

A new perspective on the complexity of interior point methods for linear programming

Coralia Cartis^{*,‡} and Raphael A. Hauser^{†,‡}

March 15, 2007

Abstract

In a dynamical systems paradigm, many optimization algorithms are equivalent to applying forward Euler method to the system of ordinary differential equations defined by the vector field of the search directions. Thus the stiffness of such vector fields will play an essential role in the complexity of these methods. We first exemplify this point with a theoretical result for general linesearch methods for unconstrained optimization, which we further employ to investigating the complexity of a primal short-step path-following interior point method for linear programming. Our analysis involves showing that the Newton vector field associated to the primal logarithmic barrier is nonstiff in a sufficiently small and shrinking neighbourhood of its minimizer. Thus, by confining the iterates to these neighbourhoods of the primal central path, our algorithm has a nonstiff vector field of search directions, and we can give a worst-case bound on its iteration complexity. Furthermore, due to the generality of our vector field setting, we can perform a similar (global) iteration complexity analysis when the Newton direction of the interior point method is computed only approximately, using some direct method for solving linear systems of equations.

1 Introduction

The Nesterov–Nemirovskii self-concordant barriers theory constructs a class of functions whose associated Newton vector fields can be used to solve LP problems in polynomial time. We aim to introduce in what follows a minimal set of conditions that a parametric family of vector fields needs to satisfy in order to ensure that the complexity of the resulting methods can be estimated for LP. Our approach opens the possibility that other vector fields (search directions), besides Newton, can

^{*}Computational Science and Engineering Department, Rutherford Appleton Laboratory, Chilton, Oxfordshire, OX11 0QX, United Kingdom. Email: c.cartis@rl.ac.uk

[†]Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford, Oxfordshire, OX1 3QD, United Kingdom. Email: hauser@comlab.ox.ac.uk

[‡]This work was supported by the EPSRC grant GR/S34472.

be employed in interior point methods, and could make LP solvable in polynomial time. We give some illustrative examples of vector fields that satisfy these minimal conditions. We show that the Newton vector field of the logarithmic barriers functionals associated to a given LP falls into this category. Then we show that an approximate Newton direction, where the approximation comes from inexact arithmetic computations of the step, also satisfies these minimal set of conditions.

The reasoning behind the particular choice of minimal conditions for vector fields springs from stability considerations for dynamical systems. Indeed, in a dynamical systems paradigm, many optimization algorithms are equivalent to applying forward Euler method (with variable stepsize) to the system of ordinary differential equations defined by the vector field of the search directions. Since forward Euler is not A-stable, the stiffness of such vector fields will play an essential role in the complexity of these methods. Thus well-conditioned and non-stiff vector fields should be the focus of our attention. As we shall see, the Newton vector field of the logarithmic barrier is perfectly well-conditioned in a sufficiently small neighbourhood of its minimizer on the central path.

In confining the polynomial complexity results mostly to algorithms employing the Newton vector field, an implicit assumption has been created, that it is the Q-quadratic convergence properties that this vector field has that are in part responsible for this complexity. Our results show this not to be the case, in the sense that it is enough that the search direction vector field possesses linear convergence to ensure polynomial complexity of the algorithm.

Notations. Throughout, let $\|\cdot\|$ denote the Euclidean norm on \mathbb{R}^n ; the same notation is used for the operator norm induced by the Euclidean norm. Also, vector components will be denoted by subscripts, and iteration numbers, by superscripts. Furthermore, I is the $n \times n$ identity matrix and e , the vector of all ones where its dimension can be deduced from the context. Given a vector, say x , the diagonal matrix having the components of x as entries will be denoted by X .

2 Some useful preliminary results

Let $\beta \in (0, 1)$ and

$$\mathcal{N} := \{x \in \mathbb{R}^n : \|x - x^\dagger\| < R\}, \quad (2.1)$$

for some $R > 0$ and $x^\dagger \in \mathbb{R}^n$. Letting $\overline{\mathcal{N}}$ denote the closure of \mathcal{N} , we assume $v : \overline{\mathcal{N}} \rightarrow \mathbb{R}^n$, $x \mapsto v(x)$ is a vector field such that

$$\text{i) } v(x) = 0 \Leftrightarrow x = x^\dagger.$$

The following two results are essential to the material in this paper.

Theorem 2.1 *Let $v : \overline{\mathcal{N}} \rightarrow \mathbb{R}^n$, $x \mapsto v(x)$ be a vector field such that property i) above holds, and also that can be expressed as*

$$v(x) = r(x) + w(x), \quad \text{for all } x \in \mathcal{N}, \quad (2.2)$$

where $r : \overline{\mathcal{N}} \rightarrow \mathbb{R}^n$, $x \mapsto r(x)$ is a radial vector field with unique stable attractor x^\dagger , i. e.,

$$r(x) = x^\dagger - x, \quad x \in \overline{\mathcal{N}}, \quad (2.3)$$

and $w : \mathcal{N} \rightarrow \mathbb{R}^n$, $x \mapsto w(x)$ is a vector field that is β -Lipschitz continuous at x^\dagger , i. e.,

$$\|w(x)\| \leq \beta \|x - x^\dagger\|, \quad x \in \mathcal{N}, \quad (2.4)$$

where we have employed that $w(x^\dagger) = 0$ which follows from i), (2.2) and (2.3).

We consider the iterative process

$$x^{l+1} = x^l + v(x^l), \quad l \geq 0, \quad (2.5)$$

where x^0 is an arbitrary starting point in \mathcal{N} . Then

$$\|x^{l+1} - x^\dagger\| \leq \beta \|x^l - x^\dagger\|, \quad l \geq 0, \quad (2.6)$$

which provides

$$x^0 \in \mathcal{N} \implies x^l \in \mathcal{N}, \quad l \geq 0, \quad (2.7)$$

and

$$x^l \rightarrow x^\dagger, \quad \text{as } l \rightarrow \infty, \quad Q\text{-linearly with convergence factor } \beta. \quad (2.8)$$

Furthermore, we have

$$\|v(x^l)\| \leq (1 + \beta) \|x^l - x^\dagger\|, \quad l \geq 0. \quad (2.9)$$

Thus

$$v(x^l) \rightarrow 0, \quad \text{as } l \rightarrow \infty, \quad R\text{-linearly with convergence factor } \beta. \quad (2.10)$$

Proof. It follows from (2.2) and (2.3) that

$$v(x) = x^\dagger - x + w(x) \quad \text{and} \quad x + v(x) - x^\dagger = w(x), \quad x \in \mathcal{N}, \quad (2.11)$$

which together with (2.4), implies

$$\|v(x)\| \leq (1 + \beta) \|x - x^\dagger\|, \quad x \in \mathcal{N}, \quad (2.12)$$

and

$$\|x + v(x) - x^\dagger\| \leq \beta \|x - x^\dagger\|, \quad x \in \mathcal{N}. \quad (2.13)$$

Now, (2.13) and (2.5) give (2.6). Also, (2.7) results from (2.6) and $\beta \in (0, 1)$.

Straightforwardly, (2.9) follows from (2.12) and (2.5). Relations (2.6) and (2.9) give (2.10), which completes the proof. \square

The next corollary gives an example of a class of vector fields that satisfies the conditions of Theorem 2.1.

Corollary 2.2 *Let $v : \overline{\mathcal{N}} \rightarrow \mathbb{R}^n$, $x \mapsto v(x)$ be a C^1 vector field such that property *i*) above holds, and also*

$$ii) \quad \|I + Dv(x)\| \leq \beta, \text{ for all } x \in \mathcal{N}.$$

Then Theorem 2.1 applies.

Proof. For any $x \in \mathcal{N}$, property *i*) provides

$$v(x) = \int_0^1 Dv(x^\dagger + t(x - x^\dagger))(x - x^\dagger) \quad (2.14)$$

$$= x^\dagger - x + \int_0^1 [I + Dv(x^\dagger + t(x - x^\dagger))(x - x^\dagger)], \quad (2.15)$$

and further, from *ii*) and \mathcal{N} being convex,

$$\|x + v(x) - x^\dagger\| \leq \int_0^1 \|I + Dv(x^\dagger + t(x - x^\dagger))\| \cdot \|x - x^\dagger\|, \quad (2.16)$$

$$\leq \beta \|x - x^\dagger\|. \quad (2.17)$$

Thus letting $r(x) := x^\dagger - x$ and $w(x) := v(x) - r(x)$, for any $x \in \mathcal{N}$, (2.17) provides that w is β -Lipschitz continuous vector field on \mathcal{N} , which further implies that the conditions of Theorem 2.1 are satisfied. \square

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a C^3 strictly convex function, and let $n : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the associated Newton vector field. Then, conforming to [12], at the minimizer x^\dagger of f , we have

$$n(x^\dagger) = 0 \quad \text{and} \quad Dn(x^\dagger) = -I. \quad (2.18)$$

Thus the conditions of Corollary 2.2 are satisfied by the Newton vector field in a sufficiently small neighbourhood of x^\dagger . Determining the size of this neighbourhood in the specific case of the Newton vector field of the logarithmic barrier function for linear programming will be the focus of a significant part of the analysis in this paper.

2.1 On parametrized families of vector fields

Let $\mathcal{U} \subseteq \mathbb{R}^n$ be an open and convex domain. Let $\mu > 0$ and

$$\mathcal{N}(x(\mu)) := \{x \in \mathbb{R}^n : \|x - x(\mu)\| < \rho\mu\}, \quad (2.19)$$

where $x(\mu) \in \mathbb{R}^n$ and ρ is a positive constant independent of μ , such that

$$\overline{\mathcal{N}(x(\mu))} \subset \mathcal{U}, \text{ for each } \mu > 0. \quad (2.20)$$

Let (v_μ) , $\mu > 0$, be a directed and parametrized family of C^1 vector fields

$$v_\mu : \mathcal{U} \rightarrow \mathbb{R}^n, \quad x \mapsto v_\mu(x),$$

satisfying the following properties, for each $\mu > 0$,

1) $v_\mu(x) = 0 \Leftrightarrow x = x(\mu)$;

2) $\|x + v_\mu(x) - x(\mu)\| \leq \beta\|x - x(\mu)\|$, for all $x \in \mathcal{N}(x(\mu))$, where $\beta \in (0, 1)$ is independent of μ .

