\left\{\psi + (1+\psi)\varepsilon_\infty^{\mathrm{a}}(\nu_0) + (2+\psi)K(3+2K)\right\}.
\end{aligned}$$

□

In the next six results, we will purify the upper bound on $\varepsilon_\infty^{\mathrm{a}}(\nu)$ obtained in Lemma 4.4 by deriving bounds on the quantities (47) and (48) in terms of $\varepsilon_\infty^{\mathrm{a}}(\nu_0)$ and ν/ν_0.

The next result is well-known. A proof of it can be found for example in [3].

Lemma 4.5 *If $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ satisfies $\phi(w) < 1$, then*

$$\max\left(\|x(\mu) - x\|_x, \|s(\mu) - s\|_s\right) \leq \frac{\phi(w)}{1-\phi(w)},$$

where $\mu := \mu(w)$ and $\phi(w)$ is defined by (6).

Lemma 4.6 *For every scalars $0 < \nu \leq \nu_0$, the point defined as*

$$w(\nu;\nu_0) := w(\nu_0) + (\nu - \nu_0)\dot{w}(\nu_0) \tag{49}$$

satisfies the following relations:

$$\mu(w(\nu;\nu_0)) = \nu, \tag{50}$$

$$\phi(w(\nu;\nu_0)) = \frac{(1-\nu/\nu_0)^2\nu_0\,\kappa(\nu_0)^2}{\nu} \leq \frac{\nu_0\sqrt{n}\,\varepsilon_\infty^{\mathrm{a}}(\nu_0)}{\nu}, \tag{51}$$

$$\max\left(\left\|\frac{x_B(\nu;\nu_0)}{x_B(\nu_0)} - e\right\|_\infty, \left\|\frac{s_N(\nu;\nu_0)}{s_N(\nu_0)} - e\right\|_\infty\right) \leq \varepsilon_\infty^{\mathrm{a}}(\nu_0). \tag{52}$$

Proof Differentiating (3) with respect to ν, we obtain $\dot{x}(\nu)s(\nu) + x(\nu)\dot{s}(\nu) = e$ for all $\nu > 0$. Using this relation together with (49), we easily see that $w(\nu;\nu_0) = (x(\nu;\nu_0), y(\nu;\nu_0), s(\nu;\nu_0))$ satisfies

$$x(\nu;\nu_0)s(\nu;\nu_0) = \nu e + (\nu_0 - \nu)^2\,\dot{x}(\nu_0)\dot{s}(\nu_0). \tag{53}$$

Differentiating (4) and (5) with respect to ν, we conclude that $A\dot{x}(\nu) = 0$ and $A^T\dot{y}(\nu) + \dot{s}(\nu) = 0$, and hence that $\dot{x}(\nu)^T\dot{s}(\nu) = 0$, for all $\nu > 0$. Relation (50) now follows by multiplying (53) on the left by $(e/n)^T$ and using the last observation. Moreover, using (50), (53) and Lemma 4.1(iii), we derive (51) as follows:

$$\begin{aligned}\phi(w(\nu;\nu_0)) &= \frac{\|x(\nu;\nu_0)s(\nu;\nu_0)-\mu(w(\nu;\nu_0))e\|}{\mu(w(\nu;\nu_0))} = \frac{(\nu_0-\nu)^2\|\dot{x}(\nu_0)\dot{s}(\nu_0)\|}{\nu}\\ &= \frac{(1-\nu/\nu_0)^2\nu_0\,\kappa(\nu_0)^2}{\nu} \le \frac{\nu_0\sqrt{n}\,\varepsilon^{\mathrm{a}}_{\infty}(\nu_0)}{\nu}.\end{aligned}$$

Also, (49) and (36) imply that

$$\begin{aligned}&\max\left(\left\|\frac{x_B(\nu;\nu_0)}{x_B(\nu_0)}-e\right\|_\infty, \left\|\frac{s_N(\nu;\nu_0)}{s_N(\nu_0)}-e\right\|_\infty\right)\\ &= (\nu_0-\nu)\max\left(\left\|\frac{\dot{x}_B(\nu_0)}{x_B(\nu_0)}\right\|_\infty, \left\|\frac{\dot{s}_N(\nu_0)}{s_N(\nu_0)}\right\|_\infty\right)\\ &= \left(1-\frac{\nu}{\nu_0}\right)\max\left(\|Rs^{\mathrm{a}}_B(\nu_0)\|_\infty, \|Rx^{\mathrm{a}}_N(\nu_0)\|_\infty\right) \le \varepsilon^{\mathrm{a}}_\infty(\nu_0).\end{aligned}$$

□

Lemma 4.7 *For $i=1,2$, assume that $u_i\in\Re^n$ and $\theta_i>0$ satisfy $\|u_i-e\|_\infty\le\theta_i$. Then, $\|u_1u_2-e\|_\infty\le\theta_1+\theta_2+\theta_1\theta_2$, and hence $u_1u_2\le(1+\theta_1)(1+\theta_2)e$. If, in addition, $\theta_i\le 1$ for $i=1,2$, then $u_1u_2\ge(1-\theta_1)(1-\theta_2)e$.*

Proof The assumption $\|u_1-e\|_\infty\le\theta_1$ implies that $\|u_1\|_\infty\le 1+\theta_1$. Hence,

$$\begin{aligned}\|u_1u_2-e\|_\infty &= \|u_1(u_2-e)+u_1-e\|_\infty\\ &\le \|u_1\|_\infty\|u_2-e\|_\infty+\|u_1-e\|_\infty\\ &\le (1+\theta_1)\theta_2+\theta_1.\end{aligned}$$

The assumption that $\|u_i-e\|_\infty\le\theta_i\le 1$ implies that $u_i\ge(1-\theta_i)e\ge 0$ for $i=1,2$. Combining these two inequalities, we obtain the last inequality of the lemma. □

Lemma 4.8 *Let $0<\nu\le\nu_0$ be scalars such that $\phi_{\nu_0}(\nu):=\phi(w(\nu;\nu_0))<1$. Then,*

$$\max\left\{\left\|\frac{x_B(\nu)}{x_B(\nu_0)}-e\right\|_\infty, \left\|\frac{s_N(\nu)}{s_N(\nu_0)}-e\right\|_\infty\right\} \le \frac{\varepsilon^{\mathrm{a}}_\infty(\nu_0)+\phi_{\nu_0}(\nu)}{1-\phi_{\nu_0}(\nu)}.$$

As a consequence, we have:

$$\frac{\nu}{\nu_0}\max\left(\frac{s_B(\nu_0)}{s_B(\nu)}, \frac{x_N(\nu_0)}{x_N(\nu)}\right) = \max\left(\frac{x_B(\nu)}{x_B(\nu_0)}, \frac{s_N(\nu)}{s_N(\nu_0)}\right) \le \frac{1+\varepsilon^{\mathrm{a}}_\infty(\nu_0)}{1-\phi_{\nu_0}(\nu)}. \quad (54)$$

If, in addition, $\varepsilon^{\rm a}_\infty(\nu_0) < 1$ *and* $\phi_{\nu_0}(\nu) < 1/2$*, we have*:

$$\frac{\nu}{\nu_0}\min\left(\frac{s_B(\nu_0)}{s_B(\nu)}, \frac{x_N(\nu_0)}{x_N(\nu)}\right) = \min\left(\frac{x_B(\nu)}{x_B(\nu_0)}, \frac{s_N(\nu)}{s_N(\nu_0)}\right) \geq \frac{1-2\phi_{\nu_0}(\nu)}{1-\phi_{\nu_0}(\nu)}(1-\varepsilon^{\rm a}_\infty(\nu_0)). \tag{55}$$

Proof By Lemma 4.5 and the assumption that $\phi(w(\nu;\nu_0)) < 1$, it follows that

$$\max\left\{\left\|\frac{x_B(\nu)}{x_B(\nu;\nu_0)} - e\right\|_\infty, \left\|\frac{s_N(\nu)}{s_N(\nu;\nu_0)} - e\right\|_\infty\right\} \leq \max\left\{\left\|\frac{x(\nu)}{x(\nu;\nu_0)} - e\right\|, \left\|\frac{s(\nu)}{s(\nu;\nu_0)} - e\right\|\right\} \leq \frac{\phi_{\nu_0}(\nu)}{1-\phi_{\nu_0}(\nu)}.$$

The result now follows from the above relation, inequality (52) and Lemma 4.7 with $\theta_1 = \varepsilon^{\rm a}_\infty(\nu_0)$ and $\theta_2 = \phi_{\nu_0}(\nu)/(1-\phi_{\nu_0}(\nu))$. □

The following lemma derives some properties of the function gap $(\cdot, \mathcal{J}(\nu_0))$ along the central path $w(\nu)$.

Lemma 4.9 *Let* $0 < \nu \leq \nu_0$ *be scalars such that* $\phi_{\nu_0}(\nu) < 1/2$ *and* $\varepsilon^{\rm a}_\infty(\nu_0) < 1$. *Then,*

$$\frac{\text{gap}\,(w(\nu), \mathcal{J}(\nu_0))}{\text{gap}\,(\nu_0)} \geq \frac{\nu}{\nu_0}(1-2\phi_{\nu_0}(\nu))(1-\varepsilon^{\rm a}_\infty(\nu_0))^2. \tag{56}$$

In addition, the following statements hold:

i) *If* gap $(\nu_0) = x_i(\nu_0)/x_j(\nu_0)$ *for some* $i \in N(\nu_0)$ *and* $j \in B(\nu_0)$*, then*

$$\frac{\text{gap}\,(w(\nu), \mathcal{J}(\nu_0))}{\text{gap}\,(\nu_0)} \leq \frac{\nu}{\nu_0}\left(\frac{(1-\phi_{\nu_0}(\nu))^2}{(1-\varepsilon^{\rm a}_\infty(\nu_0))^2(1-2\phi_{\nu_0}(\nu))^2}\right). \tag{57}$$

ii) *If* gap $(\nu) = x_i(\nu)/x_j(\nu)$ *for some* $i \in B(\nu_0)$ *and* $j \in N(\nu_0)$*, then*

$$\frac{\text{gap}\,(\nu)}{\text{gap}\,(w(\nu_0), \mathcal{J}(\nu))} \geq \frac{\nu_0}{\nu}\left(\frac{(1-2\phi_{\nu_0}(\nu))^2(1-\varepsilon^{\rm a}_\infty(\nu_0))^2}{(1-\phi_{\nu_0}(\nu))^2}\right). \tag{58}$$

Proof Using Lemma 4.8, it is easy to see that

$$\frac{\nu}{\nu_0}(1-\phi_{\nu_0}(\nu))(1-\varepsilon^{\rm a}_\infty(\nu_0)) \leq \frac{x_i(\nu)}{x_i(\nu_0)} \leq \frac{1-\phi_{\nu_0}(\nu)}{(1-\varepsilon^{\rm a}_\infty(\nu_0))(1-2\phi_{\nu_0}(\nu))}.$$

for all $i = 1, \ldots, n$. Hence, for every $i, j \in \{1, \ldots, n\}$, we have

$$\frac{x_i(\nu)}{x_j(\nu)} \geq \frac{\nu}{\nu_0}(1-2\phi_{\nu_0}(\nu))(1-\varepsilon^{\rm a}_\infty(\nu_0))^2\,\frac{x_i(\nu_0)}{x_j(\nu_0)}.$$

In view of (32), this inequality together with the fact that $\text{gap}\,(\nu_0) := \text{gap}\,(w(\nu_0), \mathcal{J}(\nu_0))$ imply (56).

