A simplification to “A Primal-Dual Interior Point Method Whose Running Time Depends Only on the Constraint Matrix”

Stephen A. Vavasis Yinyu Ye

October 30, 1996

Abstract

This note provides a simplified proof concerning the paper “A Primal-Dual Interior Point Method Whose Running Time Depends Only on the Constraint Matrix” by the same authors. In particular, we prove that Case II, one of the three cases in the method, can never occur.

1 Introduction

Consider the linear programming (LP) problems in primal-dual form:

$$\begin{array}{rl} \text{Primal: minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x}=\mathbf{b}, \\ & \mathbf{x}\ge\mathbf{0}, \end{array} \tag{1}$$

and

$$\begin{array}{rl} \text{Dual: maximize} & \mathbf{b}^T\mathbf{y} \\ \text{subject to} & \mathbf{s}=\mathbf{c}-A^T\mathbf{y}\ge\mathbf{0}, \end{array} \tag{2}$$

where $\mathbf{c}\in\mathbb{R}^n$ and $\mathbf{b}\in\mathbb{R}^m$ are given vectors, and $\mathbf{x}\in\mathbb{R}^n$ and $\mathbf{y}\in\mathbb{R}^m$ are unknown vectors.

If this problem has an optimal solution, then there exists a unique index set $B\subset\{1,...,n\}$ and $N=\{1,...,n\}\setminus B$, such that every $\mathbf{x}$ satisfying

$$A_B\mathbf{x}_B=\mathbf{b},\quad \mathbf{x}_B\ge\mathbf{0},\quad \mathbf{x}_N=\mathbf{0}$$

is an optimal solution for the primal, and every $(\mathbf{y}, \mathbf{s})$ satisfying

$$\mathbf{s}_B = \mathbf{c}_B - A_B^T\mathbf{y} = \mathbf{0}, \quad \mathbf{s}_N = \mathbf{c}_N - A_N^T\mathbf{y} \geq \mathbf{0}$$

is an optimal solution for the dual. This partition is called the strict complementarity partition. Here, for example, subvector $\mathbf{x}_B$ contains all x_i for $i \in B$.

Consider the following equations:

$$\begin{aligned} A\mathbf{x} &= \mathbf{b}, \\ A^T\mathbf{y} + \mathbf{s} &= \mathbf{c}, \\ SX\mathbf{e} &= \mu\mathbf{e}, \\ \mathbf{x} > \mathbf{0}, &\quad \mathbf{s} > \mathbf{0}, \end{aligned} \tag{3}$$

where X denotes $\mathrm{diag}(\mathbf{x})$ and S denotes $\mathrm{diag}(\mathbf{s})$, and $\mathbf{e}$ denotes the vector of all 1's. These equations always have a unique solution for any $\mu > 0$ provided that the primal and dual problem both have interior feasible points. The solution to these equations, written $(\mathbf{x}(\mu), \mathbf{y}(\mu), \mathbf{s}(\mu))$, is called the *central path point* for μ, and the aggregate of all points, as μ ranges from 0 to ∞, is the *central path* of the LP problem. Note that $(\mathbf{x}(\mu), \mathbf{y}(\mu), \mathbf{s}(\mu))$ approaches an optimal solution as μ tends to 0.

In a generic path-following interior point method, one solves (3) approximately to obtain an "approximately centered" point $(\mathbf{x}, \mathbf{y}, \mathbf{s}, \mu)$ such that

$$\eta(\mathbf{x}, \mathbf{y}, \mathbf{s}, \mu) := \|SX\mathbf{e}/\mu - \mathbf{e}\| \leq \eta_0,$$

where $\eta_0 = 0.2$ in [4]. Then each iteration decreases μ towards zero, and for each new value of μ, the approximate point $(\mathbf{x}, \mathbf{y}, \mathbf{s}, \mu)$ to (3) is updated. The layered interior point (LIP) algorithm developed in [4] also fits into this framework; the number of times that μ is decreased is finite, and on the final step μ is driven to zero. In particular, the LIP algorithm uses a main loop iteration which consists of standard path-following interior point steps, plus an occasional use of a new type of step, the *layered least-squares* (LLS) step.