The directed family (v_μ) will be referred to as *vector fields with Linearly-Scaled Domains of Attraction* (LSDA).

Condition 2) in the definition of LSDA vector fields is equivalent to requiring that $w_\mu := v_\mu - r_\mu$, where $r_\mu := x(\mu) - x$ is a radial vector field, is β -Lipschitz continuous at $x(\mu)$ over $\mathcal{N}(x(\mu))$. Recalling Theorem 2.1, we deduce the following results concerning LSDA vector fields.

Theorem 2.3 *Let (v_μ) , $\mu > 0$, be a family of LSDA vector fields. Let $\mu > 0$ be fixed, and let $x^0 \in \mathcal{N}(x(\mu))$, where $\mathcal{N}(x(\mu))$ is defined in (2.19). Consider the iterative scheme*

$$x^{l+1} := x^l + v_\mu(x^l), \quad l \geq 0. \quad (2.21)$$

Then

$$x^l \in \mathcal{N}(x(\mu)), \quad l \geq 0. \quad (2.22)$$

Also, $x^l \rightarrow x(\mu)$ and $v_\mu(x^l) \rightarrow 0$, as $l \rightarrow \infty$, and the convergence is Q - and R -linear, respectively, with convergence factor β .

Furthermore, given $\xi \in (0, 1)$, it takes a finite number of iterations l , independent of μ , with

$$l \geq \hat{l} := \left\lceil \frac{\log \xi}{\log \beta} \right\rceil, \quad (2.23)$$

to obtain an iterate x^l such that

$$\|x^l - x(\mu)\| \leq \xi \rho \mu. \quad (2.24)$$

Proof. For each μ , the vector field v_μ satisfies the conditions of Theorem 2.1 with $\mathcal{N} := \mathcal{N}(x(\mu))$, $R := \rho\mu$ and $x^\dagger := x(\mu)$. Thus Theorem 2.1 provides $x^l \in \mathcal{N}(x(\mu))$, $l \geq 0$, and the convergence claims concerning (x^l) and $(v_\mu(x^l))$ stated above.

To give an upper bound on the number of iterations l required to generate x^l satisfying (2.24), we employ (2.6) which becomes in this case

$$\|x^{l+1} - x(\mu)\| \leq \beta\|x^l - x(\mu)\|, \quad l \geq 0. \quad (2.25)$$

It follows from (2.19) and $x^0 \in \mathcal{N}(x(\mu))$

$$\|x^l - x(\mu)\| \leq \beta^l \rho \mu, \quad l \geq 0. \quad (2.26)$$

Thus (2.24) holds provided

$$\beta^l \leq \xi, \quad (2.27)$$

which is in turn, satisfied when l achieves (2.23). \square

The next corollary presents a subclass of the LSDA family of vector fields, by analogy to Corollary 2.2.

Corollary 2.4 *Let (v_μ) , $\mu > 0$, be a family of vector fields that satisfy all the conditions in the definition of LSDA vector fields apart from condition 2), instead of which they achieve the requirement*

\mathcal{L}) $\|I + Dv_\mu(x)\| \leq \beta$, for all $x \in \mathcal{N}(x(\mu))$, where $\beta \in (0, 1)$ is a constant independent of μ .

Then (v_μ) is a family of LSDA vector fields (with the Lipschitz constant β given in \mathcal{L}), and thus Theorem 2.3 holds for v_μ .

When the iterates belong to a linear or affine subspace of \mathbb{R}^n , we work fully in that subspace by taking intersections of that subspace with the neighbourhood $\mathcal{N}(x(\mu))$, and evaluating the reduced Jacobians of v_μ . Then the above results are preserved (this will become clearer later in the paper when we analyse examples of LSDA vector fields).

3 A generic Short-Step Primal (SSP) interior point algorithm for linear programming

Let a Linear Programming (LP) problem be given in the standard form

$$\min_{x \in \mathbb{R}^n} c^\top x \quad \text{subject to} \quad Ax = b, \quad x \geq 0, \quad (\text{P})$$

where $m \leq n$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, and A is a real matrix of dimension $m \times n$. Let \mathcal{F}_P denote the primal feasible set, i. e.,

$$\mathcal{F}_P := \{x \in \mathbb{R}^n : Ax = b, \quad x \geq 0\}, \quad (3.1)$$

and \mathcal{S}_P , the set of solutions of this problem. The dual problem corresponding to the primal problem (P) is

$$\max_{(y,s) \in \mathbb{R}^m \times \mathbb{R}^n} b^\top y \quad \text{subject to} \quad A^\top y + s = c, \quad s \geq 0, \quad (\text{D})$$

and, similarly to (3.1), we let \mathcal{F}_D

$$\mathcal{F}_D := \{(y, s) \in \mathbb{R}^m \times \mathbb{R}^n : A^\top y + s = c, \quad s \geq 0\}, \quad (3.2)$$

denote the dual feasible set, and \mathcal{S}_D , the dual solution set. Moreover, we let \mathcal{F}_{PD} denote the primal-dual feasible set, i. e., $\mathcal{F}_{PD} := \mathcal{F}_P \times \mathcal{F}_D$, and \mathcal{S}_{PD} , the primal-dual solution set, i. e., $\mathcal{S}_{PD} := \mathcal{S}_P \times \mathcal{S}_D$.

We assume that there exists a primal-dual strictly feasible point $w^0 = (x^0, y^0, s^0) \in \mathcal{F}_{PD}$ that satisfies

$$Ax^0 = b, \quad A^\top y^0 + s^0 = c, \quad x^0 > 0 \quad \text{and} \quad s^0 > 0, \quad (3.3)$$

and that the matrix A has full row rank. We refer to these assumptions as the **IPM conditions**, and are standard assumptions in IPM theory [23]. Let \mathcal{F}_{PD}^0 denote the set of primal-dual strictly feasible points, and \mathcal{F}_P^0 , the set of primal strictly feasible points.

Let us now construct an interior point algorithm to solve (P).

Let (v_μ) be a directed and parametrized family of vector fields that satisfies the LSDA property 1) (see page 5), and assume that for $\mu > 0$, their unique equilibrium points $x(\mu)$ form a continuous path \mathcal{P} that converges to some $x^* \in \mathcal{S}_P$ and that has the property that for any $\mu^0 > 0$, there exists a positive constant C such that

$$\|x(\mu) - x(\mu^+)\| \leq C(\mu - \mu^+), \text{ for any } 0 < \mu^+ \leq \mu \leq \mu^0, \quad (3.4)$$

$$\|x(\mu) - x^*\| \leq C\mu, \text{ for any } \mu^0 \geq \mu > 0. \quad (3.5)$$

Preferably, C should not depend on μ^0 (see Section 3.1). The next proposition gives sufficient conditions for properties (3.4) and (3.5) to hold.

Proposition 3.1 *Let \mathcal{P} be continuously differentiable with respect to μ , for $\mu > 0$, and $x(\mu) \rightarrow x^* \in \mathcal{S}_P$, as $\mu \rightarrow 0$. Then conditions (3.4) and (3.5) are achieved if*

$$\exists \lim_{\mu \rightarrow 0} \dot{x}(\mu) := \dot{x}(0), \quad (3.6)$$

and we may let C in (3.4) and (3.5) take any value such that

$$C \geq \max_{\nu \in [0, \mu^0]} \|\dot{x}(\nu)\| := C_0. \quad (3.7)$$

Proof. Since $x(\mu) \in \mathcal{C}^1((0, \mu^0])$, we have $x(\mu) \in \mathcal{C}^1([\mu^+, \mu])$, for any $0 < \mu^+ \leq \mu \leq \mu^0$. Thus $x(\mu)$ has bounded variation on the interval $[\mu^+, \mu]$, and the inequalities hold

$$\|x(\mu) - x(\mu^+)\| \leq \int_{\mu^+}^{\mu} \|\dot{x}(\nu)\| d\nu \leq (\mu - \mu^+) \max_{\nu \in [\mu^+, \mu]} \|\dot{x}(\nu)\|. \quad (3.8)$$

Letting $\mu^+ \rightarrow 0$, and recalling (3.6) and $\dot{x}(\mu) \in \mathcal{C}((0, \mu^0])$, we further deduce

$$\|x(\mu) - x^c\| \leq \mu \max_{\nu \in [0, \mu]} \|\dot{x}(\nu)\| \leq \mu \max_{\nu \in [0, \mu^0]} \|\dot{x}(\nu)\| < +\infty, \quad (3.9)$$

and thus, we may set C to satisfy (3.7). \square

The implicit dependence of C_0 and of C in (3.7) on μ^0 and on the conditioning of the problem data can be made explicit for particular choices of \mathcal{P} (see for example, Section 3.1).

Returning to constructing an algorithm for (P), let us assume that a primal strictly feasible point $x^0 \in \mathcal{F}_P^0$ is available to start this algorithm, i. e.,

$$Ax^0 = b, \quad x^0 > 0. \quad (3.10)$$

Moreover, we require that x^0 is close to the primal components of the path \mathcal{P} . Thus there exists a positive constant ρ such that

$$\|x^0 - x(\mu^0)\| \leq \xi \rho \mu^0, \quad (3.11)$$

where $\xi \in (0, 1)$ is a constant chosen at the start of the algorithm.

A constant $\theta \in (0, 1)$ is given that we employ in defining a sequence of parameters $\mu^k > 0$, $k \geq 0$, as follows

$$\mu^{k+1} := \theta \mu^k, \quad k \geq 0. \quad (3.12)$$

Then, at the current iterate x^k , with $k \geq 0$, we let $x^{k,0} := x^k$, $\mu := \mu^{k+1}$, and form

$$x^{k,l+1} := x^{k,l} + v_\mu(x^{k,l}), \quad l \geq 0. \quad (3.13)$$

We compute a fixed number \hat{l} of such steps, where \hat{l} is independent of k and μ (see (2.23) for example), and let $x^{k+1} := x^{k,\hat{l}}$. We assume that the choice of vector fields ($v_\mu(x)$) keeps the iterates $x^{k,l}$, $k \geq 1$, $l \geq 0$, feasible with respect to the primal equality constraints, and, possibly together with the choice of parameter ρ , also ensures that $x^{k,l}$, $k \geq 0$, $l \geq 0$, is strictly positive (see Section 3.1). The tangency requirement (3.6) is essential for the latter condition to hold.

