Mathematical Programming 53 (1992) 1-16
North-Holland

Unified complexity analysis for Newton LP methods

James Renegar
School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853, USA

Michael Shub
IBM T.J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598, USA

Received 15 August 1988
Revised manuscript received 20 December 1989

We show that a theorem of Smale can be used to unify the polynomial-time bound proofs of several of the recent interior algorithms for linear programming and convex quadratic programming.

Key words: Computational complexity, Newton's method, interior-point methods.

1. Introduction

This paper was born with the realization that a theorem of Smale can be applied to unify the polynomial-time bound proofs of several of the recent LP interior methods. For the sake of completeness, we deduce the version of the theorem that we use from the Kantorovich theory.

Let $\mathscr{X}$ and $\mathscr{Y}$ denote Banach spaces, where the norm on $\mathscr{X}$ is $\| \ \|$. For an operator $M:\mathscr{X}^k\to\mathscr{X}$, let $\|M\|$ denote the usual operator norm $\|M\|:=\sup\{\|M(u^{(1)},\dots,u^{(k)})\|;\ \|u^{(i)}\|=1\ \forall i\}$. Assume that $f:\mathscr{X}\to\mathscr{Y}$ is an analytic map, i.e., the Fréchet derivatives $D^kf(x):\mathscr{X}^k\to\mathscr{Y}$ exist for all $x\in\mathscr{X}$, $k\geq 1$ and if y is in a sufficiently small open neighborhood of x, then $f(y)=\sum_{k=0}^{\infty}(1/k!)D^kf(x)(y-x)^k$. If Df^{-1} exists at x, define

$$\beta(x):=\|Df(x)^{-1}f(x)\|,$$

$$\gamma(x):=\sup_{k\geq 2}\left\|\frac{1}{k!}Df(x)^{-1}D^kf(x)\right\|^{1/(k-1)}$$

Note that $\beta(x)$ is the step length at x when Newton's method is applied to approximate a zero of f.

This research is supported by NSF grants.

Theorem 1 (Smale [17]). *If $\beta(\bar{\xi})\gamma(\bar{\xi})\leq\frac{1}{8}$, then the Newton sequence $x^{(0)}=\bar{\xi}$, $x^{(i+1)}=x^{(i)}-Df(x^{(i)})^{-1}f(x^{(i)})$ is well-defined (i.e., the inverses exist), converges to a zero of f, and satisfies*

$$\|x^{(i+1)}-x^{(i)}\|\leq 2(\tfrac{1}{2})^{2^i}\|x^{(1)}-x^{(0)}\|. \qquad \square$$

The particular ease with which this theorem can be applied derives from the fact that it depends only on data at the initial point $\bar{\xi}$ as opposed to typical theorems in the Kantorovich theory requiring bounds on data at all points in a sufficiently large neighborhood of $\bar{\xi}$.

For our goals, a slight variation of the above theorem allows the most expedient application. The algorithms we consider are of the following general form. Assume that $F:\mathbb{R}^n\times\mathbb{R}_+\to\mathbb{R}$. Let $x^{(0)}$ be given. Recursively define

$$x^{(i+1)}:=x^{(i)}-D_xF(x^{(i)},t^{(i+1)})^{-1}F(x^{(i)},t^{(i+1)})$$

for some $t^{(i)}\downarrow 0$. We will be interested in finding specific $\varepsilon,\delta>0$ such that if $\|x^{(i)}-\xi^{(i)}\|\leq\varepsilon$, where $F(\xi^{(i)},t^{(i)})=0$, and if $|t^{(i+1)}-t^{(i)}|/t^{(i)}\leq\delta$, then $\|x^{(i+1)}-\xi^{(i+1)}\|\leq\varepsilon$ for some $\xi^{(i+1)}$ satisfying $F(\xi^{(i+1)},t^{(i+1)})=0$. (The norms we consider will actually depend on $\xi^{(i)}$, but we ignore this at present.)

One way to approach finding ε and δ is simply as follows. Find $\varepsilon,\delta>0$, independent of t, such that if $F(\xi,t)=0$, $\|\bar{\xi}-\xi\|\leq\varepsilon$, $|t'-t|/t\leq\delta$, $f(x):=F(x,t')$, then $\beta(\bar{\xi})$ and $\gamma(\bar{\xi})$ are sufficiently small so that the above theorem implies $\|\bar{\xi}'-\xi'\|\leq\varepsilon$, where $\bar{\xi}'=\bar{\xi}-Df(\bar{\xi})^{-1}f(\bar{\xi})$ and $f(\xi')=0$. However, because for the maps we consider the values $\beta(\xi)$ and $\gamma(\xi)$ are easier to compute than $\beta(\bar{\xi})$ and $\gamma(\bar{\xi})$, the following theorem provides an even quicker approach.

Let $B(x,r):=\{y;\ \|y-x\|<r\}$.

Theorem 2. *Assume that $f:\mathcal{U}\to\mathcal{Y}$ is analytic, where $\mathcal{U}$ is an open subset of $\mathcal{X}$. Assume that $\beta:=\beta(\xi)$, $\gamma:=\gamma(\xi)$ and $\delta\geq 0$ satisfy $\beta\leq\frac{1}{2}\delta\leq 1/(40\gamma)$ and $B(\xi,4\delta)\subseteq\mathcal{U}$. If $\|\bar{\xi}-\xi\|\leq\delta$, then the Newton sequence $x^{(0)}:=\bar{\xi}$, $x^{(i+1)}:=x^{(i)}-Df(x^{(i)})^{-1}f(x^{(i)})$ is well-defined and converges to a zero ξ' of f in $B(\xi,\frac{3}{2}\beta)$, this being the only zero of f in $B(\xi,4\delta)$. Moreover,*

$$\|x^{(i)}-\xi'\|\leq\tfrac{7}{2}(\tfrac{1}{2})^{2^i}\delta.$$

We remark that the power series of f at ξ converges for all $x\in B(\xi,1/\gamma)$ by the root test, and hence for naturally defined f the required containment $B(\xi,4\delta)\subseteq\mathcal{U}$ follows from the other assumptions. In particular, for rational functions $f:\mathbb{R}^n\to\mathbb{R}^n$ (i.e., quotients of polynomials) the containment need not be checked when applying the theorem.

We prove the theorem in Appendix A by showing that it is a consequence of a Kantorovich theorem as presented in Deuflhard and Heindl [5]. We first proved this result using Smale's [17] arguments, but felt the Kantorovich arguments would be more accessible to the optimization community. The constant $\frac{1}{8}$ in Smale's theorem cannot be replaced with $\frac{1}{7}$, as he showed. The requirements $\beta\leq\frac{1}{2}\delta\leq 1/(40\gamma)$ and

$B(\xi, 4\delta) \subseteq \mathcal{U}$ in the latter theorem can certainly be made better, but probably at the price of a longer proof.

We apply the theorem to prove $O(\sqrt{m}\, L)$ iteration bounds for several of the recent LP and QP interior algorithms. Here, m refers to the number of linear inequalities, assumed to exceed the number of variables, and L is the number of bits required to specify the problem to be solved.

We first apply Theorem 2 in Section 2 to the barrier method. Here the application is particularly simple. One is tempted to say that the barrier method was made for the theorem. Gonzaga [7] obtained an $O(\sqrt{m}\, L)$ iteration bound for the barrier method. Daya and Shetty [3] have also studied this algorithm.