In this note we present a simplification to the LIP algorithm. In particular, we prove that Case II, one of the three cases in the algorithm, can never occur. This result simplifies the proof of the main theorem and improves certain bounds established in the paper. It is assumed throughout that the reader is familiar with [4] and has a copy available.

2 Three cases in the LIP algorithm

We now briefly describe one main loop iteration of the LIP algorithm. Given a current approximate central path point $(\mathbf{x}, \mathbf{y}, \mathbf{s}, \mu)$, partition the index set $\{1, \ldots, n\}$ into p *layers* $J_1, \ldots, J_p$ by using $\mathbf{s}$ and $\mathbf{x}$ as follows. Let

$$\Delta = \mathrm{diag}(\boldsymbol{\delta}) = \mu^{1/2} S^{1/2} X^{-1/2}.$$

Note that if $(\mathbf{x}, \mathbf{y}, \mathbf{s})$ is perfectly centered, then $SX\mathbf{e} = \mu\mathbf{e}$, i.e., $\boldsymbol{\delta} = \mathbf{s} = \mu X^{-1}\mathbf{e}$. Without losing generality, assume that these quantities have increasing order:

$$\delta_1 \leq \delta_2 \leq \cdots \leq \delta_n.$$

Let $g > 1$ be a "gap size" parameter, which will be defined later. Find the leftmost ratio-gap of size greater than g in the sorted quantities, i.e., find the smallest i such that $\delta_{i+1}/\delta_i > g$. Then let $J_1 = \{1, \ldots, i\}$. Now, put $i+1, i+2, \ldots$ in J_2, until another ratio-gap greater than g is encountered, and so on. Thus, the values of δ_i for quantities indexed by J_k for any k are within a factor of $g^{|J_k|} \leq g^n$ of each other, and are separated by more than a factor of g from quantities in J_{k+1}.

To formalize this, define

$$\theta_k = \max_{i \in J_k} \delta_i \quad \text{and} \quad \phi_k = \min_{i \in J_k} \delta_i.$$

The construction above ensures that for each k,

$$\theta_k < \phi_{k+1}/g. \quad \text{and} \quad \phi_k \leq \theta_k \leq g^{|J_k|}\phi_k \leq g^n \phi_k.$$

The main-loop iteration begins with a current iterate $(\mathbf{x}, \mathbf{y}, \mathbf{s}, \mu)$ that is feasible and approximately centered with $\eta(\mathbf{x}, \mathbf{y}, \mathbf{s}, \mu) \leq \eta_0$. Let $\delta\mathbf{x}^*$ be the minimizer of $PLLS(\mathbf{x})$, the primal LLS step, and $(\delta\mathbf{y}^*, \delta\mathbf{s}^*)$ the minimizer of $DLLS(\mathbf{s})$, the dual LLS step, with partition $J_1, \ldots, J_p$ and weights Δ as defined in [4]. Let

$$\mathbf{x}^* = \mathbf{x} + \delta\mathbf{x}^*, \quad \mathbf{y}^* = \mathbf{y} + \delta\mathbf{y}^*, \quad \text{and} \quad \mathbf{s}^* = \mathbf{s} + \delta\mathbf{s}^*.$$

For each layer in the LLS step, we define the scaled primal and dual residuals:

$$\epsilon_k^{\mathrm{P}} = \|\Delta_k(\delta\mathbf{x}_k^* + \mathbf{x}_k)\|_\infty/\mu \quad \text{and} \quad \epsilon_k^{\mathrm{D}} = \|\Delta_k^{-1}(\delta\mathbf{s}_k^* + \mathbf{s}_k)\|_\infty. \tag{4}$$

Define

$$\gamma_k^{\mathrm{P}} = \|\Delta_k \delta \mathbf{x}_k^*\|_\infty / \mu \quad \text{and} \quad \gamma_k^{\mathrm{D}} = \|\Delta_k^{-1} \delta \mathbf{s}_k^*\|_\infty. \tag{5}$$

There are three possibilities for one main-loop iteration of the algorithm, depending on the following cases.

Case I. There is a layer k such that

$$\gamma_k^{\mathrm{P}} \geq \frac{1}{4\sqrt{n}} \quad \text{and} \quad \gamma_k^{\mathrm{D}} \geq \frac{1}{4\sqrt{n}}. \tag{6}$$

In this case, for one main-loop iteration we take $n_1(A)$ ordinary path-following interior points steps. The formula for $n_1(A)$ is given by (24) below.