The algorithm terminates when $\mu^k \leq \epsilon$, where $\epsilon > 0$ is a tolerance set by the user at the start of the algorithm.

The above description of the algorithm can be summarized as follows.

A Short-Step Primal (SSP) IPM:

Let $\epsilon > 0$ be a tolerance parameter, and μ^0 , a positive parameter, $\xi \in (0, 1)$. Also, let $\hat{l} \in \{1, 2, \dots\}$, $\rho > 0$ and $\theta \in (0, 1)$ be given constants (to be specified below). A point x^0 is required that satisfies (3.10) and (3.11). At the current iterate x^k , $k \geq 0$, do:

Step 1: If $\mu^k \leq \epsilon$, STOP.

Step 2: Let $\mu^{k+1} := \theta \mu^k$, $x^{k,0} := x^k$.

Perform \hat{l} iterations of the scheme (3.13) with $\mu := \mu^{k+1}$, starting at $x^{k,0}$. This generates an iterate $x^{k,\hat{l}} := x^{k+1}$.

Step 3: Let $k := k + 1$. Go to Step 1. ◇

The value of \hat{l} , θ and ρ that ensure the SSP algorithm is well-defined and has low iteration complexity need to be determined. In particular, the neighbourhood (3.11) — to which the starting point x^0 of the algorithm belongs — should scale with μ^k , such that the iterates would satisfy

$$\|x^{k+1} - x(\mu^{k+1})\| \leq \xi \rho \mu^{k+1}, \quad k \geq 0. \quad (3.14)$$

In what follows, we address these issues. Firstly, we give a useful preliminary lemma.

Lemma 3.2 *Consider the path \mathcal{P} formed by the points $(x(\mu))$ that satisfy (3.4) and (3.5). Then for any $\mu > 0$, there exists $\theta_0 \in (0, 1)$, independent of μ , such that for any $\theta \in [\theta_0, 1]$, we have*

$$\|x - x(\mu)\| \leq \xi \rho \mu \implies \|x - x(\mu^+)\| \leq \rho \mu^+, \quad (3.15)$$

where $\mu^+ := \theta\mu$, $\xi \in (0, 1)$ and $\rho > 0$. In particular,

$$\theta_0 := \frac{\xi\rho + C}{\rho + C}, \quad (3.16)$$

where C is the complexity measure in (3.4) and (3.5).

Proof. The following identities follow from $\mu^+ = \theta\mu$ and (3.4)

$$\begin{aligned} \|x - x(\mu^+)\| &\leq \|x - x(\mu)\| + \|x(\mu) - x(\mu^+)\| \\ &\leq \xi\rho\mu + (\mu - \mu^+)C = \{\xi\rho + (1 - \theta)C\}\mu. \end{aligned} \quad (3.17)$$

Requiring that $\theta \in (0, 1]$ satisfies

$$\theta \geq \theta_0 \in (0, 1), \quad (3.18)$$

where θ_0 is defined in (3.16), (3.17) further provides

$$\|x - x(\mu^+)\| \leq \rho\theta\mu = \rho\mu^+, \quad (3.19)$$

which concludes the proof. \square

Let us now show that condition (3.14) is indeed sufficient for Algorithm SSP to converge and to allow an estimation of its worst-case iteration complexity.

Theorem 3.3 *Let problem (P) satisfy the IPM conditions, and let (v_μ) be a directed and parametrized family of vector fields that satisfies LSDA property 1) and also achieves (3.4) and (3.5). Apply Algorithm SSP to problem (P), and choose $\theta \in [\theta_0, 1)$, where θ_0 is defined in (3.16). Assume that (3.14) holds. Then $\mu^k \rightarrow 0$ and $x^k \rightarrow x^*$, as $k \rightarrow \infty$.*

Furthermore, by making the choice $\theta := \theta_0$, Algorithm SSP takes at most

$$\hat{k} := \left\lceil \left\{ \log \left(\frac{1 + C\rho^{-1}}{\xi + C\rho^{-1}} \right) \right\}^{-1} \log \frac{\mu^0}{\epsilon} \right\rceil \quad (3.20)$$

outer iterations to generate an iterate $x^{\hat{k}}$ satisfying $\mu^{\hat{k}} \leq \epsilon$, where C is the complexity measure that occurs in (3.4) and (3.5).

Proof. Since $\theta \in (0, 1)$, (3.12) implies $\mu^{k+1} \rightarrow 0$, as $k \rightarrow \infty$. Further, (3.14) implies $(x^k - x(\mu^k)) \rightarrow 0$, and since $x(\mu^k) \rightarrow x^*$ due to (3.5), we deduce $x^k \rightarrow x^*$, as $k \rightarrow \infty$.

Next we obtain an upper bound on the number of outer iterations required to generate an iterate with $\mu^k \leq \epsilon$. Letting $\theta := \theta_0$ in (3.12), we deduce inductively

$$\mu^k \leq \theta_0^k \mu^0, \quad k \geq 0. \quad (3.21)$$

Thus $\mu^k \leq \epsilon$ provided $k \log \theta_0 \leq \log(\epsilon/\mu^0)$. The value (3.20) of the bound on k now follows from (3.16). \square

The iteration worst-case complexity of generating x^k with $\|x^k - x^*\| \leq \epsilon$ follows from the above bound for $\mu^k \leq \epsilon$, from (3.14) and (3.5), and the inequalities

$$\|x^k - x^*\| \leq \|x^k - x(\mu^k)\| + \|x(\mu^k) - x^*\| \leq (\xi\rho + C)\mu^k. \quad (3.22)$$

Thus whenever $\mu^k \leq \epsilon$, we are guaranteed that the current major iterate x^k is within $\epsilon(\xi\rho + C)$ of the optimum solution x^* . This justifies the termination criteria of Algorithm SSP, and indicates the choice of tolerance ($\epsilon/(\xi\rho + C)$) that would be required on μ^k to ensure x^k is within ϵ distance from x^* .

The dependence on n and other problem data in the complexity bound in Theorem 3.3 is implicitly hidden in the term C/ρ , where ρ depends on the particular choice and properties of the family of vector fields (v_μ) and C represents a complexity measure of $v(\mu)$ or/and of our problem (P).

Assuming, in addition to the conditions of Lemma 3.2, that (v_μ) is a family of LSDA vector fields, the second inequality in (3.15) further provides, together with Theorem 2.3, that \hat{l} steps (see (2.23)) of the scheme (2.21) with $\mu := \mu^+$ generates a point x^+ satisfying $\|x^+ - x(\mu^+)\| \leq \xi\rho\mu^+$. This is the main argument that we employ inductively in the next theorem in order to show (3.14), which further implies, conforming to Theorem 3.3, that Algorithm SSP is convergent and we can estimate its iteration complexity. The inclusion (3.15), as well as the shrinking of the $\mathcal{O}(\mu^+)$ neighbourhood by ξ is depicted in the left-hand side plot of Figure 1.

Theorem 3.4 *Let problem (P) satisfy the IPM conditions, and let (v_μ) be a family of LSDA vector fields that also satisfies (3.4) and (3.5). Apply Algorithm SSP to problem (P), and choose $\theta \in [\theta_0, 1)$, where θ_0 is defined in (3.16). Furthermore, let*

$$\hat{l} := \left\lceil \frac{\log \xi}{\log \beta} \right\rceil, \quad (3.23)$$

which is independent of k and μ^k .

Then

$$\|x^{k,l} - x(\mu^{k+1})\| \leq \rho\mu^{k+1}, \quad k \geq 0, \quad l \geq 0, \quad (3.24)$$

and (3.14) holds for each $k \geq 0$. Thus Theorem 3.3 holds.

Proof. We will show (3.24) and (3.14) by induction on k . Clearly, (3.24) holds for $k = l = 0$ due to (3.11). The inductive argument is the same for any $k \geq 0$. Thus let us assume that (3.24) and (3.14) hold for some $k \geq 0$ and $l = 0$, i. e.,

$$\|x^k - x(\mu^k)\| \leq \xi\rho\mu^k. \quad (3.25)$$

Then, recalling the choice of θ , (3.15) in Lemma 3.2 provides

$$\|x^k - x(\mu^{k+1})\| \leq \rho\mu^{k+1}, \quad (3.26)$$

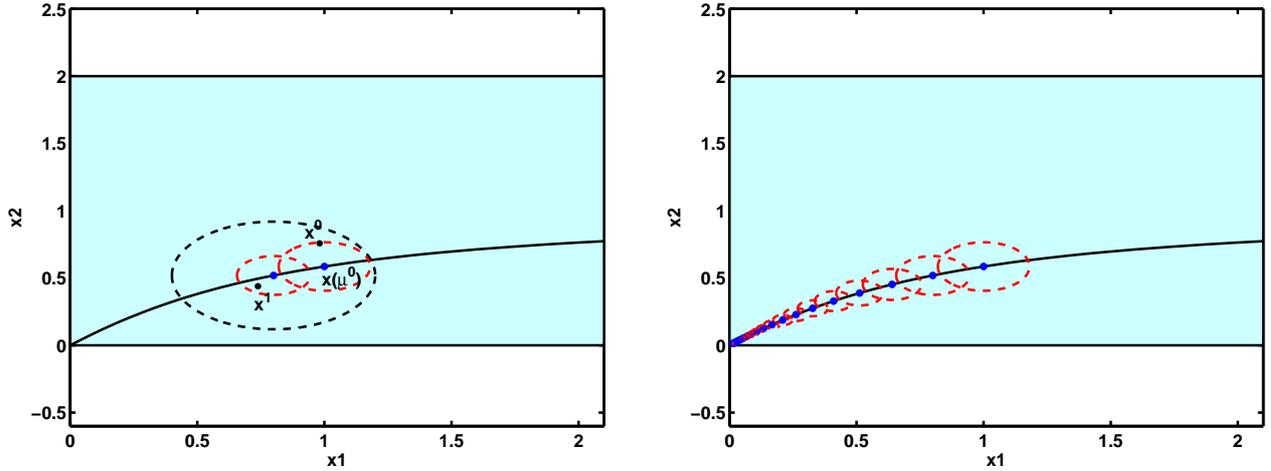


Figure 1: An iteration of Algorithm SSP. Outer iterations of Algorithm SSP.

or equivalently, $x^{k,0} := x^k \in \mathcal{N}(x(\mu^{k+1}))$. Since Theorem 2.3 applies, the inclusion (2.22) provides that (3.24) holds for all $l \geq 0$.