In Section 3, we briefly discuss the slight modifications in the LP argument required to prove an $O(\sqrt{m}\, L)$ iteration bound for the barrier method applied to convex QP. Goldfarb and Liu [6] and Ye [21] proved this bound. Daya and Shetty [4] have also obtained this bound.

In Section 4, Theorem 2 is applied to the primal algorithm studied by Renegar [14]. Sonnevend [19] proposed a similar algorithm, but gave no complexity analysis. Vaidya [20] also considered a closely related algorithm.

In Section 5, we consider the primal–dual algorithm studied by Kojima et al. [9] and Monteiro and Adler [13]. This algorithm has roots in the work of Megiddo [11]. Monteiro and Adler obtained an $O(\sqrt{m}\, L)$ iteration bound for this algorithm. Their analysis is simple and direct, but from a different vantage point than ours.

In Section 6, we briefly consider the primal–dual algorithm applied to convex QP. Kojima et al. [10] and Monteiro and Adler [13] obtained an $O(\sqrt{m}\, L)$ iteration bound for this algorithm.

All of the above algorithms follow the "central trajectory," as studied by Bayer and Lagarias [1] and Megiddo and Shub [12].

Our focus is on iteration bounds as opposed to overall arithmetic complexity. However, only a moderate amount of additional work is required to obtain the record $O(n^2 mL)$ arithmetic operation bound for LP, proven independently by Gonzaga [7] and Vaidya [20]. In Appendix B we display this for the barrier method. Minor modifications of the arguments yield the same bound for convex QP, a bound that was first proven by Kojima et al. [10] and, shortly thereafter, by Monteiro and Adler [13].

We do not discuss bit operation bounds. Of course that is what is really required to prove polynomial-time bounds.

Smale [17] motivated several papers. Royden [16] derived both the Kantorovich theory and Smale's theorem from a single theorem. He also slightly improved Smale's requisite bound on $\beta\gamma$. Rheinboldt [15] has given a direct derivation of Smale's theorem from the Kantorovich theory. In Appendix A we follow Rheinboldt's approach.

Curry [2] extended Smale's theorem to higher order methods in the case of univariate polynomials. Independently of both Smale and Curry, Kim [8] proved similar results for univariate polynomials.

Of related interest is Smale [18], especially Section 4, where a similar theorem for a path-following algorithm is discussed.

Beware that we aim at giving short proofs. Little motivation for some of the ideas in the proofs is given, in particular, for the choice of norms that make everything work out so nicely. Some motivation can be found by reading, for example, Sections 2 and 3 of Renegar [14].

Now we fix some notation. Throughout we consider problems of the form min $c^T x$, subject to $Ax \geqslant b$, although application of the theorem is also easy for the linear equalities, non-negative variables format. (Most of the aforementioned LP papers assume the latter format.) We use α_i to denote the ith row of A.

We assume $\{x; Ax > b\}$ to be non-empty and bounded, and we assume a known "good" starting point for the algorithms. Although we may make these assumptions without loss of generality, we avoid the arguments as to why. For the puzzled reader we remark that almost every paper in the area discusses this. Because all of the algorithms follow the central trajectory, good starting points for one algorithm are generally easily translated into good starting points for another. Furthermore, many papers in the area discuss how to obtain an optimal solution from a feasible point $\bar{x}$, where $c^T \bar{x}$ is sufficiently close to the optimal objective value — sufficiently close is of the form $2^{-O(L)}$ (e.g., see [14], Lemma 8.1).

We use e to denote the vector of all ones, and $\| \; \|_2$ to denote the usual Euclidean norm.

Finally, note that the conclusions of Theorem 2 imply that the Newton iterates are in $\mathcal{U}$. In particular, if as in some of our applications the natural domain $\mathcal{U}$ is the interior of the feasible region, then the iterates are feasible.

We appreciate the careful consideration and comments given by the referees.

2. The LP barrier method

In this section we consider the LP barrier method that was first analyzed by Gonzaga [7].

Let $\text{Int} = \{x; Ax > b\}$, and let $h : \text{Int} \times \mathbb{R}_+ \to \mathbb{R}$ denote the map

$$h(x, t) = c^T x - t \sum \ln(\alpha_i x - b_i).$$

For fixed t, the map $x \to h(x, t)$ is strictly convex, having a unique minimum. The sequence of minima as $t \downarrow 0$ converges to the optimal solution of the LP. The algorithm simply computes a Newton sequence $x^{(0)}, x^{(1)}, \ldots$, where

$$x^{(i+1)} := x^{(i)} - \nabla_x^2 h(x^{(i)}, t^{(i+1)})^{-1} \nabla_x h(x^{(i)}, t^{(i+1)})$$

and $t^{(i)} \downarrow 0$.

We claim that for appropriately chosen $x^{(0)}$, we may always take $t^{(i+1)} = (1 - 1/(41\sqrt{m}))t^{(i)}$ and each $x^{(i)}$ will then be a "good" approximation to the minimum of the map $x \to h(x, t^{(i)})$. Now we prove this claim.

Let $\Delta(x)$ be the diagonal matrix with ith diagonal entry $\alpha_i x - b_i$.

Assume $t > 0$ fixed. Let ξ be the minimum of $x \to h(x, t)$, i.e., $c - tA^{\mathrm{T}}\Delta(\xi)^{-1}e = 0$. Define $\|x\| := \|\Delta(\xi)^{-1}Ax\|_2$. For $t' > 0$, define ξ' and $\|\ \|'$ in the obvious way.

Assume that $\|\bar{\xi} - \xi\| \leq \frac{1}{20}$, i.e., $\bar{\xi}$ is a "good" approximation to ξ. Assume that $(1 - 1/(41\sqrt{m}))t \leq t' \leq t$. Define $f(x) := c - t'A^{\mathrm{T}}\Delta(x)^{-1}e$. Let $\bar{\xi}' := \bar{\xi} - Df(\bar{\xi})^{-1}f(\bar{\xi})$. To prove our claim, we show that $\|\bar{\xi}' - \xi'\|' \leq \frac{1}{20}$.

First note that

$$\beta := \|Df(\xi)^{-1}f(\xi)\| = \frac{t - t'}{t'}\|\Delta(\xi)^{-1}A[A^{\mathrm{T}}\Delta(\xi)^{-2}A]^{-1}A^{\mathrm{T}}\Delta(\xi)^{-1}e\|_2$$

$$\leq \frac{t - t'}{t'}\|e\|_2 \leq \tfrac{1}{40},$$

where the first inequality follows from the fact that the eliminated matrix is a projection matrix.

