Simultaneous Data Perturbations and Analytic Center Convergence

A. Holder*

May 14, 2002

Abstract

The central path is an infinitely smooth parameterization of the non-negative real line, and its convergence properties have been investigated since the middle 1980s. However, the central "path" followed by an infeasible-interior-point method relies on three parameters instead of one, and is hence a surface instead of a path. The additional parameters are included to allow for simultaneous perturbations in the cost and right-hand side vectors. This paper provides a detailed analysis of the perturbed central path that is followed by infeasible-interior-point methods, and we characterize when such a path converges. We develop a set (Hausdorff) convergence property and show that the central paths impose an equivalence relation on the set of admissible cost vectors. We conclude with a technique to test for convergence under arbitrary, simultaneous data perturbations.

Keywords: Interior Point Methods, Sensitivity Analysis, Central Path, Linear Programming

---

* Trinity University Mathematics, San Antonio, TX, USA. Research conducted at Trinity University and The University of Mississippi. Partial support through ONR grant N00014-01-1-0917.

1 Introduction

Interior point algorithms have "revolutionized" the field of mathematical programming [24], and a class of these algorithms, known as path-following interior point algorithms, follow the *central path* toward the optimal set. The central path has been studied extensively, and instead of citing the numerous articles on the subject, we direct interested readers to the three texts of Roos, Terlaky, and Vial [17], Wright [25], and Ye [26], each of which contains an extensive bibliography and a complete development of the central path.

With the amount of literature that studies the central path, one may perceive that there is little left to understand. However, this is not the case, especially in semidefinite optimization, where the general convergence of the central path has only recently been established [9]. One of the main goals of this paper is to characterize the convergence of a central "path" that depends on multiple parameters. Several researchers have investigated such convergence [1, 10, 14, 15], but none of these works completely characterized the convergence of the perturbed central path followed by many interior-point algorithms. We approach the problem as a sensitivity analysis question, and our analysis provides both a characterization of convergence, which subsequently provides an insight into algorithm design, and information about the stability of solutions. Another strength of our analysis is that it is relatively simple, requiring only an understanding of real analysis and linear programming (the down side is that the notation is a bit cumbersome). In related work, Yildrim has investigated perturbed central paths in semidefinite optimization to gain sensitivity information [27, 28].

Consider the primal and dual linear programs

$$(LP)\quad \max\{cx : Ax = b, x \geq 0\} \quad \text{and} \quad (LD)\quad \min\{yb : yA + s = c, s \geq 0\}, \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$ has full row rank, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, and y, s, and c are row vectors. The primal and dual feasible regions are denoted by $\mathcal{P}$ and $\mathcal{D}$, respectively, and their strict interiors are $\mathcal{P}^o = \{x \in \mathcal{P} : x > 0\}$ and $\mathcal{D}^o = \{(y, s) \in \mathcal{D} : s > 0\}$. The primal and dual optimal sets are $\mathcal{P}^*$ and $\mathcal{D}^*$. We assume throughout that Slater's interiority condition holds —i.e. $\mathcal{P}^o \neq \emptyset$ and $\mathcal{D}^o \neq \emptyset$. The necessary and sufficient conditions for optimality are

$$Ax = b,\ x \geq 0,\ yA + s = c,\ s \geq 0,\ x_i s_i = 0,\ i = 1, 2, \ldots, n.$$

The central path is formed by replacing the complementarity constraint, $x_i s_i = 0$, with $x_i s_i = \mu > 0$. The fact that A has full row rank implies that for each positive μ there is a unique solution, denoted $(x(\mu), y(\mu), s(\mu))$, to the system

$$Ax = b,\ x \geq 0,\ yA + s = c,\ s \geq 0,\ x_i s_i = \mu,\ i = 1, 2, \ldots, n. \tag{2}$$

An important observation is that the equations in (2) are the necessary and sufficient Lagrange conditions for the penalized linear programs

$$\min\left\{cx - \mu\sum_{i=1}^{n}\ln(x_i) : x \in \mathcal{P}^o\right\} \quad \text{and} \quad \max\left\{yb + \mu\sum_{i=1}^{n}\ln(s_i) : (y, s) \in \mathcal{D}^o\right\}. \tag{3}$$

The logarithmic barrier function in these programs is unique, in that it is the only barrier function that yields the Lagrange conditions in (2) [13]. The logarithmic barrier function is also used to define the *analytic center* of a bounded polyhedron in the following way. Let $\mathcal{S} = \{x : Ax = b, x \geq 0\}$ be a bounded polyhedron, and let I index the components of x that are positive for some feasible element —i.e. $I = \{i : x_i > 0 \text{ for some } x \in \mathcal{S}\}$. The analytic center of $\mathcal{S}$ is the unique optimizer of

$$\max\left\{\sum_{i\in I}\ln(x_i) : x \in \mathcal{S}, x_i > 0, i \in I\right\}.$$

The analytic centers of $\mathcal{P}$ and $\mathcal{D}$ are denoted by $\bar{x}$ and $(\bar{y}, \bar{s})$, provided that either $\mathcal{P}$ or $\mathcal{D}$ is bounded. Frisch [4] and Huard [11] were the first to develop algorithms using analytic centers, and Sonnevend re-introduced this concept to the mathematical programming community in [18, 19, 20, 21, 22, 23].

A result first proven by McLinden [12] is that the central path converges to an optimal analytic center as $\mu \downarrow 0$. (Note: We distinguish between a $\downarrow$ and a $\rightarrow$, the former indicating that the limit is approached from above.) To make this precise, we first define the *optimal partition*, denoted by $(B|N)$, as follows

$$B = \{i : x_i > 0 \text{ for some } x \in \mathcal{P}^*\} \quad \text{and} \quad N = \{1, 2, 3, \ldots, n\} \backslash B.$$

Allowing a set subscript on a vector (or matrix) to be the subvector (or submatrix) comprised of the coordinates (or columns) corresponding to the elements in the set, we have that the optimal partition characterizes the optimal sets,

$$\begin{aligned}
\mathcal{P}^* &= \{x \in \mathcal{P} : x_N = 0\} = \{x : A_B x_B = b, x_B \geq 0, x_N = 0\} \text{ and} \\
\mathcal{D}^* &= \{(y, s) \in \mathcal{D} : s_B = 0\} = \{(y, s) : yA_B = c_B, yA_N + s_N = c_N, s_N \geq 0, s_B = 0\}.
\end{aligned}$$

It is well known that the strict interiors of the primal and dual feasible regions being non-empty is equivalent to both $\mathcal{P}^*$ and $\mathcal{D}^*$ being bounded [17]. The *central solution*, written (x^*, y^*, s^*), is the analytic center of $\mathcal{P}^*$ and $\mathcal{D}^*$, which means that x^* and (y^*, s^*) are the unique optimal solutions to

$$\max\left\{\sum_{i \in B} \ln(x_i) : x \in \mathcal{P}^*, x_B > 0\right\} \quad \text{and} \quad \max\left\{\sum_{i \in N} \ln(s_i) : (y, s) \in \mathcal{D}^*, s_N > 0\right\}.$$

What McLinden showed in 1980 was that the central path converges to (x^*, y^*, s^*) as $\mu \downarrow 0$, a result that is stated in Theorem 1.1.

Theorem 1.1 (McLinden [12]) *We have that*

$$\lim_{\mu \downarrow 0}(x(\mu), y(\mu), s(\mu)) = (x^*, y^*, s^*).$$

Furthermore, if $\mathcal{P}$ *is bounded,* $\lim_{\mu \to \infty} x(\mu) = \bar{x}$, *and if* $\mathcal{D}$ *is bounded,* $\lim_{\mu \to \infty}(y(\mu), s(\mu)) = (\bar{y}, \bar{s})$.