Assume now that $\text{gap}\,(\nu_0) = x_i(\nu_0)/x_j(\nu_0)$ for some $i \in N(\nu_0)$ and $j \in B(\nu_0)$. Then, in view of (31), we have that $i \in \mathcal{J}_k^0$ and $j \in \mathcal{J}_{k+1}^0$, where $\mathcal{J}_k^0$ and $\mathcal{J}_{k+1}^0$ are two consecutive layers of $\mathcal{J}(\nu_0)$. Using (32) and (55), we obtain

$$\begin{aligned}\frac{\text{gap}\,(w(\nu), \mathcal{J}(\nu_0))}{\text{gap}\,(\nu_0)} &\leq \frac{x_i(\nu)/x_j(\nu)}{\text{gap}\,(\nu_0)} = \frac{x_i(\nu)/x_i(\nu_0)}{x_j(\nu)/x_j(\nu_0)}\\ &\leq \frac{\nu}{\nu_0}\frac{(1-\phi_{\nu_0}(\nu))^2}{(1-\varepsilon_\infty^{\mathrm{a}}(\nu_0))^2(1-2\phi_{\nu_0}(\nu))^2},\end{aligned}$$

showing that statement i) of the lemma holds.

To show statement ii), assume now that $\text{gap}\,(\nu) = x_i(\nu)/x_j(\nu)$ for some $i \in B(\nu_0)$ and $j \in N(\nu_0)$. Then, in view of (31), we have that $i \in \mathcal{J}_k$ and $j \in \mathcal{J}_{k+1}$, where $\mathcal{J}_k$ and $\mathcal{J}_{k+1}$ are two consecutive layers of $\mathcal{J}(\mu)$. Using (32) and (55), we obtain

$$\begin{aligned}\frac{\text{gap}\,(\nu)}{\text{gap}\,(w(\nu_0), \mathcal{J}(\nu))} &\geq \frac{\text{gap}\,(\nu)}{x_i(\nu_0)/x_j(\nu_0)} = \frac{x_i(\nu)}{x_i(\nu_0)}\,\frac{x_j(\nu_0)}{x_j(\nu)}\\ &\geq \frac{\nu_0}{\nu}\frac{(1-\varepsilon_\infty^{\mathrm{a}}(\nu_0))^2(1-2\phi_{\nu_0}(\nu))^2}{(1-\phi_{\nu_0}(\nu))^2},\end{aligned}$$

showing that statement ii) of the lemma also holds. □

With the aid of Lemmas 4.8 and 4.9, the following result purifies the upper bound on $\varepsilon_\infty^{\mathrm{a}}(\nu)$ derived in Lemma 4.4.

Lemma 4.10 *Let $0 < \nu \leq \nu_0$ be scalars such that $\nu/\nu_0 \geq 4\sqrt{n}\varepsilon_\infty^{\mathrm{a}}(\nu_0)$. Then,*

$$\varepsilon_\infty^{\mathrm{a}}(\nu) \leq \sqrt{n}\left\{\left(3+\frac{4\sqrt{n}\,\nu_0}{3\,\nu}\right)\varepsilon_\infty^{\mathrm{a}}(\nu_0) + \frac{32\bar{\chi}_A(\nu_0/\nu)}{\text{gap}\,(\nu_0)} + \frac{68\,\bar{\chi}_A^2(\nu_0/\nu)^2}{[\text{gap}\,(\nu_0)]^2}\right\}.$$

Proof First note that the assumptions and Lemma 4.6 imply that

$$\varepsilon_\infty^{\mathrm{a}}(\nu_0) \leq \frac{1}{4\sqrt{n}} \leq \frac{1}{4}, \quad \phi_{\nu_0}(\nu) \leq \frac{\nu_0\sqrt{n}\,\varepsilon_\infty^{\mathrm{a}}(\nu_0)}{\nu} \leq \frac{1}{4}. \tag{59}$$

Using these inequalities together with Lemma 4.9, we can now bound the quantities $K(\nu, \nu_0)$ and $\psi(\nu, \nu_0)$ defined in (48) and (47), respectively, as

$$K(\nu, \nu_0) \leq \frac{\bar{\chi}_A\,(\nu_0/\nu)}{\text{gap}\,(\nu_0)(1-2\phi_{\nu_0}(\nu))(1-\varepsilon_\infty^{\mathrm{a}}(\nu_0))^2} \leq \frac{32\,\bar{\chi}_A\,(\nu_0/\nu)}{9\,\text{gap}\,(\nu_0)}$$

and

$$\psi(\nu, \nu_0) \leq \frac{\varepsilon_\infty^{\mathrm{a}}(\nu_0)+\phi_{\nu_0}(\nu)}{1-\phi_{\nu_0}(\nu)} \leq \frac{4}{3}\left[\varepsilon_\infty^{\mathrm{a}}(\nu_0)+\phi_{\nu_0}(\nu)\right] \leq \frac{4}{3}\varepsilon_\infty^{\mathrm{a}}(\nu_0)\left(1+\frac{\nu_0\sqrt{n}}{\nu}\right) \leq \frac{2}{3}.$$

The last two inequalities together with Lemma 4.4 then imply that

$$\begin{aligned}
\varepsilon_\infty^{\mathrm{a}}(\nu) &\le \sqrt{n}\left\{\psi + (1+\psi)\varepsilon_\infty^{\mathrm{a}}(\nu_0) + (2+\psi)K(3+2K)\right\} \\
&\le \sqrt{n}\left\{\frac{4}{3}\varepsilon_\infty^{\mathrm{a}}(\nu_0)\left(1 + \frac{\nu_0\sqrt{n}}{\nu}\right) + \frac{5}{3}\varepsilon_\infty^{\mathrm{a}}(\nu_0) + 8K + \frac{16}{3}K^2\right\} \\
&\le \sqrt{n}\left\{\left(3 + \frac{4\sqrt{n}\,\nu_0}{3\,\nu}\right)\varepsilon_\infty^{\mathrm{a}}(\nu_0) + \frac{32\bar{\chi}_A(\nu_0/\nu)}{\mathrm{gap}\,(\nu_0)} + \frac{68\,\bar{\chi}_A^2(\nu_0/\nu)^2}{[\mathrm{gap}\,(\nu_0)]^2}\right\}.
\end{aligned}$$

□

As an immediate consequence of Lemma 4.10, we can now derive the main result of this section.

Lemma 4.11 *Let* $\nu_0 > 0$ *be such that* $4\sqrt{n}\varepsilon_\infty^{\mathrm{a}}(\nu_0) \le 1$. *Then, for every* $\nu_1 > 0$ *such that* $\nu_1/\nu_0 \in [4\sqrt{n}\varepsilon_\infty^{\mathrm{a}}(\nu_0), 1]$, *we have*

$$\int_{\nu_1}^{\nu_0} \frac{\kappa(\nu)}{\nu}\,d\nu \le \sqrt{n}\left[\left(\sqrt{12} + \frac{2n^{1/4}}{\sqrt{3}}\right)\Phi_0(\nu_0,\nu_1) + \sqrt{32\,\Phi_1(\nu_0,\nu_1)} + \sqrt{68}\,\Phi_1(\nu_0,\nu_1)\right], \tag{60}$$

where

$$\Phi_0(\nu_0,\nu_1) := \left(\varepsilon_\infty^{\mathrm{a}}(\nu_0)\frac{\nu_0}{\nu_1}\right)^{1/2}, \quad \Phi_1(\nu_0,\nu_1) := \frac{\bar{\chi}_A\,\nu_0}{\mathrm{gap}\,(\nu_0)\,\nu_1}. \tag{61}$$

Proof By Lemmas 4.1 and 4.10, we have for every $\nu \in [4\sqrt{n}\,\varepsilon_\infty^{\mathrm{a}}(\nu_0)\nu_0, \nu_0]$ that

$$\begin{aligned}
\frac{\kappa(\nu)}{\nu} &\le \frac{n^{1/4}}{\nu}\varepsilon_\infty^{\mathrm{a}}(\nu)^{1/2} \\
&\le \frac{\sqrt{n}}{\nu}\left[3\varepsilon_\infty^{\mathrm{a}}(\nu_0) + \left(\frac{4\sqrt{n}\,\varepsilon_\infty^{\mathrm{a}}(\nu_0)}{3} + \frac{32\,\bar{\chi}_A}{\mathrm{gap}\,(\nu_0)}\right)\frac{\nu_0}{\nu} + \frac{68\,\bar{\chi}_A^2\,\nu_0^2}{\mathrm{gap}\,(\nu_0)^2\,\nu^2}\right]^{1/2} \\
&\le \sqrt{n}\left[\frac{\sqrt{3\,\varepsilon_\infty^{\mathrm{a}}(\nu_0)}}{\nu} + \left(\frac{2n^{1/4}\sqrt{\varepsilon_\infty^{\mathrm{a}}(\nu_0)}}{\sqrt{3}} + \frac{\sqrt{32\bar{\chi}_A}}{\mathrm{gap}\,(\nu_0)^{1/2}}\right)\frac{\nu_0^{1/2}}{\nu^{3/2}} + \frac{\sqrt{68}\bar{\chi}_A}{\mathrm{gap}\,(\nu_0)}\frac{\nu_0}{\nu^2}\right].
\end{aligned}$$

Hence,

$$
\begin{aligned}
&\int_{\nu_1}^{\nu_0} \frac{\kappa(\nu)}{\nu}\, d\nu \\
&\quad \le \sqrt{n}\left[\sqrt{3\varepsilon_\infty^{\mathrm{a}}(\nu_0)}\log\frac{\nu_0}{\nu_1}+\left(\frac{2n^{1/4}\sqrt{\varepsilon_\infty^{\mathrm{a}}(\nu_0)}}{\sqrt{3}}+\frac{\sqrt{32}\,\bar\chi_A}{\mathrm{gap}\,(\nu_0)^{1/2}}\right)\sqrt{\frac{\nu_0}{\nu_1}}+\frac{\sqrt{68}\,\bar\chi_A\,\nu_0}{\mathrm{gap}\,(\nu_0)\,\nu_1}\right] \\
&\quad \le \sqrt{n}\left[\sqrt{12\,\varepsilon_\infty^{\mathrm{a}}(\nu_0)\,\frac{\nu_0}{\nu_1}}+\left(\frac{2n^{1/4}\sqrt{\varepsilon_\infty^{\mathrm{a}}(\nu_0)}}{\sqrt{3}}+\frac{\sqrt{32}\,\bar\chi_A}{\mathrm{gap}\,(\nu_0)^{1/2}}\right)\sqrt{\frac{\nu_0}{\nu_1}}+\frac{\sqrt{68}\,\bar\chi_A\,\nu_0}{\mathrm{gap}\,(\nu_0)\,\nu_1}\right],
\end{aligned}
$$

where in the last relation we used the fact that $\log\alpha = 2\,\log(\sqrt{\alpha}) \le 2(\sqrt{\alpha}-1) \le 2\sqrt{\alpha}$ for every $\alpha > 0$. Relation (60) now follows from the above relation and (61). $\square$

5 Proof of the main results

In this section, we prove the main results of this paper, namely Theorems 2.4 and 2.5. The structure of the analysis of this section is similar to that of Sect. 4 of [10], but requires new and more involved ideas, such as the bound developed in Lemma 4.11.