Next we compute a bound on

$$\gamma := \sup_{k \geq 2}\left\|\frac{1}{k!}Df(\xi)^{-1}D^k f(\xi)\right\|^{1/(k-1)}.$$

Observe that the operator $D^k f(\xi) : (\mathbb{R}^n)^k \to \mathbb{R}^n$ sends the tuple $(u^{(1)}, \ldots, u^{(k)}) \in (\mathbb{R}^n)^k$ to $(-1)^k t' k! A^{\mathrm{T}}\Delta(\xi)^{-1}w$ where $w \in \mathbb{R}^m$ is the vector with $w_i = \prod_j w_i^{(j)}$ and $w^{(j)} = \Delta(\xi)^{-1}Au^{(j)}$. Since $\|w^{(j)}\|_2 = \|u^{(j)}\|$, it easily follows that $\|w\|_2 \leq 1$ if $\|u^{(j)}\| \leq 1$ for all j. Consequently,

$$\left\|\frac{1}{k!}Df(\xi)^{-1}D^k f(\xi)\right\| \leq \|\Delta(\xi)^{-1}A[A^{\mathrm{T}}\Delta(\xi)^{-2}A]^{-1}A^{\mathrm{T}}\Delta(\xi)^{-1}\|_2 = 1,$$

the equality because the eliminated matrix is a projection matrix. Hence, $\gamma \leq 1$.

Theorem 2 now implies that $\|\bar{\xi}' - \xi'\| \leq \frac{7}{8} \cdot \frac{1}{20}$ and $\|\xi - \xi'\| \leq \frac{3}{80}$. Since

$$0 \leq \frac{\alpha_i\xi - b_i}{\alpha_i\xi' - b_i} = \frac{1}{1 - [\alpha_i(\xi - \xi')]/[\alpha_i\xi - b_i]} \leq \frac{1}{1 - \|\xi - \xi'\|} \leq \tfrac{80}{77}$$

we have for all v that

$$\|v\|' := \|\Delta(\xi')^{-1}Av\|_2 \leq \|\Delta(\xi')^{-1}\Delta(\xi)\|_2\|v\| \leq \tfrac{80}{77}\|v\|.$$

In particular, $\|\bar{\xi}' - \xi'\|' \leq \frac{1}{22}$, concluding the proof of the claim.

The principal remaining ingredient for proving an $O(\sqrt{m}\,L)$ iteration bound is an inequality of the form

$$c^{\mathrm{T}}\bar{\xi} - k^* < 2tm,$$

where k^* is the optimal objective value. In proving this inequality, first note that since $c - tA^{\mathrm{T}}\Delta(\xi)^{-1}e = 0$, we have

$$c^{\mathrm{T}}(\bar{\xi} - \xi) = te^{\mathrm{T}}\Delta(\xi)^{-1}A(\bar{\xi} - \xi) \leq t\|e\|_2\|\Delta(\xi)^{-1}A(\bar{\xi} - \xi)\|_2$$

$$= \sqrt{m}\,t\|\bar{\xi} - \xi\| \leq \tfrac{1}{20}\sqrt{m}\,t.$$

Hence, assuming that x^* is an optimal solution,

$$\begin{aligned} c^{\mathrm{T}}\bar{\xi} - k^* &= c^{\mathrm{T}}(\bar{\xi} - \xi) + c^{\mathrm{T}}(\xi - x^*) \\ &\leqslant \tfrac{1}{20} t\sqrt{m} + c^{\mathrm{T}}(\xi - x^*) + t\sum \frac{\alpha_i x^* - b_i}{\alpha_i \xi - b_i} \\ &= \tfrac{1}{20} t\sqrt{m} + tm + (\xi - x^*)^{\mathrm{T}}[c - tA^{\mathrm{T}}\Delta(\xi)^{-1}e] \\ &= \tfrac{1}{20} t\sqrt{m} + tm + (\xi - x^*)^{\mathrm{T}}0. \end{aligned}$$

In Appendix B we present an argument for modifying the preceding $\mathrm{O}(n^2m^{1.5}L)$ arithmetic operation algorithm into one requiring only $\mathrm{O}(n^2mL)$ operations. The latter operation count was first established for LP by Gonzaga [7] and Vaidya [20], but with much longer proofs.

3. The QP barrier method

In this section we briefly consider the barrier method applied to convex quadratic programming. The notation and definitions remain exactly the same as in the previous section, except that the objective $c^{\mathrm{T}}x$ is replaced throughout by $\frac{1}{2}x^{\mathrm{T}}Qx + c^{\mathrm{T}}x$, where Q is symmetric and positive semi-definite, and hence $h: \mathrm{Int} \times \mathbb{R}_+ \to \mathbb{R}$ is defined by

$$h(x, t) = \tfrac{1}{2}x^{\mathrm{T}}Qx + c^{\mathrm{T}}x - t\sum \ln(\alpha_i x - b_i).$$

The crucial fact here is that if M is an $m \times n$ matrix of rank n, $m \geqslant n$, then $\|M(Q + M^{\mathrm{T}}M)^{-1}M^{\mathrm{T}}\|_2 \leqslant 1$. We will prove this inequality momentarily.

Proceeding exactly as in the preceding section for computing β and γ, except that now $Df(\xi)^{-1} = [t'A^{\mathrm{T}}\Delta(\xi)^{-2}A]^{-1}$ is replaced by $Df(\xi)^{-1} = [Q + t'A^{\mathrm{T}}\Delta(\xi)^{-2}A]^{-1}$, and using the above inequality shows that we may again take $t^{(i+1)} = (1 - 1/(41\sqrt{m}))t^{(i)}$.

Now to prove the inequality. Let S denote an $m \times m$ orthogonal matrix moving the range of M onto $\mathbb{R}^n \times \{0\}$, and let $\tilde{M}$ be the $n \times n$ invertible matrix defined by $\tilde{M} = PSM$, where P is projection onto $\mathbb{R}^n$. Then

$$\begin{aligned} &\|M(Q + M^{\mathrm{T}}M)^{-1}M^{\mathrm{T}}\|_2 \\ &\quad = \|[SM](Q + [SM]^{\mathrm{T}}[SM])^{-1}[SM]^{\mathrm{T}}\|_2 \quad \text{(by othogonality of } S\text{)} \\ &\quad = \|[PSM](Q + [PSM]^{\mathrm{T}}[PSM])^{-1}[PSM]^{\mathrm{T}}\|_2 \\ &\quad = \|\tilde{M}(Q + \tilde{M}^{\mathrm{T}}\tilde{M})^{-1}\tilde{M}^{\mathrm{T}}\|_2 \\ &\quad = \|(I + \tilde{M}^{-\mathrm{T}}Q\tilde{M}^{-1})^{-1}\|_2 \\ &\quad \leqslant 1, \end{aligned}$$

where the third, fourth and fifth expressions are with respect to the Euclidean norm on $\mathbb{R}^n$, and where the final inequality is because $\tilde{M}^{-\mathrm{T}}Q\tilde{M}^{-1}$ is positive semi-definite.

4. A primal LP algorithm

In this section we consider the algorithm studied in Renegar [14]. (Also see Sonnevend [17] and Vaidya [20].)

Assume that $k^{(0)}$ is a known strict upper bound on the optimal objective value. Assume that $Ax^{(0)} > b$, $c^T x^{(0)} < k^{(0)}$. The sequence $x^{(i)}$, $i = 1, 2, \ldots$, is defined recursively as follows. Let $x^{(i+1)}$ be obtained by applying one iteration of Newton's method, beginning at $x^{(i)}$, in attempting to maximize

$$x \to m \ln(k^{(i+1)} - c^T x) + \sum_{i=1}^{m} \ln(\alpha_i x - b_i),$$

where $k^{(i+1)} = \delta c^T x^{(i)} + (1-\delta) k^{(i)}$ and $0 < \delta < 1$ is some prespecified value. (The function is strictly concave on $\{x;\ Ax > b,\ c^T x < k^{(i+1)}\}$.)