Originally, interior point algorithms assumed the existence of a strictly feasible primal and dual element. However, subsequent interior point algorithms allowed infeasible starting points, with the idea being to start with any (x^0, y^0, s^0) such that both x^0 and s^0 are positive, and define the following primal and dual residuals,

$$r_b = Ax^0 - b \;\text{ and }\; r_c = y^0 A + s^0 - c. \tag{4}$$

These residuals are scaled and added to b and c in (2) to obtain

$$Ax = b + \rho\, r_b,\; x \geq 0,\; yA + s = c + \tau\, r_c,\; s \geq 0,\; x_i s_i = \mu,\; i = 1, 2, \ldots, n. \tag{5}$$

For $\rho = \tau = 1$, (x^0, y^0, s^0) is a strictly feasible solution. The problem is that unless the residuals are zero, the right-hand side and cost vector are different from those of the original problem. So, infeasible-interior-point algorithms start with the perturbed data $b + \rho\, r_b$ and $c + \tau\, r_c$, and then decrease ρ and τ to zero while decreasing μ to zero. However, this means that the central path no longer relies on the single parameter μ, but the three parameters of μ, ρ, and τ. Unfortunately, an example in [10] shows that convergence is not guaranteed as μ, ρ and τ decrease to zero.

Explaining the convergence behavior of $(x(\mu), y(\mu), s(\mu))$ under data perturbations falls under the auspices of sensitivity analysis, and this is precisely the perspective from which we approach the problem. Because we are interested in how the central path relies on b and c, we extend our notation so that $(x(\mu, b, c), y(\mu, b, c), s(\mu, b, c))$ is the unique solution to the equations in (2). We point out that because $x(\mu, b, c)$ is the optimizer of the first math program in (3), we have for any positive α that $x(\mu, b, c) = x(\alpha\mu, b, \alpha c)$ (simply multiply the objective function by α). Similarly, $(y(\mu, b, c), s(\mu, b, c)) = (y(\alpha\mu, \alpha b, c), s(\alpha\mu, \alpha b, c))$ for $\alpha > 0$. For the data b and c, the central path, primal central path, and dual central path are respectively,

$$\begin{aligned}
CP_{(b,c)} &\equiv \{(x(\mu, b, c), y(\mu, b, c), s(\mu, b, c)) : \mu > 0\}, \\
PCP_{(b,c)} &\equiv \{x(\mu, b, c) : \mu > 0\}, \text{ and} \\
DCP_{(b,c)} &\equiv \{(y(\mu, b, c), s(\mu, b, c)) : \mu > 0\}.
\end{aligned}$$

In general, we consider sequences b^k and c^k, the use of which allows for arbitrary, simultaneous, and independent perturbations in b and c. Obviously, these data perturbations encompass the linear changes found in (5). Because x

y and s no longer depend on a single parameter, we are technically dealing with a surface and not a path. However, for intuitive and geometric concerns, we refer to a *perturbed central path* and choose sequences $x(\mu^k, b^k, c^k)$ from $PCP_{(b^k,c^k)}$.

As we shall see, allowing nonlinear perturbations in the cost coefficients significantly increases the difficulty of characterizing the convergence of the perturbed central path, and we often deal with linear changes. When this is the case, we let $b^k = b(\rho^k) = b + \rho^k \delta b$ and $c^k = c(\tau^k) = c + \tau^k \delta c$, where the direction vectors δb and δc are understood. Other notational extensions are described in Table 1.

Notation	**Explanation**	**Notation**	**Explanation**
$\mathcal{P}_b$	primal feasible region	$\mathcal{D}_c$	dual feasible region
$\mathcal{P}_b^o$	strict interior of $\mathcal{P}_b$	$\mathcal{D}_c^o$	strict interior of $\mathcal{D}_c$
$\mathcal{P}^*_{(b,c)}$	primal optimal set	$\mathcal{D}^*_{(b,c)}$	dual optimal set
$(\mathcal{P}^*_{(b,c)})^o$	strict interior of $\mathcal{P}^*_{(b,c)}$	$(\mathcal{D}^*_{(b,c)})^o$	strict interior of $\mathcal{D}^*_{(b,c)}$
$\bar{x}(b)$	analytic center of $\mathcal{P}_b$	$(\bar{y}(c), \bar{s}(c))$	analytic center of $\mathcal{D}_c$
$x^*(b,c)$	analytic center of $\mathcal{P}^*_{(b,c)}$	$(y^*(b,c), s^*(b,c))$	analytic center of $\mathcal{D}^*_{(b,c)}$
$(B(b,c)\|N(b,c))$	optimal partition		

Table 1: Notation that accounts for the dependence on b and c

All scalar sequences are in $\mathbb{R}^*_+ = \{\nu \in \mathbb{R} : \nu \geq 0\} \cup \{\infty\}$, which means that every scalar sequence has a cluster point (one of which may be ∞). The row, column, and null spaces of a matrix are denoted by $\text{row}(A)$, $\text{col}(A)$, and $\text{null}(A)$, and the projection of v onto the vector space W is denoted by $\text{proj}_W v$. The capitalization of a vector indicates the diagonal matrix formed from the vector. So, X is a diagonal matrix whose diagonal components are $x_1, x_2, \ldots, x_n$. The vector e is the all ones vector, where length is decided by the context of its use. The standard Big-O, o, Ω, and Θ notations are used [16]. Other notation is consistent with that found in the *Mathematical Programming Glossary* [5].

We have three primary goals for this paper. First, we characterize the convergence of $x(\mu^k, b^k, c + \tau^k \delta c)$ as $\mu^k \downarrow 0$, $b^k \to b$, and $\tau^k \downarrow 0$ by providing necessary and sufficient conditions on (μ^k, b^k, τ^k). Notice that nonlinear perturbations in b are allowed (but only linear changes in c). This result completely describes the convergence of the perturbed central path followed by all infeasible-path-following-interior-point algorithms. Second, we provide a set convergence result for the perturbed center path. This result shows that while the sequence $x(\mu^k, b^k, c+\tau^k \delta c)$ may not converge, the sequence of perturbed central paths does converge. Third, we remove the restriction that the perturbation in c must be linear and develop a process to calculate the limit of $x(\mu^k, b^k, c^k)$.

Before we begin, we point out that partial solutions are found in the literature. In [14], Mizuno, Todd, and Ye provide necessary conditions for the cluster points of the perturbed central path to be contained in the interior of the optimal set and the boundary of the optimal set. Bonnans and Potra [1] consider the case of a single shifted center within a specific algorithm environment for the horizontal linear complementarity problem. However, these results do not permit independent changes in b and c because the single parameter that is used controls the perturbation in both b and c. Monteiro and Tsuchiya [15] show that $x^*(\mu^k, b, c + \mu^k \delta c)$ converges as $\mu^k \downarrow 0$, but as in [1] this analysis relies on a single parameter. Holder, Strum, and Zhang [10] show that for any positive η, $x(\eta\mu^k, b + \rho^k \delta b, c + \mu^k \delta c)$ converges as $(\mu^k, \rho^k) \downarrow 0$ and that $x(\mu^k, b + \rho^k \delta b, c)$ converges as $(\mu^k, \rho^k) \downarrow 0$. Moreover, they prove that so long as $\tau^k = o(\mu^k)$ that $x(\mu^k, b + \rho^k \delta b, c + \tau^k \delta c)$ converges as $(\mu^k, \rho^k, \tau^k) \downarrow 0$. The results in [10] and [15] provide the actual limit when convergence is guaranteed. As one can see, there are many **sufficient** conditions that guarantee the convergence of $x(\mu^k, b + \rho^k \delta b, c + \tau^k \delta c)$. Our goal is different in that we characterize the convergence of $x(\mu^k, b^k, c+\tau^k \delta c)$ by providing **necessary** and **sufficient** conditions. A strength of our analysis is that we explain the entire set of cluster points of $x(\mu^k, b^k, c + \tau^k \delta c)$.