Furthermore, the same theorem gives, together with (3.26), that

$$\|x^{k,l} - x(\mu^{k+1})\| \leq \xi \rho \mu^{k+1}, \quad \text{for all } l \geq \hat{l}, \quad (3.27)$$

where \hat{l} is defined in (2.23) or equivalently, in (3.23). Thus $x^{k+1} := x^{k,\hat{l}}$ satisfies (3.14). \square

In the case when (v_μ) is a family of LSDA vector fields, Figure 1 illustrates the workings of a major iteration of Algorithm SSP in its left-hand side plot, while the right-hand side graph shows the shrinking ball neighbourhoods (3.11) and (3.14) as $k \rightarrow \infty$.

3.1 A choice for the path \mathcal{P}

In this section, we show that the *primal central path* associated to (P) [6, 23] may be chosen as the path \mathcal{P} above, since it satisfies properties (3.4) and (3.5). Also, there exists a range of values $(0, \rho_0)$ for ρ for this choice of \mathcal{P} that ensure that the iterates of Algorithm SSP are positive, provided the starting point x^0 is. As in this section we are only concerned with the existence of the constants C and ρ_0 .

For $\mu > 0$, consider the following strictly convex problem

$$\min_{x \in \mathbb{R}^n} f_\mu(x) := c^\top x - \mu \sum_{i=1}^n \log x_i \quad \text{subject to} \quad Ax = b, \quad x > 0, \quad (\text{P}_\mu)$$

which has a unique solution $x(\mu)$ provided the IPM conditions are satisfied. Letting $s(\mu)_i := \mu/x(\mu)_i$, $i \in \{1, \dots, n\}$, and $y(\mu) \in \mathbb{R}^m$ be the unique Lagrange multipliers of the equality constraints of (P_μ) , we obtain a pair $(y(\mu), s(\mu))$ that is strictly feasible for (D) and thus, the point $w(\mu) :=$

$(x(\mu), y(\mu), s(\mu))$ is the unique solution of the following system (of optimality conditions of (P_μ))

$$\begin{cases} Ax(\mu) = b \\ A^\top y(\mu) + s(\mu) = c \\ X(\mu)s(\mu) = \mu e, \\ x > 0 \quad \text{and} \quad s > 0. \end{cases} \quad (3.28)$$

As $\mu > 0$ varies, the points $w(\mu)$ define the *primal-dual central path* [6], contained in \mathcal{F}_{PD}^0 and continuously differentiable for $\mu > 0$. As μ tends to zero, the points $w(\mu)$, $\mu > 0$, converge to a well-defined primal-dual strictly complementary solution of (P) and (D), called the *analytic centre* of the primal-dual solution set denoted here by $w^c = (x^c, y^c, s^c)$ [6, 26]. Thus the conditions of Proposition 3.1 are satisfied provided (3.6) holds for $(x(\mu))$. To see this, as well as a similar property for $(s(\mu))$, we recall in what follows, a result from literature. concerning the limiting behaviour of the derivatives of the central path.

Let Z be an $n \times (n - m)$ matrix such that $AZ = 0$. Then the dual constraints $A^\top y + s = c$ are equivalent to $Z^\top s = Z^\top c$. Also, let us recall here that there exists a unique partition $(\mathcal{A}, \mathcal{I})$ of the index set $\{1, \dots, n\}$, where one of the sets \mathcal{A} and \mathcal{I} may be empty, such that the primal-dual solution set \mathcal{S}_{PD} can be expressed as

$$\mathcal{S}_{PD} = \{w^* = (x^*, y^*, s^*) \in \mathcal{F}_{PD} : x_{\mathcal{A}}^* = 0 \quad \text{and} \quad s_{\mathcal{I}}^* = 0\}, \quad (3.29)$$

where $x_{\mathcal{A}}^* := (x_i^* : i \in \mathcal{A})$ and $s_{\mathcal{I}}^* := (s_j^* : j \in \mathcal{I})$. We call the sets \mathcal{A} and \mathcal{I} the *strict complementarity index sets*.

Lemma 3.5 [[8], Corollary 3.6, Theorem 3.8] *Let problems (P) and (D) satisfy the IPM conditions, and $w(\mu) = (x(\mu), y(\mu), s(\mu))$, $\mu > 0$, denote the primal-dual central path. Let $\dot{w}(\mu) = (\dot{x}(\mu), \dot{y}(\mu), \dot{s}(\mu))$ denote its derivative with respect to $\mu > 0$. Then we have*

$$\lim_{\mu \rightarrow 0} \dot{x}_{\mathcal{A}}(\mu) = (S_{\mathcal{A}}^c)^{-1}e > 0 \quad \text{and} \quad \lim_{\mu \rightarrow 0} \dot{s}_{\mathcal{I}}(\mu) = (X_{\mathcal{I}}^c)^{-1}e > 0. \quad (3.30)$$

Moreover, the limit $\lim_{\mu \rightarrow 0} (\dot{x}_{\mathcal{I}}(\mu), \dot{s}_{\mathcal{A}}(\mu))$ exists and it is the unique optimal solution of the following convex quadratic problem

$$\min_{(\dot{x}_{\mathcal{I}}, \dot{s}_{\mathcal{A}})} \frac{1}{2} \|(X_{\mathcal{I}}^c)^{-1} \dot{x}_{\mathcal{I}}\|^2 + \frac{1}{2} \|(S_{\mathcal{A}}^c)^{-1} \dot{s}_{\mathcal{A}}\|^2 \quad \text{s.t.} \quad A_{\mathcal{I}} \dot{x}_{\mathcal{I}} = -A_{\mathcal{A}} (S_{\mathcal{A}}^c)^{-1} e, \quad Z_{\mathcal{A}}^\top \dot{s}_{\mathcal{A}} = -Z_{\mathcal{I}}^\top (X_{\mathcal{I}}^c)^{-1} e. \quad (3.31)$$

We remark that no nondegeneracy assumption on problems (PD) was necessary for the above Lemma to hold. Furthermore, no self-concordancy property is explicitly required either. The next corollary follows from Lemma 3.5 and Proposition 3.1.

Corollary 3.6 *Let problems (P) and (D) satisfy the IPM conditions, and $(w(\mu) = (x(\mu), y(\mu), s(\mu)))$, $\mu > 0$, denote the primal-dual central path. Given $\mu^0 > 0$, there exist a positive constant C such*

that

$$\|x(\mu) - x(\mu^+)\| \leq C(\mu - \mu^+), \quad \|s(\mu) - s(\mu^+)\| \leq C(\mu - \mu^+), \quad \text{for any } 0 < \mu^+ \leq \mu \leq \mu^0, \quad (3.32)$$

$$\|x(\mu) - x^c\| \leq C\mu \quad \text{and} \quad \|s(\mu) - s^c\| \leq C\mu, \quad \mu^0 \geq \mu > 0. \quad (3.33)$$

In particular, C satisfies

$$C \geq \max_{\nu \in [0, \mu^0]} \{\|\dot{x}(\nu)\|, \|\dot{s}(\nu)\|\} := C_0^{pd}. \quad (3.34)$$

Proof. The properties concerning $(x(\mu))$ follow straightforwardly from Lemma 3.5 and Proposition 3.1, while for $(s(\mu))$, a similar argument to the one in Proposition 3.1 may be employed together with Lemma 3.5. \square

The properties of the central path allow us to obtain a range of values for ρ that ensure the iterates $x^{k,l}$ of Algorithm SSP remain positive once the starting point x^0 is chosen as such. It follows from (3.24) that $x^{k,l} > 0$, $k \geq 0$, $l \geq 0$, provided $\rho < \min\{x_i(\mu^{k+1}) : i = \overline{1, n}\}/\mu^{k+1}$, for all $k \geq 0$. Let

$$\bar{\rho} := \sup \left\{ \rho > 0 : \frac{x_i(\mu)}{\mu} \geq \rho, \quad i = \overline{1, n}, \quad \text{for all } \mu \in (0, \mu^0] \right\}, \quad (3.35)$$

for any (fixed) $\mu^0 > 0$. [We remark that if we remove the condition that $\mu \leq \mu^0$ in (3.35), then $\bar{\rho}$ may be zero since when $\mu \rightarrow \infty$, $x_i(\mu)/\mu \rightarrow 0$ for i corresponding to bounded components of $x(\mu)$; see also Theorem 3.3 in [11].] Then, recalling $\mathcal{N}(x(\mu))$ defined in (2.19), we have

$$x \in \overline{\mathcal{N}(x(\mu))} \implies x > 0, \quad \text{for any } 0 < \mu \leq \mu^0 \text{ and } 0 < \rho < \bar{\rho}, \quad (3.36)$$

and in particular,

$$x^{k,l} > 0, \quad \text{for any } k \geq 0, l \geq 0 \text{ and } \rho \in (0, \bar{\rho}). \quad (3.37)$$

It follows from (3.33) in Corollary 3.6, as well as from the definition of the central path and of (x^c, s^c) , that

$$\frac{\mu}{C\mu + s_i^c} \leq x_i(\mu) \leq C\mu, \quad \text{and} \quad \frac{1}{C} \leq s_i(\mu) \leq C\mu + s_i^c, \quad i \in \mathcal{A}, \quad (3.38a)$$

$$\frac{1}{C} \leq x_j(\mu) \leq C\mu + x_j^c, \quad \text{and} \quad \frac{\mu}{C\mu + x_j^c} \leq s_j(\mu) \leq C\mu, \quad j \in \mathcal{I}, \quad (3.38b)$$

for all $\mu \in (0, \mu^0]$, and any fixed $\mu^0 > 0$. Thus

$$\bar{\rho} \geq \min \left\{ \frac{1}{C\mu^0}, \frac{1}{C\mu^0 + \max\{s_i^c : i \in \mathcal{A}\}} \right\} = \frac{1}{C\mu^0 + \|S_{\mathcal{A}}^c\|} := \rho_0 > 0, \quad \text{for any } \mu^0 > 0. \quad (3.39)$$

In what follows, we assume

$$\rho \in (0, \bar{\rho}), \quad (3.40)$$

and thus, (3.36) and (3.37) hold.