Case II. There is a layer k such that

$$\gamma_k^{\mathrm{P}} < \frac{1}{4\sqrt{n}} \quad \text{and} \quad \gamma_k^{\mathrm{D}} < \frac{1}{4\sqrt{n}}. \tag{7}$$

Case III. For every layer k, either

$$\gamma_k^{\mathrm{P}} \geq \frac{1}{4\sqrt{n}} \quad \text{and} \quad \gamma_k^{\mathrm{D}} < \frac{1}{4\sqrt{n}} \tag{8}$$

holds, or

$$\gamma_k^{\mathrm{P}} < \frac{1}{4\sqrt{n}} \quad \text{and} \quad \gamma_k^{\mathrm{D}} \geq \frac{1}{4\sqrt{n}} \tag{9}$$

holds. In this case, we define α_k for each layer as follows. If (8) holds, we take

$$\alpha_k = \min(1, 8\epsilon_k^{\mathrm{P}}\sqrt{n}) \tag{10}$$

else if (9) holds, we take

$$\alpha_k = \min(1, 8\epsilon_k^{\mathrm{D}}\sqrt{n}). \tag{11}$$

Then we define

$$\bar{\alpha} = \max\{\alpha_1, \ldots, \alpha_p\}. \tag{12}$$

Now we take a step defined by the primal and dual LLS directions; we compute a new feasible iterate

$$\mathbf{x}^+ = \mathbf{x} + (1-\bar{\alpha})\delta\mathbf{x}^* = \bar{\alpha}\mathbf{x} + (1-\bar{\alpha})\mathbf{x}^*,$$

$$\mathbf{y}^+ = \mathbf{y} + (1-\bar{\alpha})\delta\mathbf{y}^* = \bar{\alpha}\mathbf{y} + (1-\bar{\alpha})\mathbf{y}^*,$$

and

$$\mathbf{s}^+ = \mathbf{s} + (1-\bar{\alpha})\delta\mathbf{s}^* = \bar{\alpha}\mathbf{s} + (1-\bar{\alpha})\mathbf{s}^* = A^T\mathbf{y}^+ - \mathbf{c}.$$

Note that if (8) holds for k and $8\epsilon_k^{\mathrm{P}}\sqrt{n} \geq 1$, then the minimizer in (10) is 1, and thus $\bar{\alpha} = 1$ also. Thus, in this case $(\mathbf{x}^+, \mathbf{y}^+, \mathbf{s}^+) = (\mathbf{x}, \mathbf{y}, \mathbf{s})$.

On the other hand, if $\bar{\alpha} = 0$, this is the termination of our algorithm: $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{s}^*)$ is an optimal solution pair for the primal and dual as proved in [4].

If the termination criterion is not satisfied, we set $\mu^+ = \mu\bar{\alpha}$. Theorem 9 of [4] claims that $\eta(\mathbf{x}^+, \mathbf{y}^+, \mathbf{s}^+, \mu^+) < 0.65$. Then we take two Newton steps with this fixed μ^+ to restore the proximity to 0.2 by Lemma 1 of [4]. Note that if $\bar{\alpha} = 1$, then no restoration is needed.

3 Case II can never occur

We use the additional notation described in [4]. Let

$$\mathbf{x}^{[k]} = (\mathbf{0}; \ldots; \mathbf{0}; \mathbf{x}_k; \mathbf{0}; \ldots; \mathbf{0}) \quad \text{and} \quad \mathbf{s}^{[k]} = (\mathbf{0}; \ldots; \mathbf{0}; \mathbf{s}_k; \mathbf{0}; \ldots; \mathbf{0}),$$

where subvector $\mathbf{x}_k$ and $\mathbf{s}_k$ contain all x_i and all s_i in layer k, respectively. Define $\delta\mathbf{x}^{*[k]}$ to be the primal LLS minimizer of $PLLS(\mathbf{x}^{[k]})$ and $(\delta\mathbf{y}^{*[k]}, \delta\mathbf{s}^{*[k]})$ to be the dual LLS minimizer of $DLLS(\mathbf{s}^{[k]})$. Hereinafter we refer to $(\delta\mathbf{y}^{*[k]}, \delta\mathbf{s}^{*[k]})$ as the "focused" LLS step because they focus attention on what happens at the kth layer. Note that we use these vectors in our analyses only, and we do not need them in our algorithm. Recall that $\delta\mathbf{x}^*$ and $\delta\mathbf{s}^*$ are LLS minimizers for $PLLS(\mathbf{x})$ and $DLLS(\mathbf{s})$, respectively and are computed by our algorithm. We refer to these as the "original" LLS step. The first lemma shows that focused LLS step is a good approximation in the kth layer to the original LLS step for the appropriate scaled norm.