We will now give a brief outline of the analysis of this section. Letting $\bar g := 348n\bar\chi_A$, we will show in Lemmas 5.1, 5.3 and 5.7 that, for any scalar $\nu_0 > 0$ satisfying one of the following three conditions:

(i) $\mathrm{gap}\,(\nu_0) \le \bar g$ (Lemma 5.1);
(ii) $\mathrm{gap}\,(\nu_0) \ge \bar g$ and $\varepsilon_\infty^{\mathrm{a}}(\nu_0) \ge 24\sqrt{n}\bar g/\mathrm{gap}\,(\nu_0)$ (Lemma 5.3);
(iii) $\mathrm{gap}\,(\nu_0) \ge \bar g$, $\varepsilon_\infty^{\mathrm{a}}(\nu_0) \le 24\sqrt{n}\bar g/\mathrm{gap}\,(\nu_0)$ and $\mathrm{gap}\,(\nu_0) = x_i(\nu_0)/x_j(\nu_0)$ for some $i \in N(\nu_0)$ and $j \in B(\nu_0)$ (Lemma 5.7),

there exists a scalar $\nu_1 \le \nu_0$ such that $I(\nu_1,\nu_0) = \mathcal{O}(n^{1.5}\log(n+\bar\chi_A))$ and the interval $(\nu_1,\nu_0]$ contains a $\bar g^n$-crossover event. Moreover, we will show in Lemma 5.8 that if $(\nu_l,\nu_u] \subset \Re_{++}$ is an interval which does not contain a scalar ν_0 satisfying one of the above three conditions, then we must have $I(\nu_l,\nu_u) \le 25/9$. Now, a simple argument shows that it is possible to express $\Re_{++}$ as the union of disjoint (possibly empty) intervals of the above two types which alternate from one type to the other as we traverse the set $\Re_{++}$. Since the number of intervals of the first type can not exceed $n(n-1)/2$ in view of Proposition 3.1, a simpe argument reveals that $I(0,\infty) = \mathcal{O}(n^{3.5}\log(n+\bar\chi_A))$.

As a consequence of Lemmas 5.1 and 5.3, we will prove in Theorem 5.5 a geometric result about the central path which claims that the points in the path with curvature larger than a given threshold value $\bar\kappa > 0$ lie in $\mathcal{O}(n^2)$ intervals, each with logarithmic length bounded by $\mathcal{O}(n\log(\bar\chi_A^*+n)+\log(\bar\kappa^{-1}))$.

Lemma 5.1 *Let $\nu_0 > 0$ be given and let $\bar g$ be a constant such that*

$$\bar g \ge \max\left\{4\,n\bar\chi_A\,,\ 48\sqrt{n}\,\bar\chi_A\right\}. \tag{62}$$

If gap $(\nu_0) \leq \bar{g}$*, then there exists* $\nu_1 \in (0, \nu_0)$ *such that the interval* $(\nu_1, \nu_0]$ *contains a* $\bar{g}^n$*-crossover event and*

$$\log \frac{\nu_0}{\nu_1} = \mathcal{O}\left(\log(\bar{\chi}_A + n) + n \log \bar{g}\right), \tag{63}$$

$$\int_{\nu_1}^{\nu_0} \frac{\kappa(\nu)}{\nu}\, d\nu = \mathcal{O}\left(\sqrt{n}\, \log(\bar{\chi}_A + n) + n^{1.5} \log \bar{g}\right). \tag{64}$$

Proof To simplify notation, let $(Rx^{\mathrm{a}}, Rs^{\mathrm{a}}) := (Rx^{\mathrm{a}}(\nu_0), Rs^{\mathrm{a}}(\nu_0))$. Moreover, let $J = (J_1, \ldots, J_p)$ denote the VY $\bar{g}$–partition at $w(\nu_0)$ and let $(Rx^{\mathrm{ll}}, Rs^{\mathrm{ll}})$ be the residual of the LLS direction at $w(\nu_0)$ with respect to J. Note that, by Proposition 3.4, we have gap $(w(\nu_0), J) > \bar{g}$ and $\mathrm{spr}(J_l) \leq \bar{g}^n$ for every $l = 1, \ldots, n$. It is easy to see that the assumption that gap $(\nu_0) \leq \bar{g}$ and the fact that J is the VY $\bar{g}$–partition at $w(\nu_0)$ imply the existence of two indices i, j lying in some layer J_q of J, with one contained in $B(\nu_0)$ and the other in $N(\nu_0)$. Without loss of generality, assume that $i \in B(\nu_0)$ and $j \in N(\nu_0)$. By Lemma 4.1(i), we have $Rx^{\mathrm{a}}_i \geq 1/2$ and $Rs^{\mathrm{a}}_j \geq 1/2$. Since $i, j \in J_q$, this implies that $\min\{\|Rx^{\mathrm{a}}_{J_q}\|_\infty, \|Rs^{\mathrm{a}}_{J_q}\|_\infty\} \geq 1/2$. Moreover, Proposition 3.2 together with the fact that gap $(w(\nu_0), J) > \bar{g}$ and relation (62) imply that

$$\max\left\{ \left\|Rx^{\mathrm{a}} - Rx^{\mathrm{ll}}\right\|_\infty, \left\|Rs^{\mathrm{a}} - Rs^{\mathrm{ll}}\right\|_\infty \right\} \leq \frac{12\sqrt{n}\,\bar{\chi}_A}{\mathrm{gap}\,(w(\nu_0), J)} \leq \frac{1}{4}. \tag{65}$$

Hence, we have

$$\begin{aligned}
&\min\left\{\|Rx^{\mathrm{ll}}_{J_q}\|_\infty, \|Rs^{\mathrm{ll}}_{J_q}\|_\infty\right\} \\
&\geq \min\left\{ \left\|Rx^{\mathrm{a}}_{J_q}\right\|_\infty - \left\|Rx^{\mathrm{a}} - Rx^{\mathrm{ll}}\right\|_\infty, \left\|Rs^{\mathrm{a}}_{J_q}\right\|_\infty - \left\|Rs^{\mathrm{a}} - Rs^{\mathrm{ll}}\right\|_\infty \right\} \\
&\geq \min\left\{\left\|Rx^{\mathrm{a}}_{J_q}\right\|_\infty, \left\|Rs^{\mathrm{a}}_{J_q}\right\|_\infty\right\} - \frac{1}{4} \geq \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.
\end{aligned} \tag{66}$$

By Lemma 3.3 with $\mathcal{C}_q = \bar{g}^n$, we know that the interval $(\nu, \nu_0]$ contains a $\bar{g}^n$-crossover event for any $\nu > 0$ satisfying (33). Letting ν_1 be the largest $\nu > 0$ satisfying (33), we then have

$$\begin{aligned}
\log \frac{\nu_0}{\nu_1} &= \mathcal{O}\left(\log(\bar{\chi}_A + n) + \log \mathcal{C}_q\right) + \log\left(\|Rx^{\mathrm{ll}}_{J_q}\|_\infty^{-1} \|Rs^{\mathrm{ll}}_{J_q}\|_\infty^{-1}\right) \\
&= \mathcal{O}\left(\log(\bar{\chi}_A + n) + n \log \bar{g}\right),
\end{aligned}$$

where the last inequality is due to (66). The bound on the integral now follows from the observation that, for any $0 < \nu_1 < \nu_0$, we have

$$\int_{\nu_1}^{\nu_0} \frac{\kappa(\nu)}{\nu}\, d\nu \leq \sqrt{\frac{n}{2}} \log \frac{\nu_0}{\nu_1}, \tag{67}$$

in view of Lemma 2.1(i). $\square$

Lemma 5.2 *Let $\nu_0 > 0$ be given and let $\bar{g}$ be a constant satisfying (62). Let (Rx^1, Rs^1) denote the residual of the LLS direction at $w(\nu_0)$ with respect to $\mathcal{J}$, where $\mathcal{J} = (\mathcal{J}_1, \ldots, \mathcal{J}_l)$ is an ordered AS-partition at $w(\nu_0)$. Assume that*

$$\varepsilon^1_\infty := \max\left\{\left\|Rx^1_N\right\|_\infty, \left\|Rs^1_B\right\|_\infty\right\} > 0, \tag{68}$$

where $(B, N) := (B(\nu_0), N(\nu_0))$. If $\operatorname{gap}(\nu_0) > \bar{g}$, then there exists $\nu_1 \in (0, \nu_0)$ such that the interval $(\nu_1, \nu_0]$ contains a $\bar{g}^n$-crossover event and

$$\log\frac{\nu_0}{\nu_1} = \mathcal{O}\left(\log(\bar{\chi}_A + n) + n\log\bar{g}\right) + \log\left((\varepsilon^1_\infty)^{-1}\right). \tag{69}$$

Proof Assume without loss of generality that $\varepsilon^1_\infty = \|Rx^1_N\|_\infty$; the case in which $\varepsilon^1_\infty = \|Rs^1_B\|_\infty$ can be proved similarly. Then, $\varepsilon^1_\infty = |Rx^1_i|$ for some $i \in N$. Let $\mathcal{J}_t$ be the layer of $\mathcal{J}$ containing the index i and note that

$$\varepsilon^1_\infty \;=\; |Rx^1_i| \;=\; \|Rx^1_{\mathcal{J}_t}\|_\infty \;\le\; \|Rx^1_{\mathcal{J}_t}\|. \tag{70}$$

Now, let $\mathcal{I} = (\mathcal{I}_1, \ldots, \mathcal{I}_p)$ be the VY $\bar{g}$-partition of $\mathcal{J}_t$ at $w(\nu_0)$ and consider the ordered partition $\mathcal{J}'$ defined as

$$\mathcal{J}' := (\mathcal{J}_1, \ldots, \mathcal{J}_{t-1}, \mathcal{I}_1, \ldots, \mathcal{I}_p, \mathcal{J}_{t+1}, \ldots, \mathcal{J}_l).$$