2 Preliminary Results

This section contains foundational material for subsequent sections, and several of the results in this section are simple to prove. While many of these results are used in the literature, some proofs are not readily available, and we include such proofs for completeness. If a result is proven in another article, we simply cite the article. Readers familiar with the central path literature will feel comfortable browsing through the notation and results of this section.

We begin with a study of the data that we are allowed to operate over. We say that b and c are *admissible* if the strict interiors of the primal and dual are nonempty. The admissible data sets are denoted by

$$\begin{aligned} \mathcal{G} &\equiv \{(b,c) \in \mathbb{R}^m \times \mathbb{R}^n : \mathcal{P}_b^o \neq \emptyset \, , \, \mathcal{D}_c^o \neq \emptyset\}, \\ \mathcal{G}^1 &\equiv \{b \in \mathbb{R}^m : \mathcal{P}_b^o \neq \emptyset\}, \text{ and} \\ \mathcal{G}^2 &\equiv \{c \in \mathbb{R}^n : \mathcal{D}_c^o \neq \emptyset\}. \end{aligned}$$

Our definition of admissible does not correspond with the traditional definition of admissible, which means that (*LP*) and (*LD*) have finite optimal solutions. Our definition is more restrictive because only data for which $\mathcal{P}_b^o$ and $\mathcal{D}_c^o$ are not empty is included. The first result shows that $\mathcal{G}$ is open, which subsequently implies that arbitrarily small perturbations of b and c remain admissible.

Theorem 2.1 *$\mathcal{G}$ is an open set.*

Proof: Let $(\hat{b}, \hat{c}) \in \mathcal{G}$. Then, there exists $\hat{x}$ and $(\hat{y}, \hat{s})$ such that $A\hat{x} = \hat{b}$, $\hat{x} > 0$, $\hat{y}A + \hat{s} = \hat{c}$, and $\hat{s} > 0$. Let U be an open set in $\mathbb{R}^n$ that contains $\hat{x}$ and has the property that $x \in U$ implies $x > 0$. Since the rank of A is m, the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m : x \to Ax$ is onto. Furthermore, since T is a continuous mapping, the Open Mapping Theorem implies that $T(U)$ is open. Let $\epsilon = \min\{\hat{s}_i : i = 1, 2, \ldots, m\}$, and define $V = \{c : \|c - \hat{c}\| < \epsilon\}$. Then, $(\hat{b}, \hat{c}) \in T(U) \times V \subset \mathcal{G}$, and the result follows since $T(U) \times V$ is open. ■

If $x(\mu^k, b^k, c^k) \to \hat{x}$, we have that $b^k = Ax(\mu^k, b^k, c^k) \to A\hat{x}$, which means that the convergence of b^k is a necessary condition of the convergence of $x(\mu^k, b^k, c^k)$. As such, we make the following assumption throughout.

Assumption 1 *We assume throughout that (b,c) and (b^k, c^k) are in $\mathcal{G}$. Moreover, we assume that $b^k \to b$ (but we do* **not** *necessarily assume that $c^k \to c$).*

Also, for notational convenience we assume that when $(b^k, c^k) \to (b, c)$ that $(B|N)$ is the optimal partition for (b,c) —i.e. $(B|N) = (B(b,c)|N(b,c))$. The dependence that the optimal partition has on b and c is only indicated for the perturbed data b^k and c^k. Sonnevend [18] showed that $x(\mu, b, c)$ is an analytic function over $\mathbb{R}_{++} \times \mathcal{G}$ (where we abuse the notation so that the 2-tuple $(\mu, (b,c))$ is understood to be the 3-tuple (μ, b, c)). Hence,

$$\mu^0 > 0 \Rightarrow \lim_{(\mu^k, b^k, c^k) \to (\mu^0, b, c)} x(\mu^k, b^k, c^k) = x(\mu^0, b, c). \tag{6}$$

The next two results show that the primal objective function is either strictly decreasing along the central path or that the central path degenerates to a single element.

Theorem 2.2 (Fiacco and McCormick [3]) *For $0 < \mu^1 < \mu^2$, we have that $c \notin \text{row}(A)$ if, and only if,*

$$cx^*(b,c) < cx(\mu^1, b, c) < cx(\mu^2, b, c) < c\bar{x}(b).$$

Similarly, for $0 < \mu^1 < \mu^2$, we have that $b \neq 0$ if, and only if,

$$y^*(b,c)b > y(\mu^1, b, c)b > y(\mu^2, b, c)b > \bar{y}(c)b.$$

Theorem 2.3 (Roos, Terlaky, and Vial [17]) *The following are equivalent:*

1. *cx is constant on $\mathcal{P}_b$.*
2. *$x(\mu^1, b, c) = x(\mu^2, b, c)$, for all $0 < \mu^1 < \mu^2$.*
3. *$x(\mu^1, b, c) = x(\mu^2, b, c)$, for some $0 < \mu^1 < \mu^2$.*
4. *$c \in \text{row}(A)$.*

5. $s(\mu, b, c) = \mu s(1, b, c)$ *for all* $0 < \mu$.

An observation that we use later is that if $c \in \text{row}(A)$ and $(b, c) \in \mathcal{G}$, then $\mathcal{P}_b$ is bounded. This follows because $\mathcal{P}_b$ is bounded if, and only if, there does not exist δx such that $A\delta x = 0$, $\delta x \geq 0$, and $\delta x \neq 0$. From Gordon's Theorem of the alternative (a variant of Farkas' Lemma) this is the same as $\mathcal{P}_b$ is bounded if, and only if, there is a row vector y such that $yA > 0$. Suppose that $c \in \text{row}(A)$, so that $\hat{y}A = c$ for some $\hat{y}$. Then, for any positive μ we have that $0 < s(\mu, b, c) = c - y(\mu, b, c)A = (\hat{y} - y(\mu, b, c))A$, and hence $\mathcal{P}_b$ is bounded.

We now direct our attention towards linear perturbations. Recall that for the understood directions of change δb and δc we defined $b(\rho)$ to be $b + \rho\delta b$ and $c(\tau)$ to be $c + \tau\delta c$. Directions of change for which the optimal partition is invariant for sufficiently small ρ and τ are of particular interest, and we define

$$\begin{aligned} \mathcal{H}(b,c) &= \{(\delta b, \delta c) : \text{there exists } \tilde{\rho} > 0 \text{ and } \tilde{\tau} > 0 \text{ such that for all } 0 \leq (\rho, \tau) < (\tilde{\rho}, \tilde{\tau}), \\ &\quad (B(b(\rho), c(\tau))|N(b(\rho), c(\tau))) = (B(b,c)|N(b,c))\}, \\ \mathcal{H}^1(b,c) &= \{\delta b : (\delta b, 0) \in \mathcal{H}(b,c)\}, \text{ and} \\ \mathcal{H}^2(b,c) &= \{\delta c : (0, \delta c) \in \mathcal{H}(b,c)\}. \end{aligned}$$

Properties of these sets are found in [6] and [7]. The next lemma shows that the optimal partition characterizes $\mathcal{H}(b,c)$, $\mathcal{H}^1(b,c)$, and $\mathcal{H}^2(b,c)$.

Lemma 2.1 *We have that* $\mathcal{H}^1(b,c) = \text{col}(A_B)$ *and that* $\mathcal{H}^2(b,c) = \{\delta c \in \mathbb{R}^n : \delta c_B \in \text{row}(A_B)\}$.