Lemma 1 *(Lemma 18 in [4]) For each k,*

$$\|\Delta_k(\delta\mathbf{x}_k^* - \delta\mathbf{x}_k^{*[k]})\| \leq \mu(\bar{\chi}_A + 1)\sqrt{1+\eta_0} \cdot \sqrt{n}/g \tag{13}$$

and

$$\|\Delta_k^{-1}(\delta\mathbf{s}_k^* - \delta\mathbf{s}_k^{*[k]})\| \leq \bar{\chi}_A\sqrt{1+\eta_0} \cdot \sqrt{n}/g. \tag{14}$$

See [4] for the definition of the condition number $\bar{\chi}_A$ of A.

Recall that the standard affine-scaling primal-dual interior point step $(\delta\mathbf{x}, \delta\mathbf{y}, \delta\mathbf{s})$ satisfies the equation $S\delta\mathbf{x} + X\delta\mathbf{s} = -XS\mathbf{e}$. The next lemma shows that the focused LLS step also satisfies this equation in the kth layer.

Lemma 2 *For a particular layer k, the subvector of the focused LLS step $(\delta\mathbf{x}_k^{*[k]}, \delta\mathbf{s}_k^{*[k]})$ satisfies the following equation:*

$$S_k\delta\mathbf{x}_k^{*[k]} + X_k\delta\mathbf{s}_k^{*[k]} = -X_kS_k\mathbf{e}. \tag{15}$$

Proof. We reintroduce the notation used in the proof of Lemma 5 of [4], namely, we let Ξ denote the $n \times n$ diagonal matrix, defined as follows. For $j = 1, \ldots, p$, for $i \in J_j$, the (i,i) entry of Ξ is 2^j. Fix some $t > 0$ and let $\delta\mathbf{s}$ denote the solution to the weighted least squares problem of minimizing $\|\Delta^{-1}\Xi^{-t}(\mathbf{s}^{[k]} + \delta\mathbf{s})\|$ subject to $\delta\mathbf{s} \in \mathcal{R}(A^T)$, that is $\delta\mathbf{s} = A^T\delta\mathbf{y}$ for some $\delta\mathbf{y}$. We can write out the normal equations for this least-squares problem:

$$A\Delta^{-2}\Xi^{-2t}A^T\delta\mathbf{y} = -A\Delta^{-2}\Xi^{-2t}\mathbf{s}^{[k]}$$

If we define

$$\mathbf{v} = \Delta^{-2}\Xi^{-2t}A^T\delta\mathbf{y} + \Delta^{-2}\Xi^{-2t}\mathbf{s}^{[k]} \tag{16}$$

then $A\mathbf{v} = \mathbf{0}$. Rewriting (16),

$$-\mathbf{v} + \mu^{-1}XS^{-1}\Xi^{-2t}A^T\delta\mathbf{y} = -\mu^{-1}XS^{-1}\Xi^{-2t}\mathbf{s}^{[k]}.$$

Define $\delta\mathbf{x} = -2^{2kt}\mu\mathbf{v}$; substitute and multiply the preceding equation through by $2^{2kt}\mu SX^{-1}\Xi^{2t}$:

$$\begin{aligned} SX^{-1}\Xi^{2t}\delta\mathbf{x} + 2^{2kt}\cdot A^T\delta\mathbf{y} &= -2^{2kt}\cdot\mathbf{s}^{[k]} \\ &= -2^{2kt}\cdot SX^{-1}\mathbf{x}^{[k]} \\ &= -SX^{-1}\Xi^{2t}\mathbf{x}^{[k]}. \end{aligned} \tag{17}$$

and $A\delta\mathbf{x} = \mathbf{0}$. The preceding system of equations is precisely the Lagrange multiplier condition for the solution to the weighted-least squares problem of finding $\delta\mathbf{x}$ to minimize $\|\Delta\Xi^t(\delta\mathbf{x} + \mathbf{x}^{[k]})\|$ subject to $A\delta\mathbf{x} = \mathbf{0}$. This solution satisfies, according to (17),

$$SX^{-1}\Xi^{2t}\delta\mathbf{x} + 2^{2kt}\delta\mathbf{s} = -SX^{-1}\Xi^{2t}\mathbf{x}^{[k]}$$