Let $(Rx^{\text{ll}}, Rs^{\text{ll}})$ denote the residual of the LLS direction at $w(\nu_0)$ with respect to $\mathcal{J}'$. Using the definition of the LLS step, it is easy to see that $Rx^1_{\mathcal{J}_j} = Rx^{\text{ll}}_{\mathcal{J}_j}$ for all $j = t+1, \ldots, l$. Moreover, we have $\|Rx^1_{\mathcal{J}_t}\| \le \|Rx^{\text{ll}}_{\mathcal{J}_t}\|$ since $\|Rx^1_{\mathcal{J}_t}\|$ is the optimal value of the least squares problem which determines the $\Delta x^1_{\mathcal{J}_t}$-component of the LLS step with respect to $\mathcal{J}$, whereas $\|Rx^{\text{ll}}_{\mathcal{J}_t}\|$ is the objective value at a certain feasible solution for the same least squares problem. Hence, for some $q \in \{1, \ldots, p\}$, we have

$$\|Rx^{\text{ll}}_{\mathcal{I}_q}\|_\infty = \|Rx^{\text{ll}}_{\mathcal{J}_t}\|_\infty \ge \frac{1}{\sqrt{|\mathcal{J}_t|}}\|Rx^{\text{ll}}_{\mathcal{J}_t}\| \ge \frac{1}{\sqrt{n}}\|Rx^{\text{ll}}_{\mathcal{J}_t}\| \ge \frac{1}{\sqrt{n}}\|Rx^1_{\mathcal{J}_t}\|. \tag{71}$$

Combining (70) and (71), we then obtain

$$\|Rx^{\text{ll}}_{\mathcal{I}_q}\|_\infty \ge \frac{1}{\sqrt{n}}\varepsilon^1_\infty. \tag{72}$$

Let us now bound the quantity $\|Rs^{\text{ll}}_{\mathcal{I}_q}\|_\infty$ from below. First note that the assumption that $\operatorname{gap}(\nu_0) := \operatorname{gap}(w(\nu_0), \mathcal{J}) > \bar{g}$ and Proposition 3.4 imply that $\operatorname{gap}(w(\nu_0), \mathcal{J}') > \bar{g}$ and $\operatorname{spr}(\mathcal{I}_q) \le \bar{g}^n$. Using the triangle inequality for norms, Lemma 4.1(i) together with the fact that $\mathcal{I}_q \subseteq N$, Proposition 3.2 together with the fact that $\operatorname{gap}(w(\nu_0), \mathcal{J}') > \bar{g}$

and relation (62), we obtain

$$\begin{aligned}\|Rs^{\mathrm{ll}}_{\mathcal{I}_q}\|_\infty &\geq \|Rs^{\mathrm{a}}_{\mathcal{I}_q}\|_\infty - \|Rs^{\mathrm{ll}}_{\mathcal{I}_q} - Rs^{\mathrm{a}}_{\mathcal{I}_q}\|_\infty \\ &\geq \frac{1}{2} - \frac{12\sqrt{n}\,\bar{\chi}_A}{\mathrm{gap}\,(w(\nu_0), \mathcal{J}')} \geq \frac{1}{2} - \frac{1}{4} = \frac{1}{4},\end{aligned} \tag{73}$$

where $Rs^{\mathrm{a}} := Rs^{\mathrm{a}}(\nu_0)$. By Lemma 3.3 with $J = \mathcal{J}'$, $J_p = \mathcal{I}_q$ and $\mathcal{C}_q = \bar{g}^n$, we know that the interval $(\nu, \nu_0]$ contains a $\bar{g}^n$-crossover event for any $\nu > 0$ satisfying (33). Letting ν_1 be the largest $\nu > 0$ satisfying (33), we then have

$$\begin{aligned}\log\frac{\nu_0}{\nu_1} &= \mathcal{O}\left(\log(\bar{\chi}_A + n) + \log \mathcal{C}_q\right) + \log\left(\|Rx^{\mathrm{ll}}_{J_q}\|_\infty^{-1}\|Rs^{\mathrm{ll}}_{J_q}\|_\infty^{-1}\right) \\ &\leq \mathcal{O}\left(\log(\bar{\chi}_A + n) + n\log\bar{g}\right) + \log\left((\varepsilon^{\mathrm{l}}_\infty)^{-1}\right),\end{aligned}$$

where the last inequality is due to (72) and (73). $\square$

Lemma 5.3 *Let $\nu_0 > 0$ be given and let $\bar{g}$ be a constant satisfying (62). If $\mathrm{gap}\,(\nu_0) > \bar{g}$ and $\varepsilon^{\mathrm{a}}_\infty(\nu_0) \geq 24\sqrt{n}\bar{\chi}_A/\mathrm{gap}\,(\nu_0)$, then there exists $\nu_1 \in (0, \nu_0)$ such that the interval $(\nu_1, \nu_0]$ contains a $\bar{g}^n$-crossover event, relation (64) holds, and*

$$\log\frac{\nu_0}{\nu_1} = \mathcal{O}\left(\log(\bar{\chi}_A + n) + n\log\bar{g}\right) + \log\left(\varepsilon^{\mathrm{a}}_\infty(\nu_0)^{-1}\right), \tag{74}$$

Proof Let $(Rx^{\mathrm{a}}, Rs^{\mathrm{a}})$ and $(Rx^{\mathrm{l}}, Rs^{\mathrm{l}})$ denote the residuals of the AS direction at $w(\nu_0)$ and the LLS direction at $w(\nu_0)$ with respect to $\mathcal{J}$, respectively, where $\mathcal{J} = (\mathcal{J}_1, \ldots, \mathcal{J}_l)$ denotes an ordered AS-partition at $w(\nu_0)$. Using Proposition 3.2, relation (62) and the assumptions that $\mathrm{gap}\,(\nu_0) > \bar{g}$ and $\varepsilon^{\mathrm{a}}_\infty(\nu_0) \geq 24\sqrt{n}\bar{\chi}_A/\mathrm{gap}\,(\nu_0)$, we obtain

$$\max\left\{\left\|Rx^{\mathrm{a}} - Rx^{\mathrm{l}}\right\|_\infty, \left\|Rs^{\mathrm{a}} - Rs^{\mathrm{l}}\right\|_\infty\right\} \leq \frac{12\sqrt{n}\,\bar{\chi}_A}{\mathrm{gap}\,(\nu_0)} \leq \frac{\varepsilon^{\mathrm{a}}_\infty(\nu_0)}{2}.$$

Hence, we have

$$\begin{aligned}\varepsilon^{\mathrm{l}}_\infty &:= \max\left\{\|Rx^{\mathrm{l}}_N\|_\infty, \|Rs^{\mathrm{l}}_B\|_\infty\right\} \\ &\geq \max\left\{\|Rx^{\mathrm{a}}_N\|_\infty - \left\|Rx^{\mathrm{a}}_N - Rx^{\mathrm{l}}_N\right\|_\infty, \|Rs^{\mathrm{a}}_B\|_\infty - \left\|Rs^{\mathrm{a}}_B - Rs^{\mathrm{l}}_B\right\|_\infty\right\} \\ &\geq \max\left\{\|Rx^{\mathrm{a}}_N\|_\infty, \|Rs^{\mathrm{a}}_B\|_\infty\right\} - \frac{\varepsilon^{\mathrm{a}}_\infty(\nu_0)}{2} = \varepsilon^{\mathrm{a}}_\infty(\nu_0) - \frac{\varepsilon^{\mathrm{a}}_\infty(\nu_0)}{2} = \frac{\varepsilon^{\mathrm{a}}_\infty(\nu_0)}{2} > 0,\end{aligned}$$

where the strict inequality follows from Assumption A.3 and Lemma 4.1(iii). Hence, by Lemma 5.2, there exists $\nu_1 \in (0, \nu_0)$ such that the interval $(\nu_1, \nu_0]$ contains a $\bar{g}^n$-crossover event and relation (69) holds, which clearly implies (74) due to the above estimate.

It remains to show that (64) holds. If $\varepsilon^{\mathrm{a}}_\infty(\nu_0) \geq 1/(4\sqrt{n})$, then it follows from (74) that (63) holds in this case, and hence that (64) holds too due to (67). Assume then that $\varepsilon^{\mathrm{a}}_{\infty}(\nu_0) < 1/(4\sqrt{n})$ and let $\bar{\nu} := 4\sqrt{n}\,\varepsilon^{\mathrm{a}}_{\infty}(\nu_0)\nu_0 < \nu_0$. The definition of $\bar{\nu}$ and the assumption that $\varepsilon^{\mathrm{a}}_{\infty}(\nu_0) \geq 24\sqrt{n}\,\bar{\chi}_A/\mathrm{gap}\,(\nu_0)$ then imply that

$$\Phi_0(\nu_0, \bar{\nu}) := \sqrt{\varepsilon^{\mathrm{a}}_{\infty}(\nu_0)\,\frac{\nu_0}{\bar{\nu}}} = \frac{1}{2}\,n^{-1/4}, \quad \Phi_1(\nu_0, \bar{\nu}) := \frac{\bar{\chi}_A\,\nu_0}{\mathrm{gap}\,(\nu_0)\,\bar{\nu}} \leq \frac{1}{96\,n}.$$

Using Lemma 4.11 with $\nu_1 = \bar{\nu}$ and the above estimates, we then conclude that

$$\int_{\bar{\nu}}^{\nu_0} \frac{\kappa(\nu)}{\nu}\,d\nu = \mathcal{O}(\sqrt{n}). \tag{75}$$

Also, using (74) and the fact that $\nu_0/\bar{\nu} = (4\sqrt{n}\,\varepsilon^{\mathrm{a}}_{\infty}(\nu_0))^{-1}$, we easily see that

$$\log\frac{\bar{\nu}}{\nu_1} = \log\frac{\nu_0}{\nu_1} - \log\frac{\nu_0}{\bar{\nu}} = \mathcal{O}\left(\log(\bar{\chi}_A + n) + n\log\bar{g}\right),$$

which, in view of (67), implies that

$$\int_{\nu_1}^{\bar{\nu}} \frac{\kappa(\nu)}{\nu}\,d\nu = \mathcal{O}\left(\sqrt{n}\,\log(\bar{\chi}_A + n) + n^{1.5}\log\bar{g}\right) \tag{76}$$

Clearly, (64) follows by adding (75) and (76). $\square$

We will now use Lemmas 5.1 and 5.3 to establish a result about the geometric structure of the central path, namely Theorem 2.5. These two lemmas give two independent sets of conditions on a scalar $\nu_0 > 0$ which guarantee the occurrence of a crossover event in an interval of the form $(\nu_1, \nu_0]$. Both sets involve a condition on the quantity gap (ν_0), which is not scale-invariant. Moreover, the notion of crossover event itself is not scale-invariant. Nevertheless, it is possible to derive scale-invariant geometric properties of the central path in terms of its curvature using the above two lemmas as will be shown in Theorem 2.5. Before establishing this theorem, we need to state one more lemma which estimates the length of an interval $(\nu_1, \nu_0]$ containing a crossover event in terms of the curvature $\kappa(\nu_0)$.