We remark that other choices for the path \mathcal{P} include *weighted paths*, which also have the properties in Corollary 3.6 [8, 18].

For the remainder of the paper, we present examples of LSDA vector fields $(v_\mu(x))$ that generate algorithms that are globally convergent when applied to LP problems, and whose iteration complexity we can bound using Theorem 3.4. We begin by analyzing the Newton vector field of the logarithmic barrier functions (P_μ) , $\mu > 0$.

4 A choice for the family (v_μ) of LSDA vector fields

Let $w(\mu) = (x(\mu), y(\mu), s(\mu))$, $\mu > 0$, denote the primal-dual central path of (P) (see Section 3.1). Let (n_μ) denote the Newton vector field associated to the logarithmic barrier problem (P_μ) , $\mu > 0$, whose domain of values we restrict to the set \mathcal{F}_P^0 of primal strictly feasible points, as these are the points of interest to us. At any such point x , the Newton step $n_\mu(x)$ for (P_μ) is the solution of the system

$$\begin{cases} \nabla^2 f_\mu(x)n_\mu(x) + \nabla f_\mu(x) = A^\top \lambda, \\ An_\mu(x) = 0, \end{cases} \quad (4.1)$$

or equivalently,

$$\begin{cases} \mu X^{-2}n_\mu(x) + c - \mu X^{-1}e = A^\top \lambda, \\ An_\mu(x) = 0, \end{cases} \quad (4.2)$$

where X denotes the diagonal matrix with the components of x as entries and e is the n -dimensional vector of all 1s. Further, $n_\mu(x)$ has the explicit expression

$$n_\mu(x) = -\frac{1}{\mu}\{I - X^2 A^\top (AX^2 A^\top)^{-1} A\}X^2(c - \mu X^{-1}e) \quad (4.3)$$

$$= -\frac{1}{\mu}X\{I - X A^\top (AX^2 A^\top)^{-1} AX\}(Xc - \mu e) \quad (4.4)$$

$$= -\frac{1}{\mu}X\{I - X A^\top (AX^2 A^\top)^{-1} AX\}S(\mu)(x - x(\mu)), \quad (4.5)$$

where to obtain the last identity, we employed $c = A^\top y(\mu) + s(\mu)$. It follows from (4.5) that

$$n_\mu(x(\mu)) = 0, \quad \mu > 0. \quad (4.6)$$

Furthermore, since (P_μ) is a strictly convex problem, $x(\mu)$ is the unique equilibrium point of n_μ in the set of primal strictly feasible points. Thus property 1 in the definition on page 5 of LSDA vector fields is satisfied by (n_μ) . Recalling (2.18), we remark that property (4.6) needed no further mentioning, were problem (P_μ) unconstrained.

We let Algorithm SSPN below be Algorithm SSP of the previous section with (n_μ) chosen as (v_μ) .

Algorithm SSPN:

Let $\epsilon > 0$ be a tolerance parameter, and μ^0 , a positive parameter, $\xi \in (0, 1)$. Let $\hat{l}_N \in \{1, 2, 3, \dots\}$ be a given constant to be specified later. Let $\rho \in (0, \rho_0)$, where ρ_0 is defined in (3.39), to be possibly further restricted. Let $\theta \in [\theta_0, 1)$, where θ_0 is defined in (3.16). A point x^0 is required that satisfies (3.10) and (3.11), where $x(\mu^0)$ is a point on the primal central path. At the current iterate x^k ,

$k \geq 0$, do:

Step 1: If $\mu^k \leq \epsilon$, STOP.

Step 2: Let $\mu^{k+1} := \theta\mu^k$, $x^{k,0} := x^k$.

Perform \hat{l}_N iterations of Newton's method applied to (P_μ) with $\mu := \mu^{k+1}$, starting at $x^{k,0}$. This generates an iterate $x^{k,l} := x^{k+1}$.

Step 3: Let $k := k + 1$. Go to Step 1. ◇

We remark that generic short-step primal path-following IPMs for LPs — in whose framework Algorithm SSPN broadly fits — usually compute only one Newton step for each value of μ [17, 23].

The second set of equations in (4.1) implies that

$$A[x + n_\mu(x)] = b, \quad x \in \mathcal{F}_P^0, \quad \mu > 0. \quad (4.7)$$

Since x^0 satisfies (3.10), (4.7) implies that all iterates $x^{k,l}$, $k \geq 0$, $l \geq 0$, generated by Algorithm SSPN remain feasible with respect to the primal equality constraints. Furthermore, choosing ρ in Algorithm SSPN to take values in $(0, \rho_0)$, where ρ_0 is defined in (3.39), implies, conforming to the argument at the end of Section 3.1, that $x^{k,l} > 0$, $k \geq 0$, $l \geq 0$. Thus all iterates of Algorithm SSPN are primal strictly feasible, i. e., $x^{k,l} \in \mathcal{F}_P^0$, $k \geq 0$, $l \geq 0$.

For the results of Section 3 to hold for Algorithm SSPN, which would make the latter well-defined and provide a worst-case iteration complexity bound, it remains to show that property 2 in the definition on page 5 of LSDA vector fields is satisfied by (n_μ) . This may involve further restricting the range $(0, \rho_0)$ that ρ belongs to, as we show next.

4.1 On ensuring LSDA property 2 for the Newton vector field of the log barrier

If similarly to the agreement between (4.6) and the first relation in (2.18), the second relation in (2.18) holds for the Jacobian of n_μ at $x(\mu)$, then this Jacobian would remain well-conditioned in a neighbourhood of $x(\mu)$, and we would only need to prove it is of size $\mathcal{O}(\mu)$. Thus let us firstly compute the Jacobian of $n_\mu(x)$, $x \in \mathcal{F}_P^0$.

Differentiating (4.2), we deduce

$$\begin{cases} \mu X^{-2}[Dn_\mu(x) + I] - 2\mu X^{-3}N_\mu(x) = A^\top D\lambda, \\ ADn_\mu(x) = 0, \end{cases} \quad (4.8)$$

where $N_\mu(x)$ is the diagonal matrix with the components of the vector $n_\mu(x)$ as entries. We obtain the explicit expression

$$Dn_\mu(x) + I = X^2 A^\top (AX^2 A^\top)^{-1} A + 2[I - X^2 A^\top (AX^2 A^\top)^{-1} A] X^{-1} N_\mu(x), \quad (4.9)$$

which further gives, together with (4.6),

$$Dn_\mu(x(\mu)) + I = X(\mu)^2 A^\top (AX(\mu)^2 A^\top)^{-1} A. \quad (4.10)$$

Thus, due to the presence of the primal equality constraints, the property $Dn_\mu(x(\mu)) = -I$ in (2.18) continues to hold only for directions in the null space of A , i. e.,

$$[Dn_\mu(x(\mu)) + I]d = 0, \quad \text{for } d \text{ such that } Ad = 0.$$

Moreover, considering the expression (4.10), we cannot bound it so as to ensure the requirement 2 in the definition of LSDA vector fields. Thus we introduce a change of variables so that we work in the reduced space of the points x that satisfy the primal equality constraints. We will show that the reduced Newton vector field of (P_μ) has the LSDA properties. Finally, the results in Section 3 will be applied to the corresponding “reduced” iterates and Newton vector field.

4.2 A change of variables

Since A has full row rank, the dimension of its null space $\mathcal{N}(A)$ is $n - m$, and there exists $z_i \in \mathcal{N}(A)$, $i = \overline{1, n - m}$, orthogonal vectors such that

$$\mathcal{N}(A) := \{x \in \mathbb{R}^n : Ax = 0\} = \{Zu : u \in \mathbb{R}^{n-m}\}, \quad (4.11)$$

where the $n \times (n - m)$ matrix Z has columns z_i , $i = \overline{1, n - m}$, and rows $Z_j \in \mathbb{R}^{n-m}$, $j = \overline{1, n}$. Thus we have

$$AZ = 0, \quad Z^\top Z = I, \quad \|Z\| = 1, \quad (4.12)$$

where the last two properties follow from the columns of Z being orthogonal to each other. Therefore, we can represent any vector x satisfying $Ax = b$ as

$$x = Zu + x(\mu), \quad \text{for some (unique) } u \in \mathbb{R}^{n-m}, \quad (4.13)$$

where $\mu > 0$. Thus problem (P) is equivalent to

$$\min_{u \in \mathbb{R}^{n-m}} (Z^\top c)^\top u \quad \text{subject to} \quad Zu \geq -x(\mu). \quad (4.14)$$

Its dual is

$$\max_{s \in \mathbb{R}^n} (-x(\mu))^\top s \quad \text{subject to} \quad Z^\top s = Z^\top c, \quad s \geq 0. \quad (4.15)$$

Problem (P_μ) is equivalent to

$$\min_{u \in \mathbb{R}^{n-m}} f_\mu^r(u) := c^\top Zu - \mu \sum_{i=1}^n \log(Z_i^\top u + x_i(\mu)) \quad \text{subject to} \quad Zu > -x(\mu). \quad (P_{\mu,u})$$

If the IPM conditions are satisfied by (P) and (D), then they also hold for the above reduced problems. For $\mu > 0$, the solution of $(P_{\mu,u})$ is $u(\mu) = 0$. We will now “reduce” all the quantities of interest (the Newton step, its Jacobian, etc.), to the lower dimensional space of the vectors u .