Restricting this equation to subscripts indexed by layer k, and dividing through by 2^{2kt} yields

$$S_k X_k^{-1} \delta \mathbf{x}_k + \delta \mathbf{s}_k = -S_k X_k^{-1} \mathbf{x}_k = -S_k \mathbf{e}$$

(recall that $\mathbf{x}_k^{[k]} = \mathbf{x}_k$). Taking the limit as $t \to \infty$, we know that $\delta \mathbf{x} \to \delta \mathbf{x}^{*[k]}$ and $\delta \mathbf{s} \to \delta \mathbf{s}^{*[k]}$ by Lemmas 4 and 5 of [4]. Therefore the preceding equation holds for the LLS solutions:

$$S_k X_k^{-1} \delta \mathbf{x}_k^{*[k]} + \delta \mathbf{s}_k^{*[k]} = -S_k \mathbf{e}.$$

Multiplying through by X_k yields (15). ∎

We now derive two important corollaries from (15).

Corollary 1 *Case II of the LIP algorithm can never occur, if we choose*

$$g \geq 6(\bar{\chi}_A + 1)\sqrt{1 + \eta_0} \cdot \sqrt{n}. \tag{18}$$

Proof. If we multiply (15) by by $X_k^{-1/2} S_k^{-1/2} \mu^{-1/2}$ we can rewrite it as

$$\Delta_k^{-1} \delta \mathbf{s}_k^{*[k]} + \Delta_k \delta \mathbf{x}_k^{*[k]} / \mu = X_k^{1/2} S_k^{1/2} \mathbf{e} / \mu^{1/2}. \tag{19}$$

Now we substitute (13) and (14) to replace the focused LLS step by the original LLS step

$$\| \Delta_k^{-1} \delta \mathbf{s}_k^* + \Delta_k \delta \mathbf{x}_k^* / \mu - X_k^{1/2} S_k^{1/2} \mathbf{e} / \mu^{1/2} \| \leq 2(\bar{\chi}_A + 1)\sqrt{1 + \eta_0}\sqrt{n}/g.$$

Switching to the ∞-norm,

$$\| \Delta_k^{-1} \delta \mathbf{s}_k^* + \Delta_k \delta \mathbf{x}_k^* / \mu - X_k^{1/2} S_k^{1/2} \mathbf{e} / \mu^{1/2} \|_\infty \leq 2(\bar{\chi}_A + 1)\sqrt{1 + \eta_0}\sqrt{n}/g.$$

Next, we use the fact that

$$\begin{aligned} \| X_k^{1/2} S_k^{1/2} \mathbf{e} / \mu^{1/2} - \mathbf{e} \|_\infty &\leq 1 - \sqrt{1 - \eta_0} \\ &\leq 0.6 \eta_0 \end{aligned}$$

since $\eta_0 \leq 0.2$. Thus,

$$\| \Delta_k^{-1} \delta \mathbf{s}_k^* + \Delta_k \delta \mathbf{x}_k^* / \mu - \mathbf{e} \|_\infty \leq 0.6\eta_0 + 2(\bar{\chi}_A + 1)\sqrt{1 + \eta_0}\sqrt{n}/g.$$

If g is chosen as in (18), then

$$0.6\eta_0 + 2(\bar{\chi}_A + 1)\sqrt{1+\eta_0}\sqrt{n}/g \leq 0.5,$$

i.e.

$$\|\Delta_k^{-1}\delta\mathbf{s}_k^* + \Delta_k\delta\mathbf{x}_k^*/\mu - \mathbf{e}\|_\infty \leq 0.5.$$

This implies by the triangle inequality (since $\|\mathbf{e}\|_\infty = 1$) that

$$\|\Delta_k^{-1}\delta\mathbf{s}_k^*\|_\infty + \|\Delta_k\delta\mathbf{x}_k^*\|_\infty/\mu \geq 0.5.$$

Examining (5), we see that $\gamma_k^{\mathrm{P}} + \gamma_k^{\mathrm{D}} \geq 0.5$. This implies that (7) can never hold, i.e., there is never a layer in Case II. ∎

The second corollary shows a relation between ϵ_k^{P} and γ_k^{D} and a similar relation in the dual case.