Lemma 5.4 *Let* $\bar{\kappa} \in (0, \sqrt{n}]$ *be given and define*

$$\bar{g}(\bar{\kappa}) := \max\left\{4\,n\,\bar{\chi}_A\,,\ 48\sqrt{n}\,\bar{\chi}_A\,,\ \frac{24\,n\,\bar{\chi}_A}{\bar{\kappa}^2}\right\}. \tag{77}$$

Then, for every scalar $\nu_0 > 0$ *such that* $\kappa(\nu_0) \geq \bar{\kappa}$, *there exists* $\nu_1 \in (0, \nu_0)$ *such that the interval* $(\nu_1, \nu_0]$ *contains a* $[\bar{g}(\bar{\kappa})]^n$*-crossover event and*

$$\log\frac{\nu_0}{\nu_1} = \mathcal{O}\left(n\log(\bar{\chi}_A + n) + n|\log\bar{\kappa}|\right).$$

Proof Assume that $\nu_0 > 0$ is such that $\kappa(\nu_0) \geq \bar{\kappa}$ and let $\mathcal{J} = (\mathcal{J}_1, \ldots, \mathcal{J}_l)$ denote an ordered AS-partition at $w(\nu_0)$. There are two cases to consider, namely: i) $\text{gap}\,(\nu_0) \leq \bar{g}(\bar{\kappa})$; and ii) $\text{gap}\,(\nu_0) > \bar{g}(\bar{\kappa})$. If case i) holds then the conclusion of the lemma follows from Lemma 5.1 and definition (77) of $\bar{g}(\bar{\kappa})$. Consider now case ii) in which $\text{gap}\,(\nu_0) > \bar{g}(\bar{\kappa})$. Using the assumption that $\kappa(\nu_0) \geq \bar{\kappa}$, relation (77) and Lemma 4.1(iii), we conclude that in this case we have

$$\varepsilon^{\mathrm{a}}_{\infty}(\nu_0) \geq \frac{\kappa(\nu_0)^2}{\sqrt{n}} \geq \frac{\bar{\kappa}^2}{\sqrt{n}} = \frac{24\sqrt{n}\bar{\chi}_A}{\bar{g}(\bar{\kappa})} \geq \frac{24\sqrt{n}\bar{\chi}_A}{\text{gap}\,(\nu_0)}.$$

The conclusion of the lemma under case ii) now follows immediately from Lemma 5.3, the above inequality and relation (77). $\square$

Theorem 5.5 *For any constant $\bar{\kappa} \in (0, \sqrt{n}]$, there exist $l \leq n(n-1)/2$ closed intervals $I_k = [d_k, e_k]$, $k = 1, \ldots, l$, such that*:

i) $d_k \geq e_{k+1}$ *for all* $k = 1, \ldots, l-1$;
ii) $\{\nu > 0 : \kappa(\nu) \geq \bar{\kappa}\} \subseteq \cup_{k=1}^{l} I_k$;
iii) $\log(e_k/d_k) = \mathcal{O}\left(n \log(\bar{\chi}^*_A + n) + n|\log \bar{\kappa}|\right)$ *for all* $k = 1, \ldots, l$.

Proof The intervals I_k can be constructed inductively as follows. Suppose that $q \geq 0$ intervals $I_k = [d_k, e_k]$, $k = 1, \ldots, q$, have been constructed so that: a) properties i) and iii) hold with l replaced by q; b) $\{\nu \geq d_q : \kappa(\nu) \geq \bar{\kappa}\} \subseteq \cup_{k=1}^{q} I_k$, and; c) each interval I_k, $k = 1, \ldots, q$, contains a $[\bar{g}(\bar{\kappa})]^n$-crossover event. Note that, by Proposition 3.1, the latter property implies that q can not exceed $n(n-1)/2$. If the set $\{\nu \leq d_q : \kappa(\nu) \geq \bar{\kappa}\}$ is empty then property ii) obviously hold with $l = q$, and the conclusion of the theorem holds with $l = q$. Otherwise, let $e_{q+1} := \max\{\nu \leq d_q : \kappa(\nu) \geq \bar{\kappa}\}$. (Here, by convention, $d_0 := \infty$.) By Lemma 5.4, there exists $d_{q+1} < e_{q+1}$ such that $\log(e_{q+1}/d_{q+1}) = \mathcal{O}\left(n \log(\bar{\chi}_A + n) + n|\log \bar{\kappa}|\right)$ and the interval $I_{q+1} = [d_{q+1}, e_{q+1}]$ contains a $[\bar{g}(\bar{\kappa})]^n$-crossover event. Clearly, the intervals I_k, $k = 1, \ldots, q+1$, satisfy the above statements a), b) and c) with q replaced by $q+1$. Since the number of intervals I_k satisfying a), b) and c) above can not exceed $n(n-1)/2$, it is clearly that the above construction eventually yields intervals I_k's satisfying the conclusion of the theorem. $\square$

We will now continue with our goal of proving Theorem 2.4. Our next objective is to derive one more set of conditions on a scalar ν_0 which guarantees the occurrence of a crossover event in an interval of the form $(\nu_1, \nu_0]$ over which the curvature integral can be bounded as in (64). This set of conditions is slightly more general than the third set of conditions introduced at the beginning of this section. But first we need to establish the following technical result.

Lemma 5.6 *Let scalars $0 < \nu_1 < \nu_0$ be given. Then, the following implications hold*:

i) *if* $(B(\nu_1), N(\nu_1)) \neq (B(\nu_0), N(\nu_0))$, *then there exists* $\nu' \in [\nu_1, \nu_0]$ *such that* $\varepsilon^{\mathrm{a}}_{\infty}(\nu') \geq 1/2$;
ii) *if* $\mathcal{J}(\nu_1) \neq \mathcal{J}(\nu_0)$, *then there exists* $\nu' \in [\nu_1, \nu_0]$ *such that either* $\varepsilon^{\mathrm{a}}_{\infty}(\nu') \geq 1/2$ *or* $\text{gap}\,(w(\nu')) = 1$.

Proof We first prove i). If $\varepsilon^{\rm a}_\infty(\nu_0) \geq 1/2$, then the conclusion of implication i) obviously holds. Assume then that $\varepsilon^{\rm a}_\infty(\nu_0) < 1/2$. The assumption that $(B(\nu_1), N(\nu_1)) \neq (B(\nu_0), N(\nu_0))$ implies that either $B(\nu_1) \cap N(\nu_0) \neq \emptyset$ or $N(\nu_1) \cap B(\nu_0) \neq \emptyset$. Without loss of generality, assume that $B(\nu_1) \cap N(\nu_0) \neq \emptyset$ and let i be an index lying in this intersection. Then, $|Rx^{\rm a}_i(\nu_0)| \leq \varepsilon^{\rm a}_\infty(\nu_0) < 1/2$ since $i \in N(\nu_0)$. Moreover, by Lemma 4.1(i) and the fact that $i \in B(\nu_1)$, we have that $Rx^{\rm a}_i(\nu_1) \geq 1/2$. The intermediate value theorem applied to the continuous function $Rx^{\rm a}_i(\cdot)$ then implies the existence of a scalar $\nu' \in [\nu_1, \nu_0]$ such that $Rx^{\rm a}_i(\nu') = 1/2$. Since $Rx^{\rm a}_i(\nu') + Rs^{\rm a}_i(\nu') = 1$, we conclude that $Rs^{\rm a}_i(\nu') = 1/2$, and hence that $\varepsilon^{\rm a}_\infty(\nu') \geq \min\{|Rx^{\rm a}_i(\nu')|, |Rs^{\rm a}_i(\nu')|\} = 1/2$.

To prove ii), assume that $\mathcal{J}(\nu_1) \neq \mathcal{J}(\nu_0)$. We may also assume that $\mathrm{gap}\,(\nu_0) > 1$ and that $(B(\nu_1), N(\nu_1)) = (B(\nu_0), N(\nu_0))$ since otherwise the conclusion of ii) would either obviously hold or follow from i). The fact that $\mathcal{J}(\nu_1) \neq \mathcal{J}(\nu_0)$ and $(B(\nu_1), N(\nu_1)) = (B(\nu_0), N(\nu_0))$ together with Lemma 3.5 imply that $\mathrm{gap}\,(w(\nu_1), \mathcal{J}(\nu_0)) < 1$. Noting that $1 < \mathrm{gap}\,(\nu_0) := \mathrm{gap}\,(w(\nu_0), \mathcal{J}(\nu_0))$, it follows from the intermediate value theorem applied to the continuous function $\nu \to \mathrm{gap}\,(w(\nu), \mathcal{J}(\nu_0))$ that there exists $\nu' \in [\nu_1, \nu_0]$ such that $\mathrm{gap}\,(w(\nu'), \mathcal{J}(\nu_0)) = 1$. In view of Lemma 3.5, this implies that $\mathcal{J}(\nu_0) = \mathcal{J}(\nu')$ and hence that $\mathrm{gap}\,(\nu') := \mathrm{gap}\,(w(\nu'), \mathcal{J}(\nu')) = 1$. $\square$

We are now ready to establish the result mentioned in the paragraph before Lemma 5.6.

Lemma 5.7 *Let* $\nu_0 > 0$ *be given and let* $\bar{g}$ *be a constant satisfying* (62). *Assume that* $\mathrm{gap}\,(\nu_0) > \bar{g}$,

$$\varepsilon^{\rm a}_\infty(\nu_0) \leq \frac{\bar{g}}{16\sqrt{n}\,\mathrm{gap}\,(\nu_0)}, \tag{78}$$

and $\mathrm{gap}\,(\nu_0) = x_i(\nu_0)/x_j(\nu_0)$ *for some* $i \in N(\nu_0)$ *and* $j \in B(\nu_0)$. *Then, there* exists $\nu_1 \in (0, \nu_0)$ *such that the interval* $(\nu_1, \nu_0]$ *contains a* $\bar{g}^n$*-crossover event, relation* (64) *holds, and*

$$\begin{aligned}\log\frac{\nu_0}{\nu_1} &= \mathcal{O}\left(\log(\bar{\chi}_A + n) + n\log\bar{g}\,\right) + \log\left(\mathrm{gap}\,(\nu_0)\right),\\ &\leq \mathcal{O}\left(\log(\bar{\chi}_A + n) + n\log\bar{g}\,\right) + \log\left(\varepsilon^{\rm a}_\infty(\nu_0)^{-1}\right)\end{aligned} \tag{79}$$

Proof Let $\bar{\nu} := (1-\bar{t})\nu_0$, where

$$\bar{t} := 1 - \frac{\bar{g}}{4\,\mathrm{gap}\,(\nu_0)}. \tag{80}$$

In view of (78) and the assumption that $\mathrm{gap}\,(\nu_0) > \bar{g}$, we have $\bar{\nu}/\nu_0 = (1-\bar{t}) = \bar{g}/(4\mathrm{gap}\,(\nu_0)) \in [4\sqrt{n}\varepsilon^{\rm a}_\infty(\nu_0), 1]$. This conclusion together with (62) then imply that

$$\log\frac{\nu_0}{\bar{\nu}} = \log\left(\frac{4\,\mathrm{gap}\,(\nu_0)}{\bar{g}}\right) \leq \log(\mathrm{gap}\,(\nu_0)) \tag{81}$$

and

$$\Phi_0(\nu_0,\bar\nu) := \sqrt{\varepsilon_\infty^{\mathrm{a}}(\nu_0)\,\frac{\nu_0}{\bar\nu}} \le \frac{1}{2}\,n^{-1/4}, \quad \Phi_1(\nu_0,\bar\nu) := \frac{\bar\chi_A\,\nu_0}{\mathrm{gap}\,(\nu_0)\,\bar\nu} = \frac{4\bar\chi_A}{\bar g} \le \frac{1}{n}.$$