Proof: The partition $(B|N)$ is optimal for the right hand side $b(\rho)$ if, and only if, the following system is consistent

$$A_B x_B = b(\rho),\ x_B > 0,\ yA_B = c_B, \quad \text{and} \quad yA_N < c_N.$$

If $\delta b \in \text{col}(A_B)$, there exists x' such that $A_B(\rho x') = \rho\delta b$. Since $x^*_B(b,c) - \rho x'$ is positive for sufficiently small ρ, the above conditions remain consistent for arbitrarily small ρ. Hence, $\text{col}(A_B) \subseteq \mathcal{H}^1(b,c)$. If the optimal partition is invariant for sufficiently small ρ, then there exists $x_B(\rho)$ such that $A_B x_B(\rho) = b(\rho)$. Since $A_B(x_B(\rho) - x^*_B) = \rho\delta b$, we have that δb is in $\text{col}(A_B)$.

The argument for $\mathcal{H}^2(b,c)$ is similar, the difference being that the optimality conditions are

$$A_B x_B = b,\ yA_B = c_B(\tau),\ yA_N < c_N(\tau),\ x_B > 0.$$

■

The remainder of this section is concerned with establishing the existence of limits. Lemmas 2.2 and 2.4 provide bounds so that sequences have cluster points, and Lemma 2.3 and Theorem 2.4 use these bounds to establish limits. Consider the level set

$$\mathcal{L}(b, c, M) = \{(x, y, s) \in \mathcal{P}_b \times \mathcal{D}_c : sx \leq M\}.$$

The next Lemma shows that the union over k of the level sets $\mathcal{L}(b^k, c^k, M)$ is bounded, provided that c^k is bounded. The level set argument is similar to Theorem I.4 in [17] and Lemma 4.2 in [10], the differences being that c^k need not converge and independent, arbitrary perturbations in the right-hand side and the cost coefficients are allowed (Theorem I.4 does not permit data perturbations and Lemma 4.2 allows only linear changes in b and c that converge).

Lemma 2.2 *If* c^k *is bounded, then for* $M \geq 0$ *we have that* $\bigcup_k \mathcal{L}(b^k, c^k, M)$ *is bounded.*

Proof: Let $M \geq 0$ and $\mu^0 > 0$. Also, let $x^k = x(\mu^0, b^k, c^k)$ and $s^k = s(\mu^0, b^k, c^k)$. Then, for any $x \in \mathcal{P}_{b^k}$ and $(y, s) \in \mathcal{D}_{c^k}$, we have that $x^k - x \in \text{null}(A)$, $s^k - s \in \text{row}(A)$, and

$$0 = (s^k - s)(x^k - x) = s^k x^k - sx^k - s^k x + sx. \tag{7}$$

So, for any $(x, y, s) \in \mathcal{L}(b^k, c^k, M)$ we have that

$$s^k_i x_i \leq s^k x + sx^k = s^k x^k + sx \leq s^k x^k + M.$$

Since $s^k > 0$ and $s^k x^k = \mu^0 n$, we have that $x_i \le (M+\mu^0 n)/s^k_i$. A similar argument shows that $s_i \le (M+\mu^0 n)/x^k_i$. Since y relates to s in a one-to-one, linear fashion, we have for each k that $\mathcal{L}(b^k, c^k, M)$ is bounded.

To establish that $\cup_k \mathcal{L}(b^k, c^k, M)$ is bounded, we first show that $x(\mu^0, b^k, c^k)$ and $s(\mu^0, b^k, c^k)$ are $\Omega(1)$. Suppose for the sake of attaining a contradiction that there is a subsequence $(\mu^0, b^{k_j}, c^{k_j})$ such that $x_i(\mu^0, b^{k_j}, c^{k_j}) \downarrow 0$ for some i. Since c^{k_j} is bounded, it contains a convergent subsequence, and we assume without loss of generality that $c^{k_j} \to c$. However, this provides a contradiction since from (6) we have that $x(\mu^0, b^{k_j}, c^{k_j}) \to x(\mu^0, b, c) > 0$. Hence, $x(\mu^0, b^k, c^k) = \Omega(1)$. An analogous argument shows that $s(\mu^0, b^k, c^k) = \Omega(1)$. We now have that there is a positive λ^1 and λ^2 such that $x(\mu^0, b^k, c^k) > \lambda^1$ and $s(\mu^0, b^k, c^k) > \lambda^2$. So,

$$x_i \le \frac{M + \mu^0 n}{s^k_i} < \frac{2(M + \mu^0 n)}{\lambda^2} \quad \text{and} \quad s_i \le \frac{M + \mu^0 n}{x^k_i} < \frac{2(M + \mu^0 n)}{\lambda^1}.$$

Since these bounds are independent of k, we have that $\bigcup_k \mathcal{L}(b^k, c^k, M)$ is bounded. ■

The statement in Lemma 2.2 that bounds $x(\mu^k, b^k, c^k)$ does not require the convergence of c^k, but only that c^k be bounded. From Lemma 2.2 we have that if $\mu^k \downarrow 0$ and c^k is bounded, then the sequence

$$(x(\mu^k, b^k, c^k), y(\mu^k, b^k, c^k), s(\mu^k, b^k, c^k))$$

has a cluster point. However, an example in [10] shows that these sequences need not converge, which means a straight forward extension of Theorem 1.1 is not available. The next lemma shows that x_N and s_B approach zero with μ.

Lemma 2.3 *If $\mu^k \downarrow 0$ and $c^k \to c$, we have that $x_N(\mu^k, b^k, c^k) \to 0$ and $s_B(\mu^k, b^k, c^k) \to 0$.*

Proof: Lemma 2.2 implies that $\left(x(\mu^k, b^k, c^k), y(\mu^k, b^k, c^k), s(\mu^k, b^k, c^k)\right)$ has a convergent subsequence, say

$$\lim_{i \to \infty} \left(x(\mu^{k_i}, b^{k_i}, c^{k_i}), y(\mu^{k_i}, b^{k_i}, c^{k_i}), s(\mu^{k_i}, b^{k_i}, c^{k_i})\right) = (\hat{x}, \hat{y}, \hat{s}).$$

Set $x^i = x(\mu^{k_i}, b^{k_i}, c^{k_i})$, $y^i = y(\mu^{k_i}, b^{k_i}, c^{k_i})$, and $s^i = s(\mu^{k_i}, b^{k_i}, c^{k_i})$. Since

$$\left.\begin{array}{rcl} Ax^i & = & b^{k_i}, \;\; x^i > 0 \\ y^i A + s^i & = & c^{k_i}, \;\; s^i > 0 \\ s^i x^i & = & n\mu^{k_i} \end{array}\right\} \Rightarrow \left\{\begin{array}{rcl} A\hat{x} & = & b, \;\; \hat{x} \ge 0 \\ \hat{y}A + \hat{s} & = & c, \;\; \hat{s} \ge 0 \\ \hat{s}\hat{x} & = & 0, \end{array}\right. \tag{8}$$

we have that $\hat{x} \in \mathcal{P}^*_{(b,c)} = \{x \in \mathcal{P}_b : x_N = 0\}$ and $(\hat{y}, \hat{s}) \in \mathcal{D}^*_{(b,c)} = \{(y, s) \in \mathcal{D}_{\bar{c}} : s_B = 0\}$, which proves the result. ■

If $\mu^k \downarrow 0$, $c^k \to c$, and $x(\mu^k, b^k, c^k) \to \hat{x}$, Lemma 2.3 identifies a subvector of $\hat{x}$ that is zero. Unfortunately, this is not necessarily the largest subvector of $\hat{x}$ that is zero, an issue that we address in Section 5.

The final objective of this section is to develop sufficient conditions for $x(\mu^k, b^k, c^k)$ to converge to the analytic center of a polytope, a result that relies on Lemmas 2.4 and 2.5.

Lemma 2.4 (Caron, Greenberg, and Holder [2]) *If $\mathcal{P}_b$ is bounded, $\bigcup_k \mathcal{P}_{b^k}$ is bounded.*

From Lemma 2.4 we have that a bounded polytope remains bounded under right-hand side perturbation. We now introduce the concept of *set convergence* [8] (typically called Hausdorff convergence), an idea that we use now to establish the existence of a particular sequence and later to show that the central path converges as a set. We say that a sequence of sets H^k converges to the set H if the following two conditions hold,

1. if $h^k \in H^k$ and $h^k \to h$, then h must be in H, and
2. if $h \in H$, then there exists $h^k \in H^k$ such that $h^k \to h$.