Corollary 2 *We have the following lower bounds on ϵ_k^{P} and ϵ_k^{D}:*

$$\epsilon_k^{\mathrm{P}} \geq \gamma_k^{\mathrm{D}} - \frac{1}{8\sqrt{n}} \quad \textit{and} \quad \epsilon_k^{\mathrm{D}} \geq \gamma_k^{\mathrm{P}} - \frac{1}{8\sqrt{n}}, \tag{20}$$

if we choose

$$g \geq 16(\bar{\chi}_A + 1)\sqrt{1+\eta_0}\cdot n. \tag{21}$$

Proof. Observe that we can rewrite the right-hand side of (19) as

$$\begin{aligned}\Delta_k^{-1}\delta\mathbf{s}_k^{*[k]} + \Delta_k\delta\mathbf{x}_k^{*[k]}/\mu &= -\Delta_k^{-1}S_k\mathbf{e}\\ &= -\Delta_k^{-1}\mathbf{s}_k\end{aligned}$$

that is

$$\Delta_k^{-1}(\mathbf{s}_k + \delta\mathbf{s}_k^{*[k]}) + \Delta_k\delta\mathbf{x}_k^{*[k]}/\mu = \mathbf{0}.$$

Now we substitute (13) and (14) to obtain bounds for the original LLS step and switch to the ∞-norm

$$\|\Delta_k^{-1}(\mathbf{s}_k + \delta\mathbf{s}_k^*) + \Delta_k\delta\mathbf{x}_k^*/\mu\|_\infty \leq 2(\bar{\chi}_A + 1)\sqrt{1+\eta_0}\sqrt{n}/g$$

which implies by the triangle inequality that

$$\|\Delta_k^{-1}(\mathbf{s}_k + \delta\mathbf{s}_k^*)\|_\infty \geq \|\Delta_k\delta\mathbf{x}_k^*\|_\infty/\mu - 2(\bar{\chi}_A + 1)\sqrt{1+\eta_0}\sqrt{n}/g,$$

i.e.

$$\epsilon_k^{\mathrm{D}} \geq \gamma_k^{\mathrm{P}} - 2(\bar{\chi}_A + 1)\sqrt{1+\eta_0}\sqrt{n}/g.$$

If g is chosen as in (21), then the second term on the right is bounded above by $1/(8\sqrt{n})$, which proves the second inequality of (20). The first inequality is proved in a similar manner. ∎

4 A simplified Section 7

We now review crossover events. We start with a simple lemma:

Lemma 3 *(Lemma 12 of [4]) Consider the current approximately-centered iterate* $(\mathbf{x}, \mathbf{y}, \mathbf{s}, \mu)$ *in Algorithm LIP. Suppose* $i, j \in \{1, \ldots, n\}$ *are two indices such that* $i \in J_1 \cup \cdots \cup J_k$ *and* $j \in J_k \cup \cdots \cup J_p$ *for some* $k \in \{1, \ldots, p\}$*. Then*

$$s_i(\mu) \leq 3g^n s_j(\mu).$$

A crossover is defined as follows (see Definition 13 of [4]):

Definition. Given an LP problem in primal-dual form, and given a current approximately centered iterate $(\mathbf{x}, \mathbf{y}, \mathbf{s}, \mu)$, we say that the 4-tuple $(\mu', \mu, i \in N, j \in B)$ *crossover event* occurs if $j \succeq i$, i.e., for some $k \geq 1$

$$i \in (J_1 \cup \cdots \cup J_k) \quad \text{and} \quad j \in (J_k \cup \cdots \cup J_p)$$

and the algorithm reduces μ to μ', $0 < \mu' < \mu$, such that,

$$s_i(\mu'') \geq 5g^n s_j(\mu'').$$

In other words, for every approximately centered point $(\mathbf{x}'', \mathbf{y}'', \mathbf{s}'', \mu'')$, $\mu'' \in (0, \mu']$, $i \succ j$ meaning that i is in a layer strictly higher than that containing j.