We claim that if $x^{(0)}$ is chosen appropriately and $0 < \delta \leq 1/(42\sqrt{m})$, then for all i, $x^{(i)}$ is a "good" approximation to the actual maximum it is meant to approximate. We now prove this claim.

Let $\bar{A}$ denote the $2m \times n$ matrix with ith row α_i if $i \leq m$, $-c^T$ if $i > m$.

Assume that k is a strict upper bound for the optimal objective value. Let $\Delta(x)$ be the $2m \times 2m$ diagonal matrix with ith diagonal entry $\alpha_i x - b_i$ if $i \leq m$, $k - c^T x$ if $i > m$. Let ξ be the unique zero in $\{x;\ Ax > b,\ c^T x < k\}$ of the function

$$x \to \bar{A}^T \Delta(x)^{-1} e (= \nabla_x [m \ln(k - c^T x) + \textstyle\sum \ln(\alpha_i x - b_i)]).$$

Assume that $\bar{\xi}$ satisfies $\|\Delta(\xi)^{-1} \bar{A}(\bar{\xi} - \xi)\|_2 \leq \frac{1}{21}$, i.e., $\bar{\xi}$ is a "good" approximation to ξ. Let $k' := \delta c^T \bar{\xi} + (1-\delta) k$, where $0 \leq \delta \leq 1/(42\sqrt{m})$. Define $\Delta'(x)$ and ξ' in the obvious way. Let $f(x) := \bar{A}^T \Delta'(x)^{-1} e$, and let $\bar{\xi}' := \xi - Df(\bar{\xi})^{-1} f(\bar{\xi})$. To prove our claim we show that $\|\Delta'(\xi')^{-1} \bar{A}(\bar{\xi}' - \xi')\|_2 \leq \frac{1}{21}$.

Let $\hat{\delta} = (k - k')/(k - c^T \xi)$. Then

$$0 \leq \hat{\delta} = \delta\left(1 + \frac{c^T(\xi - \bar{\xi})}{k - c^T \xi}\right) \leq \delta(1 + \|\Delta(\xi)^{-1} \bar{A}(\bar{\xi} - \xi)\|_2) < 1/(40\sqrt{m}).$$

Defining $f(\xi) := \bar{A}^T \Delta'(\xi)^{-1} e$, note that

$$f(\xi) = \bar{A}^T [\Delta'(\xi)^{-1} - \Delta(\xi)^{-1}] e = \hat{\delta} \bar{A}^T \Delta'(\xi)^{-1} \tilde{e},$$

where $\tilde{e}_i = 0$ if $i \leq m$, $\tilde{e}_i = 1$ if $i > m$. Thus, defining $\|v\| := \|\Delta'(\xi)^{-1} \bar{A} v\|_2$ for all $v \in \mathbb{R}^n$, we have that

$$\begin{aligned}\beta &:= \|Df(\xi)^{-1} f(\xi)\| = \hat{\delta} \|\Delta'(\xi)^{-1} \bar{A}(\bar{A}^T \Delta'(\xi)^{-2} \bar{A})^{-1} \bar{A}^T \Delta'(\xi)^{-1} \tilde{e}\|_2 \\ &\leq \hat{\delta} \|\tilde{e}\|_2 \leq \tfrac{1}{40},\end{aligned}$$

the first inequality following from the fact that the eliminated matrix is a projection matrix. Moreover, by exactly the same argument as in Section 2 we have that

$$\gamma := \sup_{k \geq 2} \left\| \frac{1}{k!} Df(\xi)^{-1} D^k f(\xi) \right\|^{1/(k-1)} \leq 1.$$

Since

$$\|\bar{\xi}-\xi\| \leqslant \|\Delta'(\xi)^{-1}\Delta(\xi)\|_2\|\Delta(\xi)^{-1}\bar{A}(\bar{\xi}-\xi)\|_2 \leqslant \frac{1}{1-\bar{\delta}}\cdot\tfrac{1}{21}<\tfrac{1}{20},$$

the theorem implies that $\|\bar{\xi}'-\xi'\| \leqslant \frac{7}{8}\cdot\frac{1}{20}$ and $\|\xi-\xi'\| \leqslant \frac{3}{80}$. Hence,

$$\|\Delta'(\xi')^{-1}\bar{A}(\bar{\xi}'-\xi')\|_2 \leqslant \|\Delta'(\xi')^{-1}\Delta'(\xi)\|_2\|\bar{\xi}'-\xi'\| \leqslant \frac{\|\bar{\xi}'-\xi'\|}{1-\|\xi'-\xi\|}<\tfrac{1}{21},$$

concluding the proof of the claim.

The principal remaining ingredient for proving an $O(\sqrt{m}\,L)$ iteration bound is an inequality of the form

$$k'-k^* \leqslant (1-\tfrac{1}{3}\delta)(k-k^*),$$

where k^* is the optimal objective value. In proving this inequality, assume that x^* is an optimal solution. Then

$$k-k^*=(k-c^{\mathrm{T}}\xi)\left(\frac{k-c^{\mathrm{T}}x^*}{k-c^{\mathrm{T}}\xi}\right).$$

However,

$$0 \leqslant m\frac{k-c^{\mathrm{T}}x^*}{k-c^{\mathrm{T}}\xi} \leqslant m\frac{k-c^{\mathrm{T}}x^*}{k-c^{\mathrm{T}}\xi}+\sum\frac{\alpha_i x^*-b_i}{\alpha_i\xi-b_i}$$
$$=2m+(x^*-\xi)^{\mathrm{T}}\bar{A}^{\mathrm{T}}\Delta(\xi)^{-1}e=2m+(x^*-\xi)^{\mathrm{T}}0.$$

Hence, $k-c^{\mathrm{T}}\xi \geqslant \frac{1}{2}(k-k^*)$, and thus

$$k'-k^*=(k-k^*)-(k-k')=(k-k^*)-\hat{\delta}(k-c^{\mathrm{T}}\xi) \leqslant (1-\tfrac{1}{2}\hat{\delta})(k-k^*).$$

Since

$$\hat{\delta}=\delta\left(1+\frac{c^{\mathrm{T}}(\xi-\bar{\xi})}{k-c^{\mathrm{T}}\xi}\right) \geqslant \delta(1-\|\Delta(\xi)^{-1}\bar{A}(\bar{\xi}-\xi)\|_2) \geqslant \tfrac{20}{21}\delta,$$

it now follows that

$$k'-k^* \leqslant (1-\tfrac{1}{3}\delta)(k-k^*).$$

5. The primal–dual LP algorithm

In this section we consider the primal–dual LP algorithm studied by Kojima et al. [9] and Monteiro and Adler [13]. This algorithm has roots in the work of Megiddo [11].