From [8] we know that $b^k \to b$ implies that $\mathcal{P}_{b^k} \to \mathcal{P}_b$, which is important because we need the result that elements within the strict interior of the feasible set may be approached by strictly positive elements. To see this that this is true, let $x \in \mathcal{P}_b^o$. Then, since $\mathcal{P}_{b^k} \to \mathcal{P}_b$, there is a sequence $x^k \in \mathcal{P}_{b^k}$ such that $x^k \to x$, and because x is positive, we have that x^k is positive for sufficiently large k. We state this fact in Lemma 2.5.

Lemma 2.5 *If x is in $\mathcal{P}_b^o$, there exists a sequence $x^k \in \mathcal{P}_{b^k}^o$ such that converges $x^k \to x$.*

The next theorem provides sufficient conditions for $x(\mu^k, b^k, c^k)$ to converge to the analytic center of a polytope.

Theorem 2.4 *Let $\mathcal{P}_b$ be bounded. Then, if the vector sequence c^k/μ^k is bounded and has the property that every cluster point is in* row(A), *we have that $x(\mu^k, b^k, c^k) \to \bar{x}(b)$.*

Proof: From Lemma 2.4 we have that $x(\mu^k, b^k, c^k)$ is bounded. So, there exists a subsequence such that

$$x(\mu^{k_i}, b^{k_i}, c^{k_i}) \to \hat{x} \quad \text{and} \quad c^{k_i}/\mu^{k_i} \to \hat{c}.$$

Let $x^i = x(\mu^{k_i}, b^{k_i}, c^{k_i})$, $y^i = y(\mu^{k_i}, b^{k_i}, c^{k_i})$, and $s^i = s(\mu^{k_i}, b^{k_i}, c^{k_i})$. Similar to (8), we have that $\hat{x} \in \mathcal{P}_b$. For any i, the necessary and sufficient conditions describing (x^i, y^i, s^i) are

$$Ax = b^{k_i}, \;\; x > 0, \;\; yA + s = c^{k_i}, \;\; s > 0, \text{ and } \;\; Sx = \mu^{k_i} e,$$

which means that

$$\begin{aligned} Ax^i &= b^{k_i}, \\ -\frac{y^i}{\mu^{k_i}} A &= (X^i)^{-1} - \frac{c^{k_i}}{\mu^{k_i}}, \\ x^i &> 0. \end{aligned} \tag{9}$$

From the full row rank of A, we have that

$$-\frac{y^i}{\mu^{k_i}} = \left((X^i)^{-1} - \frac{c^{k_i}}{\mu^{k_i}} \right) A^T \left(AA^T \right)^{-1}.$$

We prove that $\hat{x}$ is positive, so that this last equality implies the sequence $\{y^i/\mu^{k_i}\}$ has a limit. Then, since $\hat{c}$ is in row(A), equation (9) implies that $(X^i)^{-1}$ is in row(A). Subsequently, we have that there is a $\hat{y}$ such that

$$A\hat{x} = b, \;\; \hat{y}A = \hat{X}^{-1}, \;\; \hat{x} > 0,$$

and because these are the necessary and sufficient conditions describing $\bar{x}(b)$, the result is established once we show that $\hat{x}$ is positive.

From Lemma 2.5 there is a sequence, $\tilde{x}^i \in \mathcal{P}_{b^{k_i}}^o$ such that $\tilde{x}^i \to \tilde{x} \in \mathcal{P}_b^o$. The optimality of x^i implies that

$$\frac{c^{k_i}}{\mu^{k_i}} x^i - \sum_{j=1}^{n} \ln(x_j^i) \le \frac{c^{k_i}}{\mu^{k_i}} \tilde{x}^i - \sum_{j=1}^{n} \ln(\tilde{x}_j^i),$$

which is equivalent to

$$\sum_{j=1}^{n} \ln(\tilde{x}_j^i) \le \frac{c^{k_i}}{\mu^{k_i}} (\tilde{x}^i - x^i) + \sum_{j=1}^{n} \ln(x_j^i). \tag{10}$$

Since $\tilde{x}^i$ is $\Omega(1)$, the left-hand side of this last inequality is bounded below. Suppose for the sake of attaining a contradiction that as $i \to \infty$, $x_j^i \to 0$, for some j. The boundedness of x^i implies that $\sum_{j=1}^{n} \ln(x_j^i) \to -\infty$. Hence, the inequality in (10) implies that $(c^{k_i}/\mu^{k_i})(\tilde{x}^i - x^i) \to \infty$. However, since $\hat{c} \in \text{row}(A)$ and $(\tilde{x} - \hat{x}) \in \text{null}(A)$, we have that

$$(c^{k_i}/\mu^{k_i})(\tilde{x}^i - x^i) \to \hat{c}(\tilde{x} - \hat{x}) = 0.$$

So, no such j exists, and $\hat{x} > 0$. ■

Corollary 2.1 *If $\mathcal{P}_b$ is bounded, $c^k \to c$, and $\mu^k \to \infty$, then $x(\mu^k, b^k, c^k) \to \bar{x}(b)$.*

Proof: The proof follows immediately from Theorem 2.4 because $c^k/\mu^k \to 0 \in \text{row}(A)$. ■

While only providing sufficient conditions for the convergence of $x(\mu^k, b^k, c^k)$, Theorem 2.4 is used in the next section to develop necessary and sufficient conditions. We point out that none of μ^k, c^k, or c^k/μ^k had to converge for $x(\mu^k, b^k, c^k)$ to converge. Because of this, Theorem 2.4 highlights the difficulty of allowing simultaneous perturbations in μ, b, and c.

3 Characterizing the Convergence of The Central Path Under Simultaneous Parameterization

The goal of this section is to develop necessary and sufficient conditions on $(\mu^k, b^k, c(\tau^k))$ so that $x(\mu^k, b^k, c(\tau^k))$ converges as $\mu^k \downarrow 0$ and $\tau^k \downarrow 0$, and we assume throughout this section that $\tau^k \downarrow 0$. These conditions are stated in Theorem 3.2, and they completely characterize the convergence of the perturbed central path followed by an infeasible-path-following-interior-point algorithm. In this section, we allow arbitrary perturbations in b and linear changes in c. The case of independent, arbitrary, nonlinear changes in both b and c is addressed in Section 5. The first lemma of this section shows that $c_B u$ is constant over the null space of A_B.

Lemma 3.1 *We have that $c_B u = 0$ for all $u \in \text{null}(A_B)$.*

Proof: Let $u \in \text{null}(A_B)$ and $(x_B^*, 0) \in (\mathcal{P}^*_{(b,c)})^0$. Then, there exists a positive α such that $(x_B^* + \alpha u, 0) \in (\mathcal{P}^*_{(b,c)})^0$. Since $c_B(x_B^* + \alpha u) = c_B x_B^*$, we have that $c_B u = 0$. ■

Lemma 3.1 is used to show that the objective function is constant on "cuts" of the feasible region, which are defined for any k and positive μ as

$$\mathcal{C}(\mu, k) = \{x_B : A_B x_B = b^k - A_N x_N(\mu, b^k, c(\tau^k)), x_B \geq 0\}.$$

$\mathcal{C}(\mu, k)$ is the sub-polyhedron of $\mathcal{P}_{b^k}$ formed by fixing x_N to be $x_N(\mu, b^k, c(\tau^k))$. Lemma 3.2 shows that $c_B x_B$ is constant on each $\mathcal{C}(\mu, k)$.