In simple terms, (μ', μ, i, j) specifies a crossover event provided that $s_j \in B$ is on the same order as, or is much larger than $s_i \in N$ for an approximately centered iterate $(\mathbf{x}, \mathbf{y}, \mathbf{s}, \mu)$, but $s'_j \ll s'_i$ for every approximately centered point $(\mathbf{x}', \mathbf{y}', \mathbf{s}', \mu')$, and furthermore, the relation $s''_j \ll s''_i$ remains in force for every approximately centered point $(\mathbf{x}'', \mathbf{y}'', \mathbf{s}'', \mu'')$ where $\mu'' \in (0, \mu']$. Thus, for a specific pair of $(i \in N, j \in B)$, the cross-over event (μ', μ, i, j) can happen at most once for some $0 \leq \mu' < \mu \leq \mu^0$, if the algorithm generates a sequence approximate centers (x, y, s, μ) where μ is strictly decreasing.

The following lemma is unchanged from [4]:

Lemma 4 *(Lemma 17 of [4]) Given a point* $(\mathbf{x}, \mathbf{y}, \mathbf{s}, \mu)$ *in Algorithm LIP, let* ϵ_k^{P} *and* ϵ_k^{D} *be defined by (4). Then there is an index* $i \in J_1 \cup \cdots \cup J_k$ *such that for all* $\mu' \in (0, \mu]$,

$$s_i(\mu') \geq \frac{\epsilon_k^{\mathrm{D}} \phi_k}{(\bar{\chi}_A + 1) n^{1.5}}. \tag{22}$$

There is also an index $j \in J_k \cup \cdots \cup J_p$ such that for all $\mu' \in (0, \mu]$,

$$x_j(\mu') \geq \frac{\mu \epsilon_k^{\mathrm{P}}}{\bar{\chi}_A \theta_k n^{1.5}}. \tag{23}$$

In the following proof, we use the result that a standard primal-dual path-following interior-point algorithm can reduce μ to μ' $(< \mu)$ in $c_0\sqrt{n}\log(\mu/\mu')$ iterations, while maintaining proximity η_0 to the central path. The constant c_0 depends on η_0 and on the particular algorithm but has 10 as an upper bound (see, e.g., [2]) for $\eta_0 = 0.2$.

Theorem 1 *Consider one main loop iteration of Algorithm LIP. (If Case III holds for the iteration, assume further that $\bar{\alpha} > 0$, i.e., the algorithm does not terminate.) This iteration causes a crossover event to take place. In other words, if μ denotes the central path parameter at the beginning of the iteration and μ' at the end, then there exist two indices $i \in N, j \in B$ such that (μ', μ, i, j) is a crossover event.*

Proof. Recall that Corollary 1 states that Case II never occurs. Therefore, the proof breaks down into two cases, based on whether Case I or Case III holds.

Case I holds for the iteration. In this case, γ_k^{D} and γ_k^{P} are both greater than or equal to $1/(4\sqrt{n})$ for some layer k. For this k, by (20) we conclude that $\epsilon_k^{\mathrm{P}} \geq 1/(8\sqrt{n})$ and $\epsilon_k^{\mathrm{D}} \geq 1/(8\sqrt{n})$. Next, we apply Lemma 4 to conclude that there is an $i \in J_1 \cup \cdots \cup J_k$ such that for all $\mu' \in (0, \mu]$,

$$s_i(\mu') \geq \frac{\phi_k}{8(\bar{\chi}_A + 1)n^2}, \quad i \in N.$$

There is also an index $j \in J_k \cup \cdots \cup J_p$ such that for all $\mu' \in (0, \mu]$,

$$x_j(\mu') \geq \frac{\mu}{8\bar{\chi}_A \theta_k n^2}, \quad j \in B.$$

Since $x_j(\mu')s_j(\mu') = \mu'$, we can rewrite the preceding equation as

$$s_j(\mu') \leq \frac{8\bar{\chi}_A \theta_k n^2 \mu'}{\mu}.$$

Therefore, a crossover event takes place when μ' is chosen sufficiently small so that