The Newton step for the (unconstrained) problem $(P_{\mu,u})$ is

$$n_{\mu}^r(u) = -[Z^{\top} \nabla^2 f_{\mu}(Zu + x(\mu))Z]^{-1} Z^{\top} \nabla f_{\mu}(Zu + x(\mu)) \quad (4.16)$$

$$= -\frac{1}{\mu} [Z^{\top} (\text{diag}(Zu + x(\mu)))^{-2} Z]^{-1} Z^{\top} (c - \mu (\text{diag}(Zu + x(\mu)))^{-1} e) \quad (4.17)$$

$$= -\frac{1}{\mu} (Z^{\top} X^{-2} Z)^{-1} Z^{\top} (c - \mu X^{-1} e), \quad \text{where } x = Zu + x(\mu), \quad (4.18)$$

$$= -\frac{1}{\mu} (Z^{\top} X^{-2} Z)^{-1} Z^{\top} (s(\mu) - \mu X^{-1} e), \quad \text{where } x = Zu + x(\mu), \quad (4.19)$$

and where to obtain the last identity, we employed $c = A^{\top} y(\mu) + s(\mu)$. The following relation connects the Newton step n_{μ} to its reduced variant n_{μ}^r , and it can be easily verified,

$$n_{\mu}(x) = Zn_{\mu}^r(u), \quad \text{where } x = Zu + x(\mu). \quad (4.20)$$

Given a strictly feasible point x^0 of (P), there exists a unique u^0 such that $x^0 = Zu^0 + x(\mu)$. Letting

$$x^{l+1} := x^l + n_{\mu}(x^l), \quad \text{and} \quad u^{l+1} := u^l + n_{\mu}^r(u^l), \quad l \geq 0, \quad (4.21)$$

then it follows from (4.20) that

$$x^l = Zu^l + x(\mu), \quad l \geq 0. \quad (4.22)$$

Thus, provided the starting points x^0 and u^0 are related by (4.22), the Newton iterates are the same in the u and x -spaces, expect for a translation.

From (4.20), we also deduce

$$n_{\mu}^r(u) = 0 \iff n_{\mu}(x) = 0 \iff x = x(\mu) \iff u = u(\mu) = 0, \quad (4.23)$$

where to obtain the second equivalence, we recalled the argument in Section 4.1. It follows from (4.23) that property 1 in the definition on page 5 of a family of LSDA vector fields is satisfied by (n_{μ}^r) . Now we proceed to the computation of the reduced Jacobian of n_{μ}^r , in order to then show that property 2 of LSDA vector fields is also achieved by (n_{μ}^r) .

To calculate $Dn_{\mu}^r(u)$, we differentiate (4.16) with respect to u and obtain

$$D[Z^{\top} \nabla^2 f_{\mu}(Zu + x(\mu))Z]n_{\mu}^r(u) + Z^{\top} \nabla^2 f_{\mu}(Zu + x(\mu))Z Dn_{\mu}^r(u) = -Z^{\top} \nabla^2 f_{\mu}(Zu + x(\mu))Z, \quad (4.24)$$

and it is easy to see that

$$Dn_{\mu}^r(u(\mu)) = -I,$$

where $u(\mu) = 0$.

Now we want to compute $D[Z^{\top} \nabla^2 f_{\mu}(Zu + x(\mu))Z]n_{\mu}^r(u)$. We have

$$Z^{\top} \nabla^2 f_{\mu}(Zu + x(\mu))Z n_{\mu}^r(u) = \left[\sum_{i=1}^n \frac{Z_i Z_i^{\top}}{(Z_i^{\top} u + x_i(\mu))^2} \right] n_{\mu}^r(u) = \sum_{i=1}^n \frac{Z_i (Z_i^{\top} n_{\mu}^r(u))}{(Z_i^{\top} u + x_i(\mu))^2}, \quad (4.25)$$

which we differentiate with respect to u , while considering $n_\mu^r(u)$ to be independent of u . We deduce

$$D[Z^\top \nabla^2 f_\mu(Zu + x(\mu))Z]n_\mu^r(u) = -2Z^\top \text{diag} \left(\frac{Z_i^\top n_\mu^r(u)}{(Z_i^\top u + x_i(\mu))^3} : i = \overline{1, n} \right) Z, \quad (4.26)$$

which further provides together with (4.24),

$$Dn_\mu^r(u) + I = 2 \left(Z^\top (\text{diag}(Zu + x(\mu)))^{-2} Z \right)^{-1} Z^\top \text{diag} \left(\frac{Z_i^\top n_\mu^r(u)}{(Z_i^\top u + x_i(\mu))^3} : i = \overline{1, n} \right) Z, \quad (4.27)$$

$$= 2 \left(Z^\top X^{-2} Z \right)^{-1} Z^\top N_\mu(x) X^{-3} Z, \text{ with } x = Zu + x(\mu), n_\mu(x) = Zn_\mu^r(u). \quad (4.28)$$

The upper bound on $\|Dn_\mu^r(u) + I\|$ that we will deduce in order to show property 2 of LSDA vector fields is satisfied by (n_μ^r) will depend on some condition numbers of problems (P) and (4.14) that we describe next.

4.2.1 Some relevant condition numbers of our problems

Conforming to [20], let

$$\chi(Z^\top) := \sup\{\|(Z^\top DZ)^{-1} Z^\top D\| : D \text{ } n \times n \text{ positive definite diagonal matrix}\} < \infty, \quad (4.29)$$

and

$$\bar{\chi}(A) := \sup\{\|A^\top (ADA^\top)^{-1} AD\| : D \text{ } n \times n \text{ positive definite diagonal matrix}\} < \infty, \quad (4.30)$$

where it is required that the matrices A and Z^\top are full row rank, condition that is satisfied in our case.

The two measures (4.29) and (4.30) are related as follows. We remark that the null space of A coincides with the range space of Z , and the range space of A^\top with the null space of Z^\top , i. e.,

$$\mathcal{N}(A) = \mathcal{R}(Z) \quad \text{and} \quad \mathcal{R}(A^\top) = \mathcal{N}(Z^\top). \quad (4.31)$$

Similarly, when we scale A by a positive definite diagonal matrix $D^{1/2}$, we have

$$\mathcal{N}(AD^{1/2}) = \mathcal{R}(D^{-1/2}Z) \quad \text{and} \quad \mathcal{R}(D^{1/2}A^\top) = \mathcal{N}(Z^\top D^{-1/2}). \quad (4.32)$$

Thus the orthogonal projection matrices into $\mathcal{N}(AD^{1/2})$ and $\mathcal{R}(D^{-1/2}Z)$

$$P_{\mathcal{N}(AD^{1/2})} := I - D^{1/2}A^\top (ADA^\top)^{-1} AD^{1/2}, \quad (4.33)$$

$$P_{\mathcal{R}(D^{-1/2}Z)} := D^{-1/2}Z(Z^\top D^{-1}Z)Z^\top D^{-1/2}. \quad (4.34)$$

represent the same linear operator (the projection operator). Moreover, we can show these matrices coincide by proving that their image coincides on all the vectors in \mathbb{R}^n , because A and Z have full

row and column rank, respectively. Thus, further multiplying (4.33) and (4.34) on the right by $Z^\top D^{1/2}$, and on the left by $D^{-1/2}$, we obtain the identity

$$(Z^\top D^{-1} Z)^{-1} Z^\top D^{-1} = Z^\top [I - DA^\top (ADA^\top)^{-1} A], \quad (4.35)$$

where we also employed $Z^\top Z = I$. Thus taking norms in (4.35), we obtain

$$\|(Z^\top D^{-1} Z)^{-1} Z^\top D^{-1}\| \leq \|Z^\top\| \cdot [1 + \|DA^\top (ADA^\top)^{-1} A\|], \quad (4.36)$$

$$\leq \|Z^\top\| \cdot [1 + \bar{\chi}(A)], \quad (4.37)$$

where the second inequality follows since transposition of matrices preserves their two-norm. Further, we employ the bound

$$\|Z^\top\| \leq \sqrt{n} \max\{\|Z_i\| : i = \overline{1, n}\} \leq \sqrt{n}, \quad (4.38)$$

and pass to the supremum over positive definite diagonal matrices on the left-hand side of (4.36), to deduce

$$\chi(Z^\top) \leq \sqrt{n}(1 + \bar{\chi}(A)). \quad (4.39)$$

The upper bound ρ_0 defined in (3.39) occurs naturally in Algorithm SSPN. It depends, however, on another condition number of our problem, $\|s_{\mathcal{A}}^c\|$, or equivalently, $\|s^c\|$. In the remainder of this subsection, we relate it to the condition number $\bar{\chi}(A)$, in the case when (P) has a unique nondegenerate solution. Let us recall that since (y^c, s^c) is a dual solution, it satisfies

$$A_{\mathcal{A}}^\top y^c + s_{\mathcal{A}}^c = c_{\mathcal{A}} \quad \text{and} \quad A_{\mathcal{I}}^\top y^c = c_{\mathcal{I}}. \quad (4.40)$$

Moreover, when (P) has a unique nondegenerate solution, the matrix $A_{\mathcal{I}}$ must be nonsingular, and we obtain the expression

$$s_{\mathcal{A}}^c = c_{\mathcal{A}} - (A_{\mathcal{I}}^{-1} A_{\mathcal{A}})^\top c_{\mathcal{I}}. \quad (4.41)$$

Now Lemma 3 in [20] implies $\|(A_{\mathcal{I}})^{-1} A_{\mathcal{A}}\| \leq \bar{\chi}(A)$, and since $\|A_{\mathcal{I}}^{-1} A_{\mathcal{A}}\| \leq \|A_{\mathcal{I}}^{-1} A\|$, we deduce the following bound

$$\|A_{\mathcal{I}}^{-1} A_{\mathcal{A}}\| \leq \bar{\chi}(A), \quad (4.42)$$

which provides, together with (4.41),

$$\|s_{\mathcal{A}}^c\| \leq \|c_{\mathcal{A}}\| + \bar{\chi}(A) \|c_{\mathcal{I}}\|, \quad (4.43)$$

or equivalently,

$$\|s_{\mathcal{A}}^c\| \leq [1 + \bar{\chi}(A)] \cdot \|c\|. \quad (4.44)$$

It follows from (3.39)

$$\rho_0 := \frac{1}{C\mu^0 + \|S_{\mathcal{A}}^c\|} \geq \frac{1}{C\mu^0 + [1 + \bar{\chi}(A)] \cdot \|c\|} := \rho_1, \quad (4.45)$$

and we may further restrict the range of admissible values for ρ in Algorithm SSPN to $(0, \rho_1)$ when (P) has a unique nondegenerate solution (see Corollary 4.2).

Now we return to computing bounds on the Jacobian of the reduced Newton direction.

4.2.2 Ensuring LSDA property 2 for the reduced Newton vector field of the log barrier

The next lemma gives the promised bound on the Jacobian of the reduced Newton vector field, implying that $(n_r(u))_\mu$ satisfies the LSDA properties.