Lemma 3.2 *For any k and positive μ, $c_B x_B$ is constant on $\mathcal{C}(\mu, k)$. Consequently, $x_B(\mu, b^k, c(\tau^k))$ is the unique solution to*

$$\min \left\{ \tau^k \delta c_B x_B - \mu \sum_{i \in B} \ln(x_i) : A_B x_B = b^k - A_N x_N(\mu, b^k, c(\tau^k)),\ x_B > 0 \right\}. \quad (11)$$

Proof: By definition, $x(\mu, b^k, c(\tau^k))$ is the unique solution to

$$\min \left\{ cx + \tau^k \delta c x - \mu \sum_{i=1}^{n} \ln(x_i) : x \in (\mathcal{P}_{b^k})^o \right\}.$$

Holding the components of $x_N(\mu, b^k, c(\tau^k))$ constant, we have that $x_B(\mu, b^k, c(\tau^k))$ is the unique solution to

$$\min \left\{ c_B x_B + \tau^k \delta c_B x_B - \mu \sum_{i \in B} \ln(x_i) : A_B x_B = b^k - A_N x_N(\mu, b^k, c(\tau^k)),\ x_B > 0 \right\}.$$

So, the result follows once we show that $c_B x_B$ is constant on $\mathcal{C}(\mu, k)$. If the columns of A_B are linearly independent, the result is immediate because $\mathcal{C}(\mu, k)$ contains a single element. Otherwise, let x_B^1 and x_B^2 be in $\mathcal{C}(\mu, k)$. Then, $x_B^1 - x_B^2 \in \text{null}(A_B)$, and from Lemma 3.1 we have for all $\alpha \in [0, 1]$ that

$$c_B(x_B^1 + \alpha(x_B^2 - x_B^1)) = c_B x_B^1.$$

For $\alpha = 1$ we have that $c_B x_B^1 = c_B x_B^2$, which proves the result. ■

The fact that $x_B(\mu, b^k, c(\tau^k))$ is the unique optimal solution to the math program in (11) is paramount in our analysis. To aid our development, for any positive η we define $z_B(\eta, b, c_B)$ to be the unique solution of

$$\min\left\{c_B z_B - \eta \sum_{i \in B} \ln(z_i) : A_B z_B = b, z_B > 0\right\}, \tag{12}$$

which means that $\{z_B(\eta, b, c_B) : \eta > 0\}$ is the central path of the linear program

$$\min\left\{c_B z_B : A_B z_B = b, z_B \geq 0\right\}. \tag{13}$$

Because $\{z_B(\eta, b, c_B) : \eta > 0\}$ is a central path for a fixed b and c, $z(\eta, b, c_B)$ has a limit as $\eta \downarrow 0$, which is denoted by $z_B^*(b, c_B)$. The feasible region of the math program in (13) is equipotent to $\mathcal{P}^*_{(b,c)}$ (just remove x_N). Since (b, c) being in $\mathcal{G}$ implies that $\mathcal{P}^*_{(b,c)}$ is bounded, we have that the feasible region of (13) is bounded, and subsequently that $z_B(\eta, b, c_B)$ converges as $\eta \to \infty$ to the analytic center of $\{z_B : A_B z_B = b, z_B \geq 0\}$. Since $x_B^*(b, c)$ is this analytic center, we have that $\lim_{\eta \to \infty} z_B(\eta, b, c_B) = x_B^*(b, c)$. In addition to the convergence properties of $z_B(\eta, b, c_B)$, we have from Lemma 3.2 that

$$x_B(\mu, b^k, c(\tau^k)) = z_B(\mu/\tau^k, b^k - A_N x_N(\mu, b^k, c(\tau^k)), \delta c_B). \tag{14}$$

The following example illustrates the relationship between $x(\mu, b, c(\tau))$ and $z_B(\eta, b, \delta c_B)$.

Example 3.1 *Consider the linear program*

$$(LP^1) \quad \min\{x_3 : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, 0 \leq x_3 \leq 1\}.$$

Allowing x_4, x_5, and x_6 to be the slack variables, we have that the optimal partition is $(\{1,2,4,5,6\}|\{3\})$. Let $b^k = b$, so there is no right-hand side perturbation, and $\delta c = (1, 1/10, 0, 0, 0, 0)$, so $c^k = c + \tau^k \delta c = (\tau^k, \tau^k/10, 1, 0, 0, 0)$. Figure 1 illustrates four central paths associated with perturbations of (LP^1). The vertical line is the unperturbed central path for (LP^1), and the curve in the x_1, x_2-plane is the central path for

$$(LP^2) \quad \min\{\delta c_B x_B : x \in \mathcal{P}^*_{(b,c)}\} = \min\{x_1 + \frac{1}{10}x_2 : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, x_3 = 0\}.$$

The curve from $(1/2, 1/2, 1/2)$ to $(0, 0, 0)$ is the perturbed central path for $\tau^k = 1$, and hence, corresponds to the linear program

$$(LP^3) \quad \min\{x_1 + (1/10)x_2 + x_3 : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, 0 \leq x_3 \leq 1\}.$$

The plane passing through the feasible region is $\mathcal{C}(1, k)$, where τ^k is 1, and the curve on this sub-polyhedron is the central path of

$$(LP^4) \quad \min\{x_1 + (1/10)x_2 : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, x_3 = x_3(1, b, c(1))\}.$$

The x_1, x_2, x_4, x_5, and x_6 values of this central path form the z variables defined by (12). Notice that the only difference between (LP^2) and (LP^4) is the value of x_3. In (LP^2), x_3 is zero, and in (LP^4), x_3 is $x_3(1, b, c(1))$. This means that the central paths for (LP^2) and (LP^4) are the same except for the shift in x_3. Equation (14) shows how the shifted central path of (LP^4) intersects the perturbed central path of (LP^3).

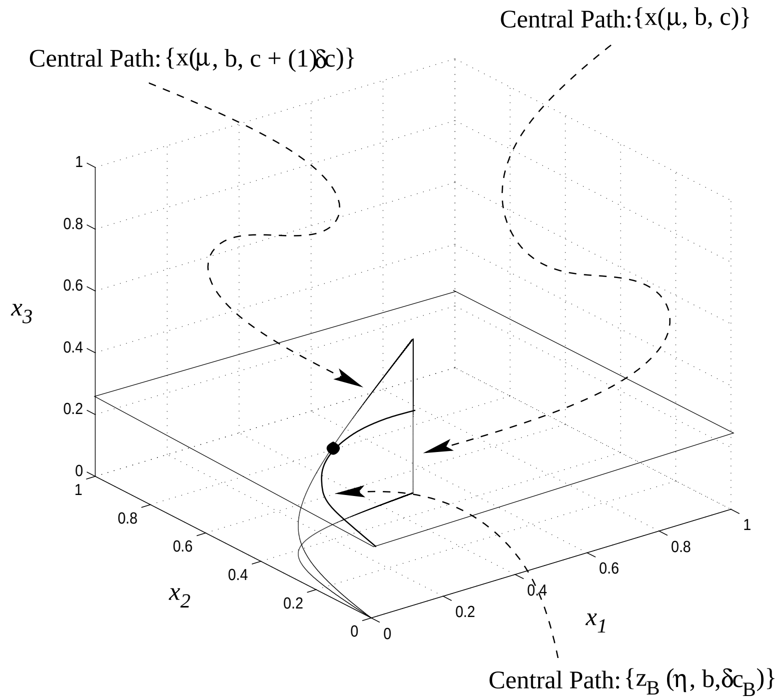

Figure 1: Four central paths associated with (LP^1) and how they intersect.

The equality in (14) is important because z has the perturbations in both b and c modeled as right-hand side perturbations —i.e. there is no perturbation of the cost vector δc_B that is used to describe z_B. This observation indicates that we need to understand the convergence properties of a central path under right-hand side perturbation. Lemma 3.3 states that the central solution is continuous with respect to b, and Lemma 3.4 shows that a perturbed central path converges to the analytic center of the unperturbed optimal set so long as there is no movement in c. We note that Lemma 3.4 is similar to Theorem 4.1 in [10], the difference being that our result allows arbitrary perturbations in b.