$$\frac{8\bar{\chi}_A\theta_k n^2\mu'}{\mu}\cdot(5g^n)=\frac{\phi_k}{8(\bar{\chi}_A+1)n^2},$$

i.e.,

$$\begin{aligned}\mu' &= \frac{\mu\phi_k}{320\bar{\chi}_A(\bar{\chi}_A+1)\theta_k n^4 g^n}\\ &\geq \frac{\mu}{320\bar{\chi}_A(\bar{\chi}_A+1)n^4 g^{2n}}.\end{aligned}$$

Thus, a crossover event takes place after

$$n_1(A)=c_0\sqrt{n}(2\log\bar{\chi}_A+4\log n+2n\log g+\log 320) \tag{24}$$

interior point iterations.

Case III holds for the iteration, and $\bar{\alpha}>0$. Let k be the layer such that $\bar{\alpha}=\alpha_k$ in the definition LLS step. For now, assume that for this layer k, (8) is satisfied. The other case is similar. In this case, $\alpha_k=\min(1,8\epsilon_k^{\mathrm{P}}\sqrt{n})$. By taking the reciprocal of (23), we arrive at a $j\in J_k\cup\cdots\cup J_p$ such that for all $\mu'\in(0,\mu]$,

$$\begin{aligned}s_j(\mu') &\leq \frac{\mu'\bar{\chi}_A\theta_k n^{1.5}}{\mu\epsilon_k^{\mathrm{P}}}\\ &\leq \frac{8\mu'\bar{\chi}_A\theta_k n^2}{\mu\bar{\alpha}}, \quad j\in B.\end{aligned}$$

On the other hand, by (8), $\gamma_k^{\mathrm{P}}\geq\frac{1}{4\sqrt{n}}$ so $\epsilon_k^{\mathrm{D}}\geq\frac{1}{8\sqrt{n}}$ by (20). Therefore, by (22) we find an $i\in J_1\cup\cdots J_k$ such that for all $\mu'\in(0,\mu]$,

$$s_i(\mu')\geq\frac{\phi_k}{8(\bar{\chi}_A+1)n^2},\quad i\in N.$$

Thus, a crossover event takes place when

$$\frac{8\mu'\bar{\chi}_A\theta_k n^2}{\mu\bar{\alpha}}=\frac{\phi_k}{5g^n\cdot 8\cdot(\bar{\chi}_A+1)n^2},$$

i.e.

$$\frac{\mu'}{\mu\bar{\alpha}}=\frac{\phi_k}{320g^n\bar{\chi}_A(\bar{\chi}_A+1)n^4\theta_k}\geq\frac{1}{320g^{2n}\bar{\chi}_A(\bar{\chi}_A+1)n^4}.$$

Since the LLS step decreases the central path parameter to $\mu\bar{\alpha}$, the number of small steps is the number required to reduce the parameter from $\mu\bar{\alpha}$ to μ' and cause a crossover is

$$n_3(A) = c_0\sqrt{n}(2n\log g + 2\log(\bar{\chi}_A + 1) + 4\log n + \log 320). \tag{25}$$

∎

Note that we no longer need Lemmas 19 and 20 in Section 7 of [4], and the bounds for g, $n_1(A)$ and $n_3(A)$ are all reduced from those in [4]. According to Gonzaga and Lara [1], expression $(\bar{\chi}_A+1)$ appeared in the above bounds can be replaced by $\bar{\chi}_A$. Moreover, Megiddo, Mizuno and Tsuchiya [3] recently developed a modification of the LIP algorithms that does not need to know g or $\bar{\chi}_A$ in advance in order to compute the LLS step but still retains all the theoretical features of the algorithm.

References

[1] C. C. Gonzaga and Hugo J. Lara, *A note on properties of condition numbers.* Department of Mathematics, Federal University of Santa Catarina, Cx. Postal 5210, 88040-970 Florianopolis, SC, Brazil, 1996.

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

[3] N. Megiddo, S. Mizuno and T. Tsuchiya, *A modified layered-step interior-point algorithm for linear programming.* IBM Technical Report, IBM Research Lab, San Jose, CA, 1996.

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