Theorem 4.1 *Let problem (P) satisfy the IPM conditions. Let μ and μ_0 be positive arbitrary parameters, such that $0 < \mu \leq \mu^0$, and x satisfies*

$$Ax = b \quad \text{and} \quad x \in \mathcal{N}(x(\mu)), \quad (4.46)$$

where $\mathcal{N}(x(\mu))$ is defined in (2.19), and $\rho \in (0, \rho_0)$, where ρ_0 is given in (3.39). Then

$$\|Dn_\mu^r(u) + I\| \leq 2\sqrt{n}(1 + \bar{\chi}(A))\frac{\rho}{\rho_0}, \quad (4.47)$$

where u is related uniquely to x by the relation $Zu = x - x(\mu)$.

In particular, letting

$$\mathcal{N}_\mu^r(0) := \{u \in \mathbb{R}^{n-m} : \|u\| < \rho\mu\}, \quad (4.48)$$

where ρ is chosen such that

$$0 < \rho < \rho_0[2\sqrt{n}(1 + \bar{\chi}(A))]^{-1}, \quad (4.49)$$

then the reduced Newton vector fields $(n_\mu^r(u))$ satisfy the LSDA properties in the neighbourhoods $\mathcal{N}_\mu^r(0)$.

Proof. Relation (4.28) may be written equivalently

$$Dn_\mu^r(u) + I = 2[(Z^\top X^{-2}Z)^{-1}Z^\top X^{-2}]N_\mu(x)X^{-1}Z, \quad \text{with } x = Zu + x(\mu), \quad n_\mu(x) = Zn_\mu^r(u). \quad (4.50)$$

We have already established at the end of Section 3.1 that the condition (3.39) on ρ implies that x satisfying (2.19) is positive. Thus it follows from (4.29)

$$\|Dn_\mu^r(u) + I\| \leq 2\chi(Z^\top)\|Z\| \cdot \|X^{-1}n_\mu(x)\| \leq 2\chi(Z^\top)\|X^{-1}n_\mu(x)\|, \quad (4.51)$$

where in the last inequality we employed $\|Z\| = 1$. To evaluate the length of $X^{-1}n_\mu(x)$, we return to the expression (4.5), and deduce

$$X^{-1}n_\mu(x) = -\frac{1}{\mu}\{I - XA^\top(A X^2 A^\top)^{-1}AX\}S(\mu)(x - x(\mu)).$$

Recalling (4.33), $I - XA^\top(A X^2 A^\top)^{-1}AX$ is the matrix of the orthogonal projection into the null space of AX , and thus, $\|I - XA^\top(A X^2 A^\top)^{-1}AX\| \leq 1$. It follows from (2.19) and (4.46) that

$$\|X^{-1}n(x)\| \leq \rho\|S(\mu)\|. \quad (4.52)$$

The inequalities (3.38) and (3.39) finally provide

$$\|X^{-1}n(x)\| \leq \rho[C\mu^0 + \|S_A^c\|] = \frac{\rho}{\rho_0}, \quad (4.53)$$

which, together with (4.39) and (4.51), implies (4.47).

To show the second part of the theorem, recall that (4.23) implies the first condition in the definition of LSDA vector fields is satisfied by (n_μ^r) .

To prove condition 2, let $u \in \mathcal{N}_\mu^r(0)$. Then $x := Zu + x(\mu)$ satisfies $x \in \mathcal{N}(x(\mu))$, since $\|Z\| \leq 1$. Furthermore, the choice (4.49) of ρ implies $\rho < \rho_0$ since $\bar{\chi}(A) > 0$. Thus the conditions of the first part of the theorem are satisfied, providing the bound (4.47) holds. This, together with (4.49), imply $\|Dn_\mu^r(u) + I\| < 1$ and bounded away from 1. We conclude the second LSDA requirement (see page 5) is achieved by (n_μ^r) . \square

Employing (4.45), the expression of the bound in (4.47) can be uniformly described in terms of only one condition number — $\bar{\chi}(A)$ — by expressing the results in terms of ρ_1 rather than ρ_0 , as the following corollary states.

Corollary 4.2 *In the conditions of Theorem 4.1, assume additionally that (P) has a unique non-degenerate solution and that ρ in (2.19) satisfies $0 < \rho \leq \rho_1$, where $\rho_1 \in (0, \rho_0]$ is defined in (4.45). Then we have*

$$\|Dn_r(u) + I\| \leq 2\sqrt{n}(1 + \bar{\chi}(A))\frac{\rho}{\rho_1}, \quad (4.54)$$

where u is related uniquely to x by the relation $Zu = x - x(\mu)$.

In particular, choosing $\rho > 0$ in (4.48) such that

$$\rho < \frac{\rho_1}{2\sqrt{n}(1 + \bar{\chi}(A))}, \quad (4.55)$$

and letting

$$\beta := 2\sqrt{n}(1 + \bar{\chi}(A))\frac{\rho}{\rho_1}, \quad (4.56)$$

then the reduced Newton vector fields $(n_\mu^r(u))$ satisfy the LSDA conditions in the neighbourhoods $\mathcal{N}_\mu^r(0)$, with the above constants.

Proof. The proof follows from (4.45) and Theorem 4.1. \square

In the conditions of the second part of Theorem 4.1 or of Corollary 4.2, since (n_μ^r) is a family of LSDA vector fields in the neighbourhoods (4.48), Theorem 2.3 applies (directly) to (n_μ^r) and the u -iterates in (4.21). In the next theorem, we employ this result to deduce a variant of Theorem 2.3 for the x iterates and the (full) Newton vector field (n_μ) . This will provide us with a suitable value for the number of inner iterations \hat{l}_N required by Algorithm SSPN.

4.3 Determining the inner iterations \hat{l}_N required by Algorithm SSPN

Theorem 4.3 *Let problem (P) have a unique nondegenerate solution and satisfy the IPM conditions. Let μ and μ_0 be positive arbitrary parameters, such that $0 < \mu \leq \mu_0$, and let x^0 satisfy (4.46), where $0 < \rho < \rho_2$ with*

$$\rho_2 := \rho_1[2n(1 + \bar{\chi}(A))]^{-1}, \quad (4.57)$$

where ρ_1 is defined in (3.39). Consider the sequence (x^l) , $l \geq 0$, generated according to the first recurrence in (4.21).

Then $x^l \rightarrow x(\mu)$ and $(n_\mu(x^l)) \rightarrow 0$ R-linearly, as $l \rightarrow \infty$, and the convergence factor is

$$\tilde{\beta} := 2n(1 + \bar{\chi}(A))\rho/\rho_1. \quad (4.58)$$

Furthermore, given $\xi \in (0, 1)$, there exists \hat{l}_N , independent of μ , where

$$\hat{l}_N := \left\lceil \frac{\log(\xi \cdot n^{-1/2})}{\log \tilde{\beta}} \right\rceil, \quad (4.59)$$

such that

$$\|x^l - x(\mu)\| \leq \xi\rho\mu, \quad l \geq \hat{l}_N. \quad (4.60)$$

In particular, letting $\xi := 1/2$ and $\rho := \rho_2/(2\sqrt{n})$, then $\tilde{\beta} = 1/(2\sqrt{n})$ and $\hat{l}_N = 1$.

Proof. We define the following neighbourhoods in the u -space

$$\tilde{\mathcal{N}}_\mu^r(0) := \{u \in \mathfrak{R}^{n-m} : \|u\| < \tilde{\rho}\mu\}, \quad (4.61)$$

where $\tilde{\rho} := \sqrt{n}\rho$. Thus it follows from (4.57) that $\tilde{\rho}$ satisfies

$$\tilde{\rho} < \frac{\rho_1}{2\sqrt{n}(1 + \bar{\chi}(A))}, \quad (4.62)$$

which is condition (4.55). Then Corollary 4.2 applies and gives that (n_μ^r) is a LSDA family of vector fields in the neighbourhoods $\tilde{\mathcal{N}}_\mu^r(0)$.

Furthermore, for x^0 , there exists a unique u^0 such that $x^0 = Zu^0 + x(\mu)$. Thus $\|u^0\| \leq \|Z^\top\| \cdot \|x^0 - x(\mu)\| \leq \sqrt{n}\|x^0 - x(\mu)\|$. Since x^0 is assumed to satisfy (4.46), it follows that $\|u^0\| < \sqrt{n}\rho\mu := \tilde{\rho}\mu$, and $u^0 \in \tilde{\mathcal{N}}_\mu^r(0)$. Starting from this u^0 , we define recursively the sequence (u^l) by $u^{l+1} := u^l + n_r(u^l)$, $l \geq 0$, which is the second relation in (4.21). It follows from the first paragraph of this proof that the conditions of Theorem 2.3 hold for this sequence (u^l) and the LSDA vector fields (n_μ^r) . It follows that

$$u^l \rightarrow u(\mu) = 0 \text{ and } n_\mu^r(u^l) \rightarrow 0, \text{ Q- and R-linearly, respectively, as } l \rightarrow \infty, \quad (4.63)$$

and the convergence factor is $\tilde{\beta}$ defined in (4.58).

Recalling (4.21) and (4.22), we deduce

$$\|x^l - x(\mu)\| \leq \|u^l\| \leq \sqrt{n}\|x^l - x(\mu)\|, \quad (4.64)$$

and the first inequality, together with (4.63), implies $x^l \rightarrow x(\mu)$ R-linearly, with convergence factor $\tilde{\beta}$. Similarly, (4.20), (4.63) and $\|Z\| \leq 1$ imply $n_\mu(x^l) \rightarrow 0$ R-linearly, with convergence factor $\tilde{\beta}$.

Now, concerning the complexity of shrinking the neighbourhoods of $x(\mu)$, recall that $u^0 \in \tilde{\mathcal{N}}_\mu^r(0)$ implies $\|u^0\| \leq \tilde{\rho}\mu = \sqrt{n}\rho\mu$. Let

$$\tilde{\xi} := \xi/\sqrt{n} \in (0, 1). \quad (4.65)$$

Then the second part of Theorem 2.3 provides that

$$\|u^l\| \leq \tilde{\xi}\tilde{\rho}\mu = \xi\rho\mu, \quad l \geq \hat{l}, \quad (4.66)$$

where $\hat{l} := \lceil \log \tilde{\xi} / \log \tilde{\beta} \rceil$. It follows from (4.65) and the first inequality in (4.64) that $\hat{l} = \hat{l}_N$, where the latter is given in (4.59), and (4.60) holds for $l \geq \hat{l}_N$. \square

To summarize, we have now specified the choice of constants in Algorithm SSPN: the size ρ must belong to $(0, \rho_2)$, where ρ_2 is defined in (4.57) and it satisfies $\rho_2 < \rho_1 \leq \rho_0$; the number of inner iterations \hat{l}_N to be performed on each major iteration is given in (4.59). We remark that to satisfy the condition on ρ , it is sufficient to let $\rho := \beta_0\rho_2$ for some $\beta_0 \in (0, 1)$. Then, $\tilde{\beta} = \beta_0$, where $\tilde{\beta}$ is defined by (4.58). In particular, by changing ρ , $\tilde{\beta}$ can be adjusted to take any value in $(0, 1)$. This in turn, will determine \hat{l}_N (also by the choice of ξ). We remark however, that a small \hat{l}_N leads to a large number of outer iterations k (recall the value of θ_0 in (3.16) and its dependence on ρ).