Lemma 3.3 (Caron, Greenberg, and Holder [2]) *The analytic center of a bounded polyhedron is a continuous function of the right-hand side. That is, if $b^k \to b$ and $\mathcal{P}_b$ is bounded, $\lim_{k\to\infty} \bar{x}(b^k) = \bar{x}(b)$ (NOTE: this result is true whether or not $\mathcal{P}_b^o$ is non-empty).*

We note that since the central solution $x^*(b,c)$ is the analytic center of the polytope $\mathcal{P}^*_{(b,c)}$, we have that $x^*(b,c)$ is a continuous function of b. This is stated in the following Corollary for future reference.

Corollary 3.1 *The central solution $x^*(b,c)$ is continuous with respect to the right-hand side b.*

Lemma 3.4 *If $\mu^k \downarrow 0$, we have that $x(\mu^k, b^k, c) \to x^*(b,c)$.*

Proof: From Lemma 2.3 we have that $x_N(\mu^k, b^k, c) \to 0$, and from Lemma 3.2 we have that $x_B(\mu^k, b^k, c)$ is the unique solution to

$$\max\left\{\sum_{i\in B} \ln(x_i) : A_B x_B = b^k - A_N x_N(\mu^k, b^k, c), x_B > 0\right\}.$$

This means that $x_B(\mu^k, b^k, c)$ is the analytic center of $\{x_B : A_B x_B = b^k - A_N x_N(\mu^k, b^k, c), x_B \geq 0\}$, and from Lemma 3.3 we have that this analytic center is a continuous function of $b^k - A_N x_N(\mu^k, b^k, c)$. Since $b - A_N x_N(\mu^k, b^k, c) \to b$, we have that $x_B(\mu^k, b^k, c)$ converges to the analytic center of $\mathcal{P}_b^* = \{x : A_B x_B = b, x_B \geq 0\}$. ■

We take a moment to summarize what we have. If μ^k has a positive limit, we have from (6) that $x(\mu^k, b^k, c(\tau^k))$ converges. The more difficult situation is if μ^k decreases to 0. From Lemma 2.3 we have that $x_N(\mu^k, b^k, c(\tau^k))$ decreases to zero as well. So, what is left to know is whether or not $x_B(\mu^k, b^k, c(\tau^k))$ converges. Since

$$x_B(\mu^k, b^k, c(\tau^k)) = z_B(\mu^k/\tau^k, b^k - A_N x_N(\mu^k, b^k, c(\tau^k)), \&_B),$$

we have from Lemma 3.4 that x_B converges so long as μ^k/τ^k and $b^k - A_N x_N(\mu^k, b^k, c(\tau^k))$ converge. Again, since $x_N(\mu^k, b^k, c(\tau^k))$ decreases to zero, we have that $b^k - A_N x_N(\mu^k, b^k, c(\tau^k)) \to b$. This means that $x_B(\mu^k, b^k, c(\tau^k))$ converges so long as μ^k/τ^k converges, a result that is stated in Theorem 3.1. This condition is "nearly" necessary and sufficient for the sequence $x(\mu^k, b^k, c(\tau^k))$ to converge, the problem being that if $\&$ is in $\mathcal{H}^2(b, c)$, then $x(\mu^k, b^k, c(\tau^k))$ may converge even though μ^k/τ^k does not converge.

Theorem 3.1 *Let $\tau^k \downarrow 0$ and $\mu^k > 0$ be such that $\mu^k \to \mu^0$. Then,*

$$\lim_{k\to\infty} x(\mu^k, b^k, c(\tau^k)) = \begin{cases} x(\mu^0, b, c) & \text{if } \mu^0 > 0 \\ x^*(b, c) & \text{if } \mu^0 = 0 \text{ and } \mu^k/\tau^k \to \infty \\ (z_B(\eta, b, d), 0) & \text{if } \mu^0 = 0 \text{ and } \mu^k/\tau^k \to \eta > 0 \\ (z_B^*(b, c), 0) & \text{if } \mu^0 = 0 \text{ and } \mu^k/\tau^k \to 0. \end{cases}$$

Proof: The case of $\mu^0 > 0$ is an immediate consequence of (6). Assume $\mu^0 = 0$. From Lemma 2.3 we have that $x_N(\mu^k, b^k, c(\tau^k)) \to 0$. Consider the situation of $\mu^k/\tau^k \to \eta > 0$. Since, μ^k/τ^k is bounded away from zero, we have from (6) that

$$x_B(\mu^k, b^k, c(\tau^k)) = z_B\left(\mu^k/\tau^k, b^k - A_N x_N(\mu^k, b^k, c(\tau^k)), \&\right) \to z_B(\eta, b, \&),$$

which establishes the third case. Suppose that $\mu^k/\tau^k \to 0$. We have from Lemma 3.4 that

$$x_B(\mu^k, b^k, c(\tau^k)) = z_B(\mu^k/\tau^k, b^k - A_N x_N(\mu^k, b^k, c(\tau^k)), \&) \to z_B^*(b, \&).$$

So, the fourth case is established. Lastly, suppose that $\mu^k/\tau^k \to \infty$. Then, $\&/(\mu^k/\tau^k) = \tau^k \&/\mu^k \to 0 \in \text{row}(A_B)$, and since $\mathcal{P}^*_{(b,c)}$ is bounded, we have from Theorem 2.4 that

$$x_B(\mu^k, b^k, c(\tau^k)) = z_B(\mu^k/\tau^k, b^k - A_N x_N(\mu^k, b^k, c(\tau^k)), \&) \to \bar{z}_B(b, c) = x_B^*(b, c).$$

■

As previously stated, the reason that the conditions in Theorem 3.1 are not necessary conditions is that if $\&$ is in $\mathcal{H}^2(b, c)$, then $x(\mu^k, b^k, c(\tau^k))$ may converge even if μ^k/τ^k does not. Lemmas 3.5 and 3.6 address this issue.

Lemma 3.5 *Let $\mu^k \downarrow 0$ and $\& \notin \mathcal{H}^2(b, c)$. Suppose that the sequence μ^k/τ^k does not converge. Then, if μ^{k_i}/τ^{k_i} and μ^{k_j}/τ^{k_j} are two convergent subsequences, we have that*

$$\lim_{i\to\infty} \mu^{k_i}/\tau^{k_i} \neq \lim_{j\to\infty} \mu^{k_j}/\tau^{k_j} \quad \Rightarrow \quad \lim_{i\to\infty} x(\mu^{k_i}, b^{k_i}, c(\tau^{k_i})) \neq \lim_{j\to\infty} x(\mu^{k_j}, b^{k_j}, c(\tau^{k_j})).$$

Proof: Without loss in generality, we assume that

$$\lim_{i\to\infty} \mu^{k_i}/\tau^{k_i} < \lim_{j\to\infty} \mu^{k_j}/\tau^{k_j}.$$

From Theorem 3.1 we have that

$$\lim_{i\to\infty} x(\mu^{k_i}, b^{k_i}, c(\tau^{k_i})) = \begin{cases} (z_B^*(b, \&), 0) & \text{if } \mu^{k_i}/\tau^{k_i} \to 0 \\ (z_B(\eta^1, b, \&), 0) & \text{if } \mu^{k_i}/\tau^{k_i} \to \eta^1 > 0 \end{cases}$$

and

$$\lim_{j\to\infty} x(\mu^{k_j}, b^{k_j}, c(\tau^{k_j})) = \begin{cases} (z_B(\eta^2, b, \delta c), 0) & \text{if} \quad \mu^{k_j}/\tau^{k_j} \to \eta^2 < \infty \\ x^*(b,c) & \text{if} \quad \mu^{k_j}/\tau^{k_j} \to \infty. \end{cases}$$