As in Section 2, define $\mathrm{Int}=\{x;\ Ax>b\}$ and let $h:\mathrm{Int}\times\mathbb{R}_+\to\mathbb{R}$ be given by $h(x,t)=c^{\mathrm{T}}x-t\sum\ln(\alpha_i x-b_i)$. The unique minimum of $x\to h(x,t)$ is the point satisfying $c - tA^{\mathrm{T}}\Delta(x)^{-1}e = 0$, where $\Delta(x)$ is the $m \times m$ diagonal matrix with ith diagonal entry $\alpha_i x - b_i$. Equivalently, it is determined by the equations

$$Z\Delta(x)e - te = 0,$$

$$c - A^{\mathrm{T}}z = 0,$$

where Z is the diagonal matrix with ith diagonal entry z_i. Letting $w := (x, z) \in \mathbb{R}^{n+m}$, and letting $H(w, t) = 0$ denote the above system of equations, the algorithm simply computes a Newton sequence $w^{(0)}, w^{(1)}, \ldots$, where

$$w^{(i+1)} = w^{(i)} - D_w H(w^{(i)}, t^{(i+1)})^{-1} H(w^{(i)}, t^{(i+1)}),$$

and $t^{(i)} \downarrow 0$.

We claim that for appropriately chosen $w^{(0)}$, we may always take $t^{(i+1)} = (1 - 1/(40\sqrt{m}))t^{(i)}$ and each $w^{(i)}$ will then be a "good" approximation to the zero of $w \to H(w, t^{(i)})$ satisfying $x \in \mathrm{Int}$. Now we prove this claim.

Fix $\hat{w} = (0, \hat{z})$ satisfying $c - A^{\mathrm{T}}\hat{z} = 0$.

Let $\mathscr{X} = \{(x, z) \in \mathbb{R}^{n+m};\ A^{\mathrm{T}}z = 0\}$. We assume that $w^{(0)}$ satisfies $c - A^{\mathrm{T}}z^{(0)} = 0$. Then $w^{(i)} - \hat{w} \in \mathscr{X}$ for all i.

In what follows, $w = (x, z)$ always refers to points in $\mathscr{X}$.

Assume $t > 0$ fixed. Let $\xi = (\rho, \omega) \in \mathscr{X}$ be the unique zero of $w \to H(w + \hat{w}, t)$ satisfying $\rho \in \mathrm{Int}$. Let Ω denote the $m \times m$ diagonal matrix with ith diagonal entry ω_i. Then $(\Omega + \hat{Z})\Delta(\rho)e - te = 0$.

For $w = (x, z) \in \mathscr{X}$, define

$$\|w\| := \frac{1}{t}\|(\Omega + \hat{Z})Ax + Z\Delta(\rho)e\|_2$$

$$= \frac{1}{t}(\|(\Omega + \hat{Z})Ax\|_2^2 + \|Z\Delta(\rho)e\|_2^2)^{1/2},$$

where the equality is because $(Z\Delta(\rho)e)^{\mathrm{T}}(\Omega + \hat{Z})Ax = tz^{\mathrm{T}}Ax = 0$.

For $t' > 0$, define $\xi' = (\rho', w')$, Ω' and $\| \ \|'$ in the obvious way.

Assume that $\bar{\xi} \in \mathscr{X}$, $\|\bar{\xi} - \xi\| \leq \frac{1}{20}$, i.e., $\bar{\xi}$ is a good approximation to ξ. Assume that $(1 - 1/(40\sqrt{m}))t \leq t' \leq t$. Define $f : \mathscr{X} \to \mathbb{R}^m$ by $f(w) := (Z + \hat{Z})\Delta(x)e - t'e$. Let $\bar{\xi}' = \bar{\xi} - Df(\bar{\xi})^{-1}f(\bar{\xi})$. To prove our claim we show that $\|\bar{\xi}' - \xi'\|' \leq \frac{1}{20}$.

Observe that for $w \in \mathscr{X}$, $\|w\| = (1/t)\|Df(\xi)w\|_2$. Hence

$$\beta := \|Df(\xi)^{-1}f(\xi)\| = \frac{t - t'}{t}\|e\|_2 \leq \tfrac{1}{40}.$$

Let $w^{[i]} = (x^{[i]}, z^{[i]})$, $i = 1, 2$, denote two arbitrary vectors in $\mathscr{X}$. Note that $D^2 f(\xi)$ maps $(w^{[1]}, w^{[2]})$ to the vector $Z^{[1]}Ax^{[2]} + Z^{[2]}Ax^{[1]}$. Since $\|w\| = (1/t)\|Df(\xi)w\|_2$, it follows that $Df(\xi)^{-1}D^2 f(\xi)$ maps $(w^{[1]}, w^{[2]})$ to a vector of $\| \ \|$-length at most

$$\frac{1}{t}\|Z^{[1]}Ax^{[2]}\|_2 + \frac{1}{t}\|Z^{[2]}Ax^{[1]}\|_2.$$

Now,

$$\|Z^{[1]}Ax^{[2]}\|_2 = \frac{1}{t}\|Z^{[1]}\Delta(\rho)(\Omega+\hat{Z})Ax^{[2]}\|_2$$

$$\leq \frac{1}{t}\|Z^{[1]}\Delta(\rho)e\|_2\|(\Omega+\hat{Z})Ax^{[2]}\|_2 \leq t\|w^{[1]}\|\,\|w^{[1]}\|.$$

The same inequality holds for $\|Z^{[2]}Ax^{[1]}\|_2$. Since $D^k f \equiv 0$ for $k \geq 2$, it follows that

$$\gamma := \sup_{k\geq 2}\left\|\frac{1}{k!}Df(\xi)^{-1}D^k f(\xi)\right\|^{1/(k-1)} \leq 1.$$

The theorem now implies that $\|\bar{\xi}' - \xi'\| \leq \frac{7}{8}\cdot\frac{1}{20}$ and $\|\xi - \xi'\| \leq \frac{3}{80}$.

Note that by the non-negativity of the entries in the diagonal matrices,

$$\|\Delta(\rho')\Delta(\rho)^{-1}\|_2 \leq 1 + \|\Delta(\rho)^{-1}\Delta(\rho'-\rho)\|_2$$

$$= 1 + \frac{1}{t}\|(\Omega+\hat{Z})\Delta(\rho'-\rho)\|_2$$

$$\leq 1 + \frac{1}{t}\|(\Omega+\hat{Z})\Delta(\rho'-\rho)e\|_2$$

$$\leq 1 + \|\xi' - \xi\|.$$

Similarly, $\|(\Omega'+\hat{Z})(\Omega+\hat{Z})^{-1}\|_2 \leq 1 + \|\xi'-\xi\|$. Hence, for all $w \in \mathscr{X}$,

$$\|w\|' \leq \frac{1}{t'}(\|(\Omega'+\hat{Z})(\Omega+\hat{Z})^{-1}\|_2^2\|(\Omega+\hat{Z})Ax\|_2^2$$

$$+ \|\Delta(\rho')\Delta(\rho)^{-1}\|_2^2\|Z\Delta(\rho)e\|_2^2)^{1/2}$$

$$\leq (1+\|\xi'-\xi\|)\|w\|.$$

In particular, it now follows that $\|\bar{\xi}'-\xi'\|' \leq \frac{1}{20}$, concluding the proof of the claim.