Using Lemma 4.11 with $\nu_1 = \bar\nu$ and the last two estimates, we then conclude that

$$\int_{\bar\nu}^{\nu_0} \frac{\kappa(\nu)}{\nu}\,d\nu = \mathcal{O}(\sqrt{n}). \tag{82}$$

We will next show that gap $(w(\bar\nu),\mathcal{J}) < \bar g$, where $\mathcal{J} := \mathcal{J}(\nu_0)$. Indeed, first note that the assumption that gap $(\nu_0) > \bar g$ and relations (51), (78) and (80) imply that

$$\begin{aligned} \varepsilon_\infty^{\mathrm{a}}(\nu_0) &\le \frac{1}{16\sqrt{n}} \le \frac{1}{16}, \\ \phi_{\nu_0}(\bar t) &\le \frac{\sqrt{n}\,\varepsilon_\infty^{\mathrm{a}}(\nu_0)}{1-\bar t} \le \frac{4\sqrt{n}\,\varepsilon_\infty^{\mathrm{a}}(\nu_0)\,\mathrm{gap}\,(\nu_0)}{\bar g} \le \frac{1}{4}. \end{aligned}$$

The above estimates together with Lemma 4.9(i) and relation (80) imply that

$$\begin{aligned} \mathrm{gap}\,(w(\bar\nu),\mathcal{J}) &\le \frac{(1-\bar t)(1-\phi_{\nu_0}(\bar t))^2}{(1-\varepsilon_\infty^{\mathrm{a}}(\nu_0))^2(1-2\phi_{\nu_0}(\bar t))^2}\,\mathrm{gap}\,(\nu_0) \\ &\le \frac{64}{25}(1-\bar t)\,\mathrm{gap}\,(\nu_0) \le \frac{16}{25}\bar g < \bar g. \end{aligned}$$

Since gap $(w(\bar\nu),\mathcal{J}) < \bar g$ and $\bar g <$ gap $(\nu_0) :=$ gap $(w(\nu_0),\mathcal{J})$ by assumption, it follows from the intermediate value theorem applied to the continuous function gap $(w(\cdot),\mathcal{J})$ that there exists a scalar $\hat\nu \in (\bar\nu,\nu_0)$ such that gap $(w(\hat\nu),\mathcal{J}) = \bar g$. We will now consider the following two possible cases separately: i) $(B(\hat\nu),N(\hat\nu)) = (B(\nu_0),N(\nu_0))$; and ii) $(B(\hat\nu),N(\hat\nu)) \ne (B(\nu_0),N(\nu_0))$. If case i) holds then it follows from Proposition 3.5 that $\mathcal{J}$ is also an AS partition at $w(\hat\nu)$, and hence that gap $(\hat\nu) :=$ gap $(w(\hat\nu)) =$ gap $(w(\hat\nu),\mathcal{J}) = \bar g$. This conclusion together with Lemma 5.1 then imply the existence of a scalar $\nu_1 \in (0,\hat\nu)$ such that the interval $(\nu_1,\hat\nu]$, and hence $(\nu_1,\nu_0]$, contains a $\bar g^n$-crossover event and

$$\log\frac{\hat\nu}{\nu_1} = \mathcal{O}\left(\log(\bar\chi_A + n) + n\log\bar g\right), \tag{83}$$

$$\int_{\nu_1}^{\hat\nu} \frac{\kappa(\nu)}{\nu}\,d\nu = \mathcal{O}\left(\sqrt{n}\,\log(\bar\chi_A+n) + n^{1.5}\log\bar g\right). \tag{84}$$

Note that in this case, (79) and (64) follow from the fact that $\bar\nu < \hat\nu$ and relations (81), (82), (83) and (84). Consider now case ii). In this case, it follows from Lemma 5.6(i) that $\varepsilon_\infty^{\mathrm{a}}(\nu') \ge 1/2$ for some $\nu' \in [\hat\nu,\nu_0]$. Applying Lemma 5.3 with $\nu_0 = \nu'$, we conclude that there exists $\nu_1 \in (0, \nu')$ such that the interval $(\nu_1, \nu']$, and hence $(\nu_1, \nu_0]$, contains a $\bar{g}^n$-crossover event and

$$\log \frac{\nu'}{\nu_1} = \mathcal{O}\left(\log(\bar{\chi}_A + n) + n \log \bar{g}\right), \tag{85}$$

$$\int_{\nu_1}^{\nu'} \frac{\kappa(\nu)}{\nu}\, d\nu = \mathcal{O}\left(\sqrt{n}\, \log(\bar{\chi}_A + n) + n^{1.5} \log \bar{g}\right). \tag{86}$$

As before, (79) and (64) follow from the fact that $\bar{\nu} < \nu'$ and relations (81), (82), (85) and (86). □

Our goal now will be to estimate the curvature integral over intervals which do not contain scalars satisfying the three set of conditions (i), (ii) and (iii) introduced at the beginning of this section. Estimation of the integral over an interval of this type is done by dividing the interval into a finite number of disjoint subintervals. The first lemma below estimates the integral on each one of these subintervals while the second lemma below gives an estimate of the integral over the whole interval with the aid of the estimate obtained in the first lemma.

Lemma 5.8 *Let* $0 < \nu_1 \le \nu_0$ *and* $\bar{g} \ge 1$ *be scalars such that*

$$\frac{\nu_1}{\nu_0} \ge \frac{\bar{g}}{4\text{gap}\,(\nu_0)} \tag{87}$$

and the following conditions hold for every $\nu \in [\nu_1, \nu_0]$:

i) $\text{gap}\,(\nu) > \bar{g}$;
ii) $\varepsilon^{\text{a}}_{\infty}(\nu) \le \bar{g}/(16\sqrt{n}\,\text{gap}\,(\nu))$;
iii) *there exist* $i \in B(\nu)$ *and* $j \in N(\nu)$ *such that* $\text{gap}\,(\nu) = x_i(\nu)/x_j(\nu)$.

Then, we have:

$$\text{gap}\,(\nu_1) \ge \frac{25\nu_0}{64\nu_1} \text{gap}\,(\nu_0), \tag{88}$$

$$\int_{\nu_1}^{\nu_0} \frac{\kappa(\nu)}{\nu}\, d\nu \le \frac{4}{5} \left(\frac{\bar{g}}{\text{gap}\,(\nu_0)}\right)^{1/2}. \tag{89}$$

Proof Using (51), (87) and assumptions i) and ii), we conclude that for every $\nu \in [\nu_1, \nu_0]$, we have

$$\begin{aligned} \varepsilon^{\text{a}}_{\infty}(\nu) &\le \frac{1}{16\sqrt{n}} \le \frac{1}{16}, \\ \phi_{\nu_0}(\nu) &\le \frac{\nu_0 \sqrt{n}\, \varepsilon^{\text{a}}_{\infty}(\nu_0)}{\nu} \le \frac{4\sqrt{n}\, \varepsilon^{\text{a}}_{\infty}(\nu_0)\, \text{gap}\,(\nu_0)}{\bar{g}} \le \frac{1}{4}. \end{aligned} \tag{90}$$

Also, assumption i), relation (90) and Lemma 5.6 (ii) imply that $(B(\nu), N(\nu)) = (B(\nu_0), N(\nu_0))$ and $\mathcal{J}(\nu_0) = \mathcal{J}(\nu)$. Using these observations together with Lemma 4.9 (ii) and assumption iii), we obtain

$$\frac{\text{gap}\,(\nu)}{\text{gap}\,(\nu_0)} = \frac{\text{gap}\,(\nu)}{\text{gap}\,(w(\nu_0), \mathcal{J}(\nu_0))} = \frac{\text{gap}\,(\nu)}{\text{gap}\,(w(\nu_0), \mathcal{J}(\nu))}$$
$$\geq \left(\frac{(1-2\phi_{\nu_0}(t))^2(1-\varepsilon^{\mathrm{a}}_\infty(\nu_0))^2}{(1-\phi_{\nu_0}(t))^2}\right)\frac{\nu_0}{\nu} \geq \frac{25}{64}\frac{\nu_0}{\nu}. \tag{91}$$

The above relation with $\nu = \nu_1$ yields (88). We will now show that (89) holds. Lemma 4.1(iii), assumption (ii) and relation (91) imply

$$\kappa(\nu)^2 \leq \sqrt{n}\varepsilon^{\mathrm{a}}_\infty(\nu) \leq \frac{\bar{g}}{16\text{gap}\,(\nu)} \leq \frac{4\,\bar{g}}{25\,\text{gap}\,(\nu_0)}\,\frac{\nu}{\nu_0}.$$

This implies that

$$\int_{\nu_1}^{\nu_0} \frac{\kappa(\nu)}{\nu}\,d\nu \leq \frac{2}{5}\left(\frac{\bar{g}}{\text{gap}\,(\nu_0)}\right)^{1/2}\int_{\nu_1}^{\nu_0}\frac{1}{\sqrt{\nu_0\nu}}\,d\nu \leq \frac{4}{5}\left(\frac{\bar{g}}{\text{gap}\,(\nu_0)}\right)^{1/2}.$$

□

Lemma 5.9 *Let* $\bar{g}$ *be a constant satisfying* (62) *and let* $0 < \nu_1 \leq \nu_0$ *be scalars such that conditions* i)–iii) *of Lemma* 5.8 *holds for every* $\nu \in [\nu_1, \nu_0]$. *Then, we have*

$$\int_{\nu_1}^{\nu_0}\frac{\kappa(\nu)}{\nu}\,d\nu \leq \frac{25}{9}.$$

Proof Let $\nu_1 = \mu_l < \cdots < \mu_1 < \mu_0 = \nu_0$ be a subdivision of the interval $[\nu_1, \nu_0]$ such that

$$\mu_{k+1} = \max\left\{\nu_1, \frac{\bar{g}\mu_k}{4\text{gap}\,(w(\mu_k))}\right\}, \quad k = 0, \ldots, l-1.$$

Lemma 5.8 and the above relation then imply that

$$\text{gap}\,(w(\mu_{k+1})) \geq \frac{25\mu_k}{64\mu_{k+1}}\text{gap}\,(w(\mu_k)) = \frac{25}{16}\frac{\text{gap}\,(w(\mu_k))^2}{\bar{g}}, \quad k = 0, \ldots, l-2,$$
$$\int_{\mu_{k+1}}^{\mu_k}\frac{\kappa(\nu)}{\nu}\,d\nu \leq \frac{4}{5}\left(\frac{\bar{g}}{\text{gap}\,(w(\mu_k))}\right)^{1/2}, \quad k = 0, \ldots, l-1.$$

Now, let $a_k := (16\bar{g})/(25\text{gap}\,(w(\mu_k))^2)$ for all $k = 0, \ldots, l-1$. Then, the first relation above is equivalent to the condition that $a_{k+1} \leq a_k^2$ for $l = 0, \ldots, l-2$. This implies that $a_k \le (a_0)^{2^k} \le a_0^{2k}$ for all $k = 0, \ldots, l-1$. Moreover, the last relation above implies that

$$\int_{\mu_{k+1}}^{\mu_k} \frac{\kappa(\nu)}{\nu}\, d\nu \le \frac{4}{5}\left(\frac{25a_k}{16}\right)^{1/2} \le \sqrt{a_k} \le a_0^k \le \left(\frac{16}{25}\right)^k, \quad k = 0, \ldots, l-1.$$

Hence, we obtain

$$\int_{\nu_1}^{\nu_0} \frac{\kappa(\nu)}{\nu}\, d\nu = \sum_{k=0}^{l-1} \int_{\mu_{k+1}}^{\mu_k} \frac{\kappa(\nu)}{\nu}\, d\nu \le \sum_{k=0}^{l-1} \left(\frac{16}{25}\right)^k \le \frac{25}{9}.$$

□

Finally, we are ready to prove the main theorem.