Please note that the assumption that (P) has a unique nondegenerate solution can be removed if we use Theorem 4.1, instead of Corollary 4.2. Then, however, the complexity results will depend not only on $\bar{\chi}(A)$, but also on the condition number $\|s^c\|$.

4.4 The worst-case outer iteration complexity of Algorithm SSPN

Now we return to analysing the overall convergence and iteration complexity of Algorithm SSPN, when applied to (P). Theorem 4.3 implies that every outer iteration of the algorithm will perform \hat{l}_N inner iterations where \hat{l}_N is prescribed by (4.59). In the next theorem, we investigate the number of outer iterations required to obtain an approximate solution of (P). It is, in fact, a straightforward employment of the general Theorem 3.3, in the particular context of the family of Newton vector fields (n_μ) of the parametrized and constrained logarithmic barrier.

Theorem 4.4 *Let (P) satisfy the IPM conditions and assume that it has a unique nondegenerate solution x^c . Let Algorithm SSPN be applied to (P), where we perform \hat{l}_N inner Newton iterations for each μ^k , with \hat{l}_N prescribed by (4.59), and where $\rho \in (0, \rho_2)$, with ρ_2 defined in (4.57). Then*

$$\|x^k - x(\mu^k)\| \leq \xi\rho\mu^k, \quad k \geq 0, \quad (4.67)$$

and $\mu^k \rightarrow 0$, $x^k \rightarrow x^c = 0$, as $k \rightarrow \infty$.

Furthermore, by making the choice $\theta := \theta_0$, Algorithm SSPN takes at most \hat{k}_N outer iterations, where

$$\hat{k}_N := \left\lceil \frac{6n(1 + \bar{\chi}(A))C \max\{C\mu^0, (1 + \bar{\chi}(A))\|c\|\}}{\tilde{\beta}(1 - \xi)} \log \frac{\mu^0}{\epsilon} \right\rceil, \quad (4.68)$$

to generate an iterate $x^{\hat{k}_N}$ satisfying $\mu^{\hat{k}_N} \leq \epsilon$, where C is the complexity measure introduced in Lemma 3.6 and where we can view $\tilde{\beta}$ as an arbitrary constant in $(0, 1)$, but connected to ρ via $\rho = \tilde{\beta}\rho_2$ and may dependent on n (see below). Since ξ is a user-chosen parameter, it follows that

$$\hat{k}_N = \mathcal{O} \left(n(1 + \bar{\chi}(A))C\tilde{\beta}^{-1} \max\{C\mu^0, (1 + \bar{\chi}(A))\|c\|\} \log \frac{\mu^0}{\epsilon} \right). \quad (4.69)$$

Proof. The conditions of Theorem 3.3 are satisfied. This theorem provides (4.67), and the convergence properties of (μ^k) and (x^k) . Furthermore, relation (3.20) gives a value for \hat{k}_N , where now $\rho = \tilde{\beta}\rho_2$ and C occurs in Lemma 3.6. It is possible to simplify the expression of the inverse logarithm in (3.20) by returning to the expression (3.12) and (3.16) and writing $\mu^k = \theta_0^k \mu^0$ as

$$\mu^k = \left(1 - \frac{(1-\xi)\rho}{C+\rho}\right)^k \mu^0 \leq e^{-(1-\xi)\rho(C+\rho)^{-1}k} \mu^0.$$

Thus an alternative value for \hat{k}_N is $\lceil (C/\rho + 1) \log(\mu^0/\epsilon)/(1-\xi) \rceil$. Then replacing $\rho = \tilde{\beta}\rho_2$, ρ_2 from (4.57), and ρ_1 from (4.45), in this latter expression for \hat{k}_N , we deduce (4.68). \square

In the conditions of theorems 4.3 and 4.4, if $\tilde{\beta}$ is a constant in $(0, 1)$, independent of n , then these theorems provide

$$\hat{l}_N = \mathcal{O}(\log n) \quad \text{and} \quad \hat{k}_N = \mathcal{O}\left(n(1 + \bar{\chi}(A))C \max\{C\mu^0, (1 + \bar{\chi}(A))\|c\|\} \log \frac{\mu^0}{\epsilon}\right). \quad (4.70)$$

If $\tilde{\beta}$ depends on n , for example, $\tilde{\beta} := \mathcal{O}(n^p)$, with $p > 0$, then the same theorems give the estimates

$$\hat{l}_N = \mathcal{O}(1) \quad \text{and} \quad \hat{k}_N = \mathcal{O}\left(n^{1+p}(1 + \bar{\chi}(A))C \max\{C\mu^0, (1 + \bar{\chi}(A))\|c\|\} \log \frac{\mu^0}{\epsilon}\right). \quad (4.71)$$

5 Conclusions

A minimal set of conditions was introduced that a vector field needs to satisfy in order for the resulting discrete dynamical system to have a provable worst-case iteration complexity for LP. We then showed that the Newton vector field of the logarithmic barrier, as well as an approximation of this field due to inexact arithmetic, satisfy these conditions.

References

- [1] A. Forsgren, P. E. Gill, and J. R. Shinnerl. Stability of symmetric ill-conditioned systems arising in interior methods for constrained optimization. *SIAM J. Matrix Anal. Appl.*, 17(1):187–211, 1996.
- [2] G. H. Golub and C. F. Van Loan. *Matrix computations*, volume 3 of *Johns Hopkins Series in the Mathematical Sciences*. Johns Hopkins University Press, Baltimore, MD, second edition, 1989.
- [3] C. C. Gonzaga. Path-following methods for linear programming. *SIAM Rev.*, 34(2), 1992.
- [4] O. Güler and Y. Ye. Convergence behavior of interior-point algorithms. *Math. Programming*, 60(2, Ser. A):215–228, 1993.

- [5] Z. Lu, R. D. C. Monteiro, and J. W. O’Neal. An iterative solver-based infeasible primal-dual path-following algorithm for convex quadratic programming. *SIAM J. Optim.*, 17(1):287–310 (electronic), 2006.
- [6] N. Megiddo. Pathways to the optimal set in linear programming. In *Progress in mathematical programming (Pacific Grove, CA, 1987)*, pages 131–158. Springer, New York, 1989.
- [7] R. D. C. Monteiro and J. W. O’Neal. Convergence analysis of a long-step primal-dual infeasible interior-point linear programming algorithm based on iterative linear solvers. Technical report, School of ISyE, Georgia Tech, USA, 2003. Available at <http://www2.isye.gatech.edu/~monteiro/publications>.
- [8] R. D. C. Monteiro and T. Tsuchiya. Limiting behavior of the derivatives of certain trajectories associated with a monotone horizontal linear complementarity problem. *Math. Oper. Res.*, 21(4):793–814, 1996.
- [9] R. D. C. Monteiro and T. Tsuchiya. A strong bound on the integral of the central path curvature and its relationship with the iteration complexity of primal-dual path-following lp algorithms. Technical report, School of ISyE, Georgia Tech, USA, 2005. Available at <http://www2.isye.gatech.edu/~monteiro/publications>.
- [10] Y. Nesterov and A. Nemirovskii. *Interior-point polynomial algorithms in convex programming*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1994.
- [11] M. A. Nunez and R. M. Freund. Condition measures and properties of the central trajectory of a linear program. *Math. Programming*, 83(1, Ser. A):1–28, 1998.
- [12] J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables*, volume 30 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. Reprint of the 1970 original.
- [13] J. Renegar. A polynomial-time algorithm, based on Newton’s method, for linear programming. *Math. Programming*, 40(1, (Ser. A)):59–93, 1988.
- [14] J. Renegar. Some perturbation theory for linear programming. *Math. Programming*, 65(1, Ser. A):73–91, 1994.
- [15] J. Renegar. Incorporating condition measures into the complexity theory of linear programming. *SIAM J. Optim.*, 5(3):506–524, 1995.
- [16] J. Renegar. Linear programming, complexity theory and elementary functional analysis. *Math. Programming*, 70(3, Ser. A):279–351, 1995.
- [17] J. Renegar. *A mathematical view of interior-point methods in convex optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2001.

- [18] C. Roos, T. Terlaky, and J.-P. Vial. *Theory and algorithms for linear optimization*. John Wiley & Sons Ltd., Chichester, 1997.
- [19] S. A. Vavasis. Stable numerical algorithms for equilibrium systems. *SIAM J. Matrix Anal. Appl.*, 15(4):1108–1131, 1994.
- [20] S. A. Vavasis and Y. Ye. A primal-dual interior point method whose running time depends only on the constraint matrix. *Math. Programming*, 74(1, Ser. A):79–120, 1996.
- [21] M. H. Wright. Some properties of the Hessian of the logarithmic barrier function. *Math. Programming*, 67(2, Ser. A):265–295, 1994.
- [22] M. H. Wright. Ill-conditioning and computational error in interior methods for nonlinear programming. *SIAM J. Optim.*, 9(1):84–111 (electronic), 1999.
- [23] S. J. Wright. *Primal-dual Interior-Point Methods*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1997.
- [24] S. J. Wright. Modified Cholesky factorizations in interior-point algorithms for linear programming. *SIAM J. Optim.*, 9(4):1159–1191 (electronic), 1999.
- [25] S. J. Wright. Effects of finite-precision arithmetic on interior-point methods for nonlinear programming. *SIAM J. Optim.*, 12(1):36–78 (electronic), 2001.
- [26] Y. Ye. *Interior Point Algorithms: Theory and Analysis*. John Wiley and Sons, New York, 1997.