Assuming $\bar{\xi} = (\bar{\rho}, \bar{\omega})$, the primal feasibility of $\bar{\rho}$ follows from $\omega_i + \hat{z}_i > 0$ (for all i) and

$$\tfrac{1}{20} \geq \|\bar{\xi} - \xi\| \geq \frac{1}{t}\|(\Omega+\hat{Z})A(\bar{\rho}-\rho)\|_2 = \frac{1}{t}\|(\Omega+\hat{Z})\Delta(\bar{\rho}) - te\|_2.$$

Similarly, for the dual feasibility of $\bar{w}+\hat{z}$.

The principal remaining ingredient for proving an $O(\sqrt{m}\,L)$ iteration bound is a duality gap bound of the form

$$c^{\mathrm{T}}\bar{\rho} - b^{\mathrm{T}}(\bar{w}+\hat{z}) < 2tm.$$

This is implied by combining

$$c^{\mathrm{T}}\rho - (\omega+\hat{z})^{\mathrm{T}}b = (\omega+\hat{z})^{\mathrm{T}}A\rho - e^{\mathrm{T}}[(\Omega+\hat{Z})(\Delta(\rho)+b)+te] = tm$$

with

$$|c^{\mathrm{T}}(\bar{\rho}-\rho)| = |e^{\mathrm{T}}(\Omega+\hat{Z})A(\bar{\rho}-\rho)| \leq \sqrt{m}\,\|\bar{\xi}-\xi\| \leq \tfrac{1}{20}\sqrt{m}$$

and similarly, $|b^{\mathrm{T}}(\bar{\omega}-\omega)| \leq \frac{1}{20}\sqrt{m}$.

6. The primal–dual QP algorithm

In this section we briefly consider the primal-dual algorithm applied to convex quadratic programming, as has been analyzed by Kojima et al. [10] and Monteiro and Adler [13]. Replacing the objective by $\frac{1}{2}x^{\mathrm{T}}Qx + c^{\mathrm{T}}x$ where Q is positive semi-definite, the same motivation as in the preceeding section leads us to consider the system of equations

$$Z\Delta(x) - te = 0,$$

$$Qx - A^{\mathrm{T}}z + c = 0,$$

which again we denote by $H(w) = 0$. The algorithm is defined analogously to the LP case. Following are the changes in the analysis that need to be made.

Now fix $\hat{w} = (\hat{x}, \hat{z})$ satisfying $Q\hat{x} - A^{\mathrm{T}}\hat{z} + c = 0$. Also, assume that the initial iterate $w^{(0)}$ satisfies this equation.

Let $\mathscr{X} := \{(x, z) \in \mathbb{R}^{n+m};\ Qx - A^{\mathrm{T}}z = 0\}$.

Let $\xi = (\rho, \omega) \in \mathscr{X}$ be the unique zero of $w \to H(w + \hat{w}, t)$ satisfying $\rho + \hat{x} \in \mathrm{Int}$. Define ξ' similarly.

Beginning with the paragraph "Assume $t > 0$ fixed..." onward, replace every $\Delta(x)$ by $\Delta(x + \hat{x})$, e.g., replace $\Delta(\rho)$ by $\Delta(\rho + \hat{x})$.

For $w \in \mathscr{X}$, define

$$\|w\| := \frac{1}{t}\|(\Omega + \hat{Z})Ax + Z\Delta(\rho + \hat{x})e\|_2$$

$$\geqslant \frac{1}{t}(\|(\Omega + \hat{Z})Ax\|_2^2 + \|Z\Delta(\rho + \hat{x})e\|_2^2)^{1/2},$$

the inequality following from the fact that $z^{\mathrm{T}}Ax = x^{\mathrm{T}}Qx \geqslant 0$. Define $\|\ \|'$ similarly.

Leaving the arguments otherwise unchanged, again we find that we may take $t^{(i+1)} = (1 - 1/(40\sqrt{m}))t^{(i)}$.

Appendix A

In this appendix we prove the theorem.

We begin by noting it is not difficult to verify that the following theorem is implied by Theorem 2 of Deuflhard and Heindl [5].

Let $\bar{B}(x, r) = \{y;\ \|x - y\| \leqslant r\}$.

Theorem 3. *Let $\mathscr{B}$ be an open convex subset of a Banach space with norm $\|\ \|$, and assume that $f : \mathscr{B} \to \mathscr{Y}$ is a twice continuously Fréchet-differentiable mapping into a Banach space $\mathscr{Y}$. Assume that Df is invertible at $x^{(0)} \in \mathscr{B}$. Furthermore, assume that*

$\beta, \omega > 0$ satisfy

$$\|Df(x^{(0)})^{-1}f(x_0)\| \le \beta,$$

$$\|Df(x^{(0)})^{-1}D^2f(y)\| \le \omega \quad \textit{for all } y \in \mathscr{B},$$

$$\beta\omega \le \tfrac{4}{9},$$

$$\bar{B}(x^{(0)}, \tfrac{3}{2}\beta) \subset \mathscr{B}.$$

Then $Df(x)^{-1}$ exists for all $x \in B(x^{(0)}, \frac{3}{2}\beta)$. Moreover, the Newton sequence $x^{(0)}, x^{(1)}, \ldots,$ is contained in $B(x^{(0)}, \frac{3}{2}\beta)$, converges to the unique zero x^* of f in $\mathscr{B} \cap B(x^{(0)}, 1/\omega)$, and satisfies

$$\|x^{(i)} - x^*\| \le 2(\tfrac{1}{2})^{2^i}\beta \quad \textit{for } i \ge 1. \qquad \square$$

Now to prove our theorem.

Proof of Theorem 2. We first obtain the bound on $\|\xi - \xi'\|$. For $y \in \mathscr{B} := B(\xi, 4\delta)$ note that

$$\begin{aligned} &\|Df(\xi)^{-1}D^2f(y)\| \\ &= \left\| Df(\xi)^{-1} \sum_{i=0}^{\infty} \frac{1}{i!} D^{2+i}f(\xi)(y-\xi)^i \right\| \\ &\le \gamma \sum_{i=0}^{\infty} (i+2)(i+1)[\gamma\|y-\xi\|]^i \le \frac{2\gamma}{(1-\frac{1}{5})^3} < 4\gamma. \end{aligned} \tag{A.1}$$

Letting $x^{(0)} := \xi$, our assumed bounds and the previous theorem imply that there exists a zero ξ' of f in $B(\xi, \frac{3}{2}\beta)$, and this is the unique zero of f in $B(\xi, 4\delta)$. (The uniqueness will imply that the Newton sequence initiated at $\bar{\xi}$ converges to the same zero.)