Theorem 5.10 *There holds*

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

Proof It suffices to show that for every $\nu_0 > 0$, we have

$$\int_0^{\nu_0} \frac{\kappa(\nu)}{\nu}\, d\nu = \mathcal{O}\left(n^{3.5} \log(\bar{\chi}_A + n)\right), \tag{92}$$

where the constant of proportionality in the $\mathcal{O}(\cdot)$ does not depend on ν_0. Indeed, define $\bar{g} := 384\, n\, \bar{\chi}_A$ and let $\nu_0 > 0$ be given. Let $\Gamma(\nu_0)$ denote the set of scalars $\nu \in (0, \nu_0]$ such that at least one of the following conditions hold:

C1 gap $(\nu) \le \bar{g}$;
C2 gap $(\nu) > \bar{g}$ and $\varepsilon_\infty^{\mathrm{a}}(\nu) \ge 24\sqrt{n}\,\bar{\chi}_A/\,\mathrm{gap}\,(\nu))$;
C3 gap $(\nu) > \bar{g}$, $\varepsilon_\infty^{\mathrm{a}}(\nu) \le \bar{g}/(16\sqrt{n}\,\mathrm{gap}\,(\nu))$ and there exist $i \in N(\nu)$ and $j \in B(\nu)$ such that gap $(\nu) = x_i(\nu)/x_j(\nu)$.

We claim that there exist $l \le n(n-1)/2$ intervals $[d_k, e_k]$, $k = 1, \ldots, l$, such that: i) $d_k \ge e_{k+1} > 0$ for all $k = 1, \ldots, l-1$; ii) $\Gamma(\nu_0) \subseteq \cup_{k=1}^{l}[d_k, e_k]$; and, iii) for all $k = 1, \ldots, l$, we have

$$\int_{d_k}^{e_k} \frac{\kappa(\nu)}{\nu}\, d\nu = \mathcal{O}\left(n^{1.5} \log(\bar{\chi}_A + n)\right), \tag{93}$$

where the constant of proportionality in the $\mathcal{O}(\cdot)$ does not depend on ν_0. Indeed, the intervals $[d_k, e_k]$ can be constructed inductively as follows. Suppose that $q \ge 0$

intervals $[d_k, e_k]$, $k = 1, \ldots, q$, have been constructed so that: a) properties i) and iii) hold with l replaced by q; b) $\{\nu \geq d_q : \nu \in \Gamma(\nu_0)\} \subseteq \cup_{k=1}^{q}[d_k, e_k]$, and; c) each interval $(d_k, e_k]$, $k = 1, \ldots, q$, contains a $\bar{g}^n$-crossover event. Note that, by Proposition 3.1, the latter property implies that q can not exceed $n(n-1)/2$. If the set $\{\nu < d_q : \nu \in \Gamma(\nu_0)\}$ is empty then property ii) obviously hold with $l = q$, and the conclusion of the theorem holds with $l = q$. Otherwise, let $e_{q+1} := \sup\{\nu \in (0, d_q) : \nu \in \Gamma(\nu_0)\}$. It is easy to see that $e_{q+1} \in \Gamma(\nu_0)$. Applying one of the Lemmas 5.1, 5.3 or 5.7 depending on which of the conditions C1, C2 or C3 holds and noting that $\bar{g} \geq \max\{4\, n\, \bar{\chi}_A\,,\ 48\sqrt{n}\ \bar{\chi}_A\}$, we conclude the existence of a scalar $d_{q+1} < e_{q+1}$ such that (93) holds with $k = q + 1$ and the interval $(d_{q+1}, e_{q+1}]$ contains a $\bar{g}^n$-crossover event. Clearly, the intervals $[d_k, e_k]$, $k = 1, \ldots, q + 1$, satisfy the above requirements a), b) and c) with q replaced by $q + 1$. Since the number of intervals $[d_k, e_k]$ satisfying a), b) and c) above can not exceed $n(n-1)/2$, it is clear that the above construction eventually yields intervals $[d_k, e_k]$'s according to the above claim.

Using the fact that $\bar{g} = 384\, n\, \bar{\chi}_A$, we easily see that the set $(0, \nu_0) \setminus \cup_{k=1}^{l}[d_k, e_k]$ consists of at most $l + 2$ open intervals, each one satisfying the assumptions of Lemma 5.9. Hence, it follows from Lemma 5.9 that the integral of the function $\kappa(\nu)/\nu$ over each one of these intervals is bounded by 25/9. Putting all the conclusions obtained above together, we easily see that (92) holds. □

6 Concluding remarks

In this paper, we have studied the geometric structure of the central path for linear programming based on the curvature $\kappa(\nu)$ and its corresponding integral $I(\nu_f, \nu_i)$. We have provided a link between two iteration-complexity bounds for the MTY P-C algorithm. There remains several interesting topics for future research.

One topic is to investigate whether the results of this paper can be extended to other classes of convex programming such as convex quadratic programming (or, more generally, monotone linear complementarity problems) and symmetric cone programming. In this regard, the following questions arise: i) how can the curvature integral be defined in the context of symmetric cone programming?; ii) how can an analogous bound for the curvature integral be derived in the context of convex quadratic programming?

Another topic would be to investigate whether the curvature defined in this paper is related to the concept of the curvature of a path in differential or Riemannian geometry.

Acknowledgment Some of the results of this work were obtained while the authors were visiting the Oberwolfach Mathematics Institute on the occasion of the workshop on Optimization and Applications on January of 2005. The authors are grateful to the organizers for inviting them to attend the meeting and to the Oberwolfach Mathematics Institute for the support and congenial scientific atmosphere provided.

Appendix

The objective of this section is to provide a proof of Lemma 4.3.

Proposition 6.1 *Let $F_i \in \Re^{m\times n_i}$, $h_i \in \Re^{n_i}$, $z_i \in \Re^{n_i}_{++}$, for $i = 1, 2$, be given. Consider the projections $p^0 = (p_1^0, p_2^0)$ and $\tilde{p}^0 = (\tilde{p}_1^0, \tilde{p}_2^0)$ given by*

$$(p_1^0, p_2^0) := \operatorname{argmin}_{(p_1,p_2)}\{\|p_1-h_1\|^2+\|p_2-h_2\|^2 : F_1Z_1p_1+F_2Z_2p_2=0\}, \tag{94}$$

$$\tilde{p}_1^0 := \operatorname{argmin}_{p_1}\{\, \|p_1 - h_1\|^2 : F_1Z_1p_1 \in \operatorname{Im}(F_2)\,\}, \tag{95}$$

$$\tilde{p}_2^0 := \operatorname{argmin}_{p_2}\{\, \|p_2 - h_2\|^2 : F_2Z_2p_2 = 0\,\}, \tag{96}$$

where $Z_1 := \operatorname{Diag}(z_1)$ and $Z_2 := \operatorname{Diag}(z_2)$. Then:

$$\|p_1^0 - \tilde{p}_1^0\| \le 2\bar{\chi}_F\Delta(1 + \bar{\chi}_F\Delta)\|h\|, \quad \|p_2^0 - \tilde{p}_2^0\| \le \bar{\chi}_F\Delta\|h\|, \tag{97}$$

where $\Delta = \Delta(z_1, z_2) := \|z_1\|_\infty\|(z_2)^{-1}\|_\infty$ and $F := [F_1, F_2]$.

Before giving the proof of the above proposition, we present two technical results. The first result is an error bound result for a system of linear equalities whose proof can be found for example in [9].

Lemma 6.2 *Let $A \in \Re^{m\times n}$ with full row rank be given and let $(\mathcal{K}, \mathcal{L})$ be an arbitrary bipartition of the index set $\{1, \dots, n\}$. Assume that $\bar{w} \in \Re^{|\mathcal{L}|}$ is an arbitrary vector such that the system $A_\mathcal{K}u = A_\mathcal{L}\bar{w}$ is feasible. Then, this system has a feasible solution $\bar{u}$ such that $\|\bar{u}\| \le \bar{\chi}_A\|\bar{w}\|$.*

The next result characterizes the displacements $\delta_1^0 := p_1^0 - \tilde{p}_1^0$ and $\delta_2^0 := p_2^0 - \tilde{p}_2^0$ as optimal solutions of certain optimization problems.