Now we turn attention to $x^{(0)} := \bar{\xi}$. Begin by noting that

$$\begin{aligned} &\|Df(\xi)^{-1}[Df(\bar{\xi}) - Df(\xi)]\| \\ &= \left\| Df(\xi)^{-1} \sum_{i=1}^{\infty} \frac{1}{i!} D^{i+1}f(\xi)(\bar{\xi}-\xi)^i \right\| \\ &\le \sum_{i=1}^{\infty} (i+1)[\gamma\|\bar{\xi}-\xi\|]^i \le \frac{1}{(1-\frac{1}{20})^2} - 1 = \tfrac{39}{361}. \end{aligned}$$

In particular, using the identity $B^{-1} = \sum_{i=0}^{\infty} (-1)^i (A^{-1}[B-A])^i A^{-1}$ (assuming $\|A^{-1}[B-A]\| < 1$), we have that for any $v \in \mathscr{X}$,

$$\|Df(\bar{\xi})^{-1}v\| \le \tfrac{9}{8}\|Df(\xi)^{-1}v\|. \tag{A.2}$$

Consequently,

$$\bar{\beta} := \|Df(\bar{\xi})^{-1}f(\bar{\xi})\| \leq \tfrac{9}{8}\|Df(\xi)^{-1}f(\bar{\xi})\|$$
$$= \tfrac{9}{8}\left\| Df(\xi)^{-1} \sum_{i=0}^{\infty} \frac{1}{i!} D^i f(\xi)(\bar{\xi}-\xi)^i \right\|$$
$$\leq \tfrac{9}{8}(\beta + \|\bar{\xi}-\xi\| \sum_{j=0}^{\infty} [\gamma\|\bar{\xi}-\xi\|]^j)$$
$$\leq \tfrac{9}{8}\left(\beta + \frac{\delta}{1-\frac{1}{20}}\right) < \tfrac{7}{4}\delta.$$

Moreover, for $y \in B(\bar{\xi}, \frac{3}{2}\cdot\frac{7}{4}\delta)$ (and hence $y \in B(\xi, 4\delta)$), we have using (A.1) and (A.2) that

$$\|Df(\bar{\xi})^{-1}D^2f(y)\| \leq \tfrac{9}{8}\|Df(\xi)^{-1}D^2f(y)\| \leq \tfrac{9}{2}\gamma.$$

Hence

$$\bar{\omega} := \sup\{\|Df(\bar{\xi})^{-1}D^2f(y)\|;\ y \in B(\bar{\xi}, \tfrac{3}{2}\cdot\tfrac{7}{4}\delta)\} \leq \tfrac{9}{2}\gamma.$$

Since $\delta \leq 1/(20\gamma)$, we have that $\bar{\beta}\bar{\omega} \leq \frac{4}{9}$. The previous theorem applied with $\bar{\xi}$, $\bar{\beta}$, $\bar{\omega}$ now gives our theorem. □

Appendix B

In this appendix, we obtain an $O(n^2mL)$ arithmetic operation bound for the following modified barrier method.

Let $x^{(0)}$ be a feasible point, $t^{(0)} > 0$, $\Delta^{(0)} = \Delta(x^{(0)})$, and $M^{(0)} = A^T\Delta(x^{(0)})^{-2}A$. Define $t^{(i+1)} := (1 - 1/(41\sqrt{m}))t^{(i)}$,

$$x^{(i+1)} := x^{(i)} - [M^{(i)}]^{-1}(c - t^{(i+1)}A^T\Delta(x^{(i)})^{-1}e).$$

Define $\Delta^{(i+1)}$ to be the $m \times m$ diagonal matrix with jth diagonal entry

$$\delta_j^{(i+1)} := \begin{cases} \delta_j^{(i)} & \text{if } |1 - \delta_j^{(i)}/(\alpha_j x^{(i+1)} - b_j)| \leq \frac{1}{50}, \\ \alpha_j x^{(i+1)} - b_j & \text{otherwise,} \end{cases}$$

and let $M^{(i+1)} := A^T[\Delta^{(i+1)}]^{-2}A$.

The Sherman-Morrison-Woodbury formula can be applied to compute $M^{(i+1)}$ from $M^{(i)}$ in $O(n^2N^{(i)})$ arithmetic operations, where $N^{(i)} = \#\{j;\ \delta_j^{(i+1)} \neq \delta_j^{(i)}\}$. The number of arithmetic operations required by the algorithm is then $O(n^2\sum_{i=0}^{I-1} N^{(i)})$, where I is the number of iterations required to guarantee that $k^* - c^Tx^{[I]}$ is sufficiently small to determine an optimal solution x^* (implicitly assuming the iteration costs to dominate the cost of determining x^* from $x^{[I]}$, as is the case with our bounds).

Let $\xi^{(i)}$ be the unique zero of $x \to c - t^{(i)}A^T\Delta(x)^{-1}e$ in $\{x;\ Ax > b\}$. We claim that if $\|\Delta(\xi^{(0)})^{-1}A(x^{(0)} - \xi^{(0)})\|_2 \leq \frac{1}{20}$, then $\|\Delta(\xi^{(i)})^{-1}A(x^{(i)} - \xi^{(i)})\|_2 \leq \frac{1}{20}$ for all i. It then follows that $I = O(\sqrt{m}\,L)$.

Retaining the notation and assumptions of Section 2, let Δ be an $m\times m$ diagonal with jth diagonal entry δ_j satisfying $|1-\delta_j/(\alpha_j\bar{\xi}-b_j)|\leqslant\frac{1}{50}$ for all j. Let $M:=t'A^{\mathrm{T}}\Delta^{-2}A$. Since, as we have shown, $\|\bar{\xi}'-\xi'\|'\leqslant\frac{1}{22}$, to prove our claim it suffices to show that

$$\|(M^{-1}-Df(\bar{\xi})^{-1})f(\bar{\xi})\|'\leqslant\tfrac{1}{220},$$

because then $\|(\bar{\xi}-M^{-1}f(\bar{\xi}))-\xi'\|'\leqslant\frac{1}{20}$.

Assume that $\|v\|=1$, that is, $\|\Delta(\xi)^{-1}Av\|_2=1$. Then

$$\begin{aligned}
&\|Df(\bar{\xi})^{-1}(M-Df(\bar{\xi}))v\|\\
&\quad=\|\Delta(\xi)^{-1}A[A^{\mathrm{T}}\Delta(\bar{\xi})^{-2}A]^{-1}A^{\mathrm{T}}\Delta(\bar{\xi})^{-1}(\Delta^{-2}\Delta(\bar{\xi})^2-I)\Delta(\bar{\xi})^{-1}Av\|_2\\
&\quad\leqslant\|\Delta(\xi)^{-1}\Delta(\bar{\xi})\|_2\|\Delta(\bar{\xi})^{-1}A[A^{\mathrm{T}}\Delta(\bar{\xi})^{-2}A]^{-1}A^{\mathrm{T}}\Delta(\bar{\xi})^{-1}\|_2\\
&\qquad\times\|\Delta^{-2}\Delta(\bar{\xi})^2-I\|_2\|v\|\\
&\quad=\|\Delta(\xi)^{-1}\Delta(\bar{\xi})\|_2\cdot\|\Delta^{-2}\Delta(\bar{\xi})^2-I\|_2\\
&\quad<(1+\|\bar{\xi}-\xi\|)\cdot\frac{20}{(21)^2}\leqslant\tfrac{1}{21}.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\|(M^{-1}-Df(\bar{\xi})^{-1})f(\bar{\xi})\|\\
&\quad=\left\|\sum_{i=1}^{\infty}(-1)^i(Df(\bar{\xi})^{-1}[M-Df(\bar{\xi})])^iDf(\bar{\xi})^{-1}f(\bar{\xi})\right\|\\
&\quad\leqslant\tfrac{1}{20}\|Df(\bar{\xi})^{-1}f(\bar{\xi})\|.
\end{aligned}\tag{B.1}$$

Since

$$\|Df(\bar{\xi})^{-1}f(\bar{\xi})\|=\|\bar{\xi}-\bar{\xi}'\|\leqslant\|\bar{\xi}-\xi\|+\|\xi-\xi'\|+\|\bar{\xi}'-\xi'\|<\tfrac{7}{80},\tag{B.2}$$

and as we have seen, $\|v\|'\leqslant\frac{80}{77}\|v\|$ for all v, our claim is proven.

Now we obtain a bound on $\sum_{i=0}^{I-1}N^{(i)}$.

Defining $\hat{\xi}:=\bar{\xi}-M^{-1}f(\bar{\xi})$, (B.1) and (B.2) imply that $\|\hat{\xi}-\bar{\xi}\|<\frac{1}{10}$. Hence,

$$\|\Delta(\bar{\xi})^{-1}A(\hat{\xi}-\bar{\xi})\|_2\leqslant\|\Delta(\bar{\xi})^{-1}\Delta(\xi)\|_2\|\hat{\xi}-\bar{\xi}\|\leqslant\frac{\|\hat{\xi}-\bar{\xi}\|}{1-\|\bar{\xi}-\xi\|}<\tfrac{1}{9}.$$

In particular, since for any $v\in\mathbb{R}^m$ satisfying $\|v\|_2<1$,

$$\sum_j|\ln(1+v_j)|\leqslant\sqrt{m}\,\|v\|_2/(1-\|v\|_2),$$

it follows that

$$\sum_j\left|\ln\left(\frac{\alpha_jx^{(i+1)}-b_j}{\alpha_jx^{(i)}-b_j}\right)\right|\leqslant\tfrac{1}{8}\sqrt{m}.$$

Now if $\delta_j^{(i+1)}\neq\delta_j^{(i)}$, then $|\ln(\delta_j^{(i+1)}/\delta_j^{(i)})|\geqslant\ln(1+\frac{1}{50})$. Hence,

$$\sum_{i=0}^{I-1}N^{(i)}\leqslant\frac{1}{\ln(1+\frac{1}{50})}\sum_j\sum_{i=0}^{I-1}\left|\ln\left(\frac{\delta_j^{(i+1)}}{\delta_j^{(i)}}\right)\right|.$$

However, assuming for fixed j that $\delta_j^{(i+1)} \neq \delta_j^{(i)}$ precisely when $i = i_1, \ldots, i_h$ (and defining $i_0 = 1$),

$$\sum_{i=0}^{I-1} \left| \ln\left(\frac{\delta_j^{(i+1)}}{\delta_j^{(i)}}\right) \right| = \sum_{k=0}^{h-1} \left| \sum_{i=i_k}^{i_k-1} \ln\left(\frac{\alpha_j x^{(i+1)} - b_j}{\alpha_j x^{(i)} - b_j}\right) \right| \leqslant \sum_{i=0}^{I-1} \left| \ln\left(\frac{\alpha_j x^{(i+1)} - b_j}{\alpha_j x^{(i)} - b_j}\right) \right|.$$

It follows that $\sum_{i=0}^{I-1} N^{(i)} = O(\sqrt{m}\, I)$.

Note added in proof

In the period between when this paper was submitted for publication and when we received the galleys, we became aware of the work of Nesterov and Nemirovsky [22].

References

[1] D. Bayer and J.C. Lagarias, "The nonlinear geometry of linear programming, I. Affine and projective scaling trajectories, II. Legendre transform coordinates and central trajectories," *Transactions of the American Mathematical Society* 2 (1989) 499-526.

[2] J. Curry, "On zero finding methods of higher order from data at one point," *Journal of Complexity* 5 (1989) 219-237.

[3] M.B. Daya and C.M. Shetty, "Polynomial barrier function algorithms for linear programming," School of Industrial and Systems Engineering, Georgia Institute of Technology (Atlanta, GA, 1987).

[4] M.B. Daya and C.M. Shetty, "Polynomial barrier function algorithms for convex quadratic programming," School of Industrial and Systems Engineering, Georgia Institute of Technology (Atlanta, GA, 1988).

[5] P. Deuflhard and G. Heindl, "Affine invariant convergence theorems for Newton's method and extensions to related methods," *SIAM Journal of Numerical Analysis* 16 (1979) 1-10.

[6] D. Goldfarb and S. Liu, "An $O(n^3L)$ primal interior point algorithm for convex quadratic programming," *Mathematical Programming* 49 (1991) 325-340.

[7] C.C. Gonzaga, "An algorithm for solving linear programming problems in $O(n^3L)$ operations," in: N. Megiddo, ed., *Progress in Mathematical Programming, Interior Point and Related Methods* (Springer, New York, 1988) pp. 1-28.

[8] M. Kim, "On approximate roots and rootfinding algorithms for a complex polynomial," *Mathematics of Computation* 51 (1988) 707-719.

[9] M. Kojima, S. Mizuno and A. Yoshise, "A primal-dual interior point method for linear programming," in: N. Megiddo, ed., *Progress in Mathematical Programming, Interior Point and Related Methods* (Springer, New York, 1988) pp. 29-47.

[10] M. Kojima, S. Mizuno and A. Yoshise, "A polynomial-time algorithm for a class of linear complementarity problems," *Mathematical Programming* 44 (1989) 1-26.

[11] N. Megiddo, "Pathways to the optimal set in linear programming," in: N. Megiddo, ed., *Progress in Mathematical Programming, Interior Point and Related Methods* (Springer, New York, 1988) pp. 131-158.

[12] N. Megiddo and M. Shub, "Boundary behaviour of interior point algorithms in linear programming," *Mathematics of Operations Research* 14 (1989) 97-146.

[13] R.C. Monteiro and I. Adler, "Interior path following primal-dual algorithms, I. Linear programming, II. Convex quadratic programming," *Mathematical Programming* 44 (1989) 27-66.

[14] J. Renegar, "A polynomial-time algorithm, based on Newton's method, for linear programming," *Mathematical Programming* 40 (1988) 59-93.

[15] W.C. Rheinboldt, "On a theorem of S. Smale about Newton's method for analytic mappings," *Applied Mathematics Letters* 1 (1988) 69-72.

[16] H.W. Royden, "Newton's method," Department of Mathematics, Stanford University (Stanford, CA, 1986).

[17] S. Smale, "Newton's method estimates from data at one point," in: *The Merging of Disciplines in Pure, Applied, and Computational Mathematics* (Springer, New York-Berlin, 1986) pp. 185-196.

[18] S. Smale, "Algorithms for solving equations," *Proceedings of the International Congress of Mathematicians* (Berkeley, CA, 1986) pp. 87-121.

[19] G. Sonnevend, "An 'Analytic Center' for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming," in: *Lecture Notes in Control and Information Sciences No. 84* (Springer, Berlin, 1986) pp. 866-876.

[20] P. Vaidya, "An algorithm for linear programming that requires $O(((m+n)n^2+(m+n)^{1.5}n)L)$ arithmetic operations," *Proceedings of the Nineteenth ACM Symposium on the Theory of Computing* (1987) pp. 29-38.

[21] Y. Ye, "Further developments on an interior algorithm for convex quadratic programming," unpublished manuscript.

[22] Ju. E. Nesterov and A.S. Nemirovsky, *Self-Concordant Functions and Polynomial-Time Methods in Convex Programming* (Central Economic and Mathematical Institute, USSR Academy of Sciences, Moscow, 1989).