Lemma 6.3 *Let F_i and Z_i, $i = 1, 2$, be as in Proposition 6.1. Then, the following statements hold*:

a) *The vector $\delta_2^0 := p_2^0 - \tilde{p}_2^0$ is the unique optimal solution of the problem*

$$\begin{aligned} &\textit{minimize}_{\delta_2} \ \tfrac{1}{2}\|\delta_2\|^2 \\ &\textit{subject to} \quad F_2Z_2\delta_2 = -F_1Z_1p_1^0; \end{aligned} \tag{98}$$

b) *The pair (δ_1^0, p_2^0), where $\delta_1^0 := p_1^0 - \tilde{p}_1^0$, is the unique optimal solution of the problem*

$$\begin{aligned} &\textit{minimize}_{(\delta_1,p_2)} \ \tfrac{1}{2}\|\delta_1\|^2 + \tfrac{1}{2}\|p_2 - h_2\|^2 \\ &\textit{subject to} \qquad F_1Z_1\delta_1 + F_2Z_2p_2 = -F_1Z_1\tilde{p}_1^0. \end{aligned} \tag{99}$$

Proof We first show a). Since p^0 and $\tilde{p}_2^0$ are optimal solutions of (94) and (96), respectively, we have

$$\begin{pmatrix} p_1^0 - h_1 \\ p_2^0 - h_2 \end{pmatrix} \in \operatorname{Im}\begin{pmatrix} Z_1F_1^T \\ Z_2F_2^T \end{pmatrix}, \quad F_1Z_1p_1^0 + F_2Z_2p_2^0 = 0 \tag{100}$$

$$\tilde{p}_2^0 - h_2 \in \operatorname{Im}(Z_2F_2^T), \quad F_2Z_2\tilde{p}_2^0 = 0, \tag{101}$$

and hence

$$p_2^0 - \tilde{p}_2^0 \in \mathrm{Im}(Z_2 F_2^T), \quad F_2 Z_2 \delta_2^0 = -F_1 Z_1 p_1^0. \tag{102}$$

This shows that $\delta_2^0 = p_2^0 - \tilde{p}_2^0$ satisfies the optimality conditions for problem (98). Since (98) is a strictly convex quadratic program, its optimal solution is unique and hence a) follows. We next show b). Since $\tilde{p}_1^0$ is the optimal solution of (95), we have

$$\begin{pmatrix} \tilde{p}_1^0 - h_1 \\ 0 \end{pmatrix} \in \mathrm{Im}(ZF^T), \quad F_1 Z_1 \tilde{p}_1^0 \in \mathrm{Im}(F_2) \tag{103}$$

which, together with (100) and the definition of δ_1^0, yields

$$\begin{pmatrix} \delta_1^0 \\ p_2^0 - h_2 \end{pmatrix} \in \mathrm{Im}(ZF^T), \quad F_1 Z_1 \delta_1^0 + F_2 Z_2 p_2^0 = -F_1 Z_1 \tilde{p}_1^0. \tag{104}$$

This shows that (δ_1^0, p_2^0) satisfies the optimality conditions for (99). Since (99) is a strictly convex quadratic program, its optimal solution is unique and hence b) holds. □

Using the above lemma, we can now prove Proposition 6.1.

Proof of Proposition 6.1 By Lemma 6.2, there exists $u_2 \in \Re^{n_2}$ such that $F_2 u_2 = -F_1 Z_1 p_1^0$, or equivalently $Z_2^{-1} u_2$ is feasible for (98), and

$$\|u_2\| \le \bar{\chi}_F \|Z_1 p_1^0\| \le \bar{\chi}_F \|z_1\|_\infty \|p_1^0\|. \tag{105}$$

Hence, in view of Lemma 6.3(a) and the definition of Δ, we have

$$\|\delta_2^0\| \le \|Z_2^{-1} u_2\| \le \|z_2^{-1}\|_\infty \|u_2\| \le \bar{\chi}_F \Delta \|p_1^0\|,$$

from which the second inequality of (97) follows in view of the fact that $\|p_1^0\| \le \|h\|$. By (100) and (103), we have that $F_1 Z_1 p_1^0 \in \mathrm{Im}(F_2)$. Hence, by Lemma 6.2, there exists a vector v_2^0 such that

$$F_2 v_2^0 = F_1 Z_1 \delta_1^0, \qquad \|v_2^0\| \le \bar{\chi}_F \|Z_1 \delta_1^0\| \le \bar{\chi}_F \|z_1\|_\infty \|\delta_1^0\|. \tag{106}$$

Relation (104) and (106) imply that $F_2[Z_2 p_2^0 + v_2^0] = -F_1 Z_1 \tilde{p}_1^0$, and hence that the pair $(0,\ p_2^0 + Z_2^{-1} v_2^0)$ is feasible for (99). This together with Lemma 6.3(b) imply that

$$\|p_2^0 - h_2\|^2 + \|\delta_1^0\|^2 \le \|p_2^0 + Z_2^{-1} v_2^0 - h_2\|^2.$$

Rearranging this expression and using relation (106), the definition of Δ, the facts that $\max\{\|p_i^0\|, \|h_i - p_i^0\|\} \le \max\{\|p^0\|, \|h - p^0\|\} \le \|h\|$ and $\|\tilde{p}_i^0\| \le \|h_i\| \le \|h\|$ for

$i = 1, 2$ and the inequality $\|r\|^2 - \|u\|^2 \leq \|r - u\| \, \|r + u\|$ for any $r, u \in \Re^p$, we obtain

$$\begin{aligned}
\|\delta_1^0\| &\leq \|\delta_1^0\|^{-1} \left(\|p_2^0 + Z_2^{-1} v_2^0 - h_2\|^2 - \|p_2^0 - h_2\|^2 \right) \\
&\leq \|\delta_1^0\|^{-1} \|Z_2^{-1} v_2^0\| \, \| 2\,(p_2^0 - h_2) + Z_2^{-1} v_2^0 \,\| \\
&\leq \|\delta_1^0\|^{-1} \|v_2^0\| \, \|z_2^{-1}\|_\infty \left(2 \, \|p_2^0 - h_2\| + \|v_2^0\| \, \|z_2^{-1}\|_\infty \right) \\
&\leq \bar{\chi}_F \, \Delta \, \left\{ 2 \, \|h\| + \bar{\chi}_F \, \Delta \, \|\delta_1^0\| \right\} \\
&\leq \bar{\chi}_F \, \Delta \, \left\{ 2 \, \|h\| + \bar{\chi}_F \, \Delta \, (\|p_1^0\| + \|\tilde{p}_1^0\|) \right\} \\
&\leq 2\bar{\chi}_F \, \Delta \, (1 + \bar{\chi}_F \, \Delta) \, \|h\|,
\end{aligned}$$

which shows that the first inequality of (97) also holds.

We are now ready to give the proof of Lemma 4.3.

Proof of Lemma 4.3 Fix $k \in \{1, \ldots, l\}$ and define

$$\begin{aligned}
&(q_{J_1}^0, \ldots, q_{J_k}^0) \\
&\quad := \operatorname{argmin}_{(q_{J_1}, \ldots, q_{J_k}) \in \Re^{n_1} \times \ldots \times \Re^{n_k}} \left\{ \sum_{i=1}^{k} \|h_{J_i} - q_{J_i}\|^2 : \sum_{i=1}^{k} A_{J_i} Z_{J_i} q_{J_i} = 0 \right\}.
\end{aligned}$$

By Proposition 6.1 with $F_1 = [A_{J_{k+1}}, \ldots, A_{J_l}]$ and $F_2 = [A_{J_1}, \ldots, A_{J_k}]$, we conclude that

$$\|q_{J_k}^0 - p_{J_k}^0\| \leq \|(q_{J_1}^0 - p_{J_1}^0, \ldots, q_{J_k}^0 - p_{J_k}^0)\| \leq \bar{\chi}_A \frac{\max(z_{J_{k+1}}, \ldots, z_{J_l})}{\min(z_{J_1}, \ldots, z_{J_k})} \|h\| \leq K \|h\|.$$

Moreover, Proposition 6.1 with $F_1 = A_{J_k}$ and $F_2 = [A_{J_1}, \ldots, A_{J_{k-1}}]$ implies that

$$\|q_{J_k}^0 - \tilde{p}_{J_k}^0\| \leq 2K(1 + K) \|(h_{J_1}, \ldots, h_{J_k})\| \leq 2K(1 + K)\|h\|.$$

Combining the two previous inequality and using the triangle inequality for norms, we obtain (46).

References

1. Dikin, I.I.: On the speed of an iterative process. Upravlyaemye Sistemi **12**, 54–60 (1974)
2. Forsgren, A.: On linear least-squares problems with diagonally dominant weight matrices. SIAM J. Matrix Anal. Appl. **17**, 763–788 (1996)
3. Gonzaga, C.C.: Path-following methods for linear programming. SIAM Rev. **34**, 167–224 (1992)
4. Gonzaga, C.C., Lara, H.J.: A note on properties of condition numbers. Linear Algebra Appl. **261**, 269–273 (1997)
5. Karmarkar, N.: Riemannian geometry underlying interior-point methods for linear programming. Mathematical Developments Arising from Linear Programming, In: Brunswick, M.E., (ed), pp. 51–75. Contemporary Mathematics, Vol. 114, American Mathematical Society, Providence, 1990

6. Megiddo, N., Mizuno, S., Tsuchiya, T.: A modified layered-step interior-point algorithm for linear programming. Math. Program. **82**, 339–355 (1998)
7. Mizuno, S., Todd, M.J., Ye, Y.: On adaptive-step primal-dual interior-point algorithms for linear programming. Math. Oper. Res. **18**, 964–981 (1993)
8. Monteiro, R.D.C., Tsuchiya, T.: Global convergence of the affine scaling algorithm for convex quadratic programming. SIAM J. Optim. **8**(1), 26–58 (1998)
9. Monteiro, R.D.C., Tsuchiya, T.: A variant of the Vavasis-Ye layered-step interior-point algorithm for linear programming. SIAM J. Optim. **13**, 1054–1079 (2003)
10. Monteiro, R.D.C., Tsuchiya, T.: A new iteration-complexity bound for the MTY predictor-corrector algorithm. SIAM J. Optim. **15**, 319–347 (2004)
11. Sonnevend, G., Stoer, J., Zhao, G.: On the complexity of following the central path of linear programs by linear extrapolation. II. Math. Program. **52**, 527–553 (1991)
12. Stewart, G.W.: On scaled projections and pseudoinverses. Linear Algebra Appl. **112**, 189–193 (1989)
13. Todd, M.J.: A Dantzig-Wolfe-like variant of Karmarkar's interior-point linear programming algorithm. Oper. Res. **38**, 1006–1018 (1990)
14. Todd, M.J., Tunçel, L., Ye, Y.: Characterizations, bounds, and probabilistic analysis of two complexity measures for linear programming problems. Math. Program. **90**(1), 59–70 (2001)
15. Tsuchiya, T.: Global convergence property of the affine scaling method for primal degenerate linear programming problems. Math. Oper. Res. **17**, 527–557 (1992)
16. Tunçel, L.: Approximating the complexity measure of Vavasis-Ye algorithm is NP-hard. Math. Program. **86**(1), 219–223 (1999)
17. Tunçel, L.: On the condition numbers for polyhedra in Karmarkar's form. Oper. Res. Lett. **24**(4), 149–155 (1999)
18. Vanderbei, R.J., Lagarias, J.C.: I. I. Dikin's convergence result for the affine-scaling algorithm, Vol 17. Contemp. Math. **114**, 109–119 (1990)
19. Vavasis, S., Ye, Y.: A primal-dual accelerated interior point method whose running time depends only on A. Math. Program. **74**, 79–120 (1996)
20. Vavasis, S., Ye, Y.: On the relationship between layered least squares and affine scaling steps. In: The Mathematics of Numerical Analysis, pp. 857–865. Lectures in Applied Mathematics, Vol. 32. American Mathematical Society, Providence, 1996
21. Zhao, G.: On the relationship between the curvature integral and the complexity of path-following methods in linear programming. SIAM J. Optim. **6**, 57–73 (1996)
22. Zhao, G., Stoer, J.: Estimating the complexity of a class of path-following methods for solving linear programs by curvature integrals. Appl. Math. Optim. **27**, 85–103 (1993)