Since $\delta c \notin \mathcal{H}^2$, we have from Lemma 2.1 that $\delta c_B \notin \text{col}(A_B)$. The result follows because from Theorem 2.2 we have that for any $\eta^1 < \eta^2$ that

$$\delta c_B z_B^*(b, \delta c) < \delta c_B z_B(\eta^1, b, \delta c) < \delta c_B z_B(\eta^2, b, \delta c) < \delta c_B \bar{z}_B(b, c) = \delta c_B x_B^*(b, \bar{c}).$$

■

Lemma 3.6 *If $\delta c \in \mathcal{H}^2(b,c)$, we have for all positive η that*

$$\bar{z}_B(b, \delta c) = z_B(\eta, b, \delta c) = z_B^*(b, \delta c).$$

Proof: From Lemma 2.1 we have that $\delta c_B \in \text{row}(A_B)$, and from Theorem 2.3 we have that $z_B(\eta^1, b, \delta c) = z_B(\eta^2, b, \delta c)$, for all positive η^1 and η^2. Hence, for any positive η^0,

$$z_B^*(b, \delta c) = \lim_{\eta\downarrow 0} z(\eta, b, \delta c) = z_B(\eta^0, b, c) = \lim_{\eta\to\infty} z(\eta, b, \delta c) = \bar{z}_B(b, \delta c).$$

■

Theorem 3.2 states the necessary and sufficient conditions for the convergence of $x(\mu^k, b^k, c(\tau^k))$.

Theorem 3.2 *Let $\tau^k \downarrow 0$ and $\mu^k \downarrow 0$. If $\delta c \in \mathcal{H}^2(b,c)$, then $x(\mu^k, b^k, c(\tau^k)) \to x^*(b,c)$. Otherwise, $\delta c \notin \mathcal{H}^2(b,c)$, and $x(\mu^k, b^k, c(\tau^k))$ converges if, and only if, μ^k/τ^k converges.*

Proof: Suppose that $\delta c \in \mathcal{H}^2(b,c)$. From Lemma 2.3 we have that $x_N(\mu^k, b^k, c(\tau^k)) \to 0$. Also, from Lemma 3.6 we have that

$$x_B(\mu^k, b^k, c(\tau^k)) = z_B(\mu^k/\tau^k, b^k - A_N x_N(\mu^k, b^k, c(\tau^k)), \delta c) = z_B^*(b^k - A_N x_N(\mu^k, b^k, c(\tau^k)), \delta c).$$

From Lemma 3.3 we know that z_B^* is a continuous function of the right-hand side $b^k - A_N x_N(\mu^k, b^k, c(\tau^k))$. So,

$$\begin{aligned} \lim_{k\to\infty} x(\mu^k, b^k, c(\tau^k)) &= \lim_{k\to\infty} \left(z_B^*(b^k - A_N x_N(\mu^k, b^k, c(\tau^k)), \delta c), x_N(\mu^k, b^k, c(\tau^k)) \right) \\ &= (z_B^*(b, \delta c), 0) \\ &= x^*(b,c). \end{aligned}$$

Assume that $\delta c \notin \mathcal{H}^2(b,c)$. If μ^k/τ^k converges, Theorem 3.1 shows that $x(\mu^k, b^k, c(\tau^k))$ converges (and provides the limit). If μ^k/τ^k does not converge, this sequence has at least two cluster points, and hence, there are two convergent subsequences, say μ^{k_i}/τ^{k_i} and μ^{k_j}/τ^{k_j}, such that $\lim_{i\to\infty} \mu^{k_i}/\tau^{k_i} \neq \lim_{j\to\infty} \mu^{k_j}/\tau^{k_j}$. Theorem 3.1 implies that both

$$\lim_{i\to\infty} x(\mu^{k_i}, b^{k_i}, c(\tau^{k_i})) \quad \text{and} \quad \lim_{j\to\infty} x(\mu^{k_j}, b^{k_j}, c(\tau^{k_j}))$$

exist, and Lemma 3.5 implies that these limits are different. Hence, $x(\mu^k, b^k, c(\tau^k))$ does not converge. ■

We conclude this section by classifying the convergence of the perturbed central path followed by infeasible-path-following-interior-point algorithms. We require the dual counterpart of Theorem 3.2, which we state without proof.

Theorem 3.3 *Let $\mu^k \downarrow 0$, $\rho^k \downarrow 0$, and $c^k \to c$. If $\delta b \in \mathcal{H}^1(b,c)$, then $(y(\mu^k, b(\rho^k), c^k), s(\mu^k, b(\rho^k), c^k)) \to (y^*(b,c), s^*(b,c))$. Otherwise, $\delta b \notin \mathcal{H}^1(b,c)$, and $(y(\mu^k, b(\rho^k), c^k), s(\mu^k, b(\rho^k), c^k))$ converges if, and only if, μ^k/ρ^k converges.*

As mentioned in Section 1, the perturbed central path followed by infeasible-path-following-interior-point algorithms has linear perturbations in b and c, with the directions of change defined by the residuals. Table 2 shows the sequences whose convergence characterizes the convergence of the perturbed central central path.

	Cost Perturbation	
Right-Hand Side Perturbation	$\delta c \notin \mathcal{H}^2(b,c)$ $\Updownarrow$ $\delta c_B \notin \text{row}(A_B)$	$\delta c \in \mathcal{H}^2(b,c)$ $\Updownarrow$ $\delta c_B \in \text{row}(A_B)$
$\delta b \notin \mathcal{H}^1(b,c) \Leftrightarrow \delta b \notin \text{col}(A_B)$	μ^k/ρ^k & μ^k/τ^k	μ^k/ρ^k
$\delta b \in \mathcal{H}^2(b,c) \Leftrightarrow \delta b \in \text{col}(A_B)$	μ^k/τ^k	
	Must Converge	

Table 2: Let δb and δc be defined by the residuals in (4). Depending on whether or not δb is in $\mathcal{H}^1(b,c)$ and δc is in $\mathcal{H}^2(b,c)$, we have that the convergence of the indicated sequences is required for, and guarantees, the convergence of $(x(\mu^k, b(\rho^k), c(\tau^k)), y(\mu^k, b(\rho^k), c(\tau^k)), s(\mu^k, b(\rho^k), c(\tau^k)))$.

4 Set Convergence

The objective of this section is to establish a set (Hausdorff) convergence property for the perturbed central path, and Theorem 4.1 shows how the central path behaves as a set under simultaneous changes in b and c, provided that the change in c is linear. We illustrate the set convergence result with the following example.

Example 4.1 *As in Example 3.1, consider the linear program*

$$\min\{x_3 : 0 \le x_1 \le 1,\ 0 \le x_2 \le 1,\ 0 \le x_3 \le 1\}.$$

Let x_4, x_5, and x_6 be the slack variables, $b^k = b$ (so there is no right-hand side perturbation), and $\delta c = (1/4, 1/2000, 0, 0, 0, 0)$. The central paths corresponding to b and $c(\tau^k)$, for $\tau^k = 1, 0.8, 0.6, 0.4, 0.2$, are shown in Figure 2. The vertical line is the central path for the unperturbed problem —i.e. the vertical line is $PCP_{(b,c)}$. The curve in the x_1 and x_2 plane is the central path for the linear program

$$\min\{1/4x_1 + 1/2000x_2 : 0 \le x_1 \le 1,\ 0 \le x_2 \le 1,\ x_3 = 0\} = \min\{\delta c_B x_B : x \in \mathcal{P}^*\}.$$

Observe that the perturbed central paths **converge** *to these two central paths.*

Example 4.1 indicates, and Theorem 4.1 proves, that the perturbed central paths converge to the union of two central paths. The first of these paths is $PCP_{(b,c)}$ —i.e. the central path of the unperturbed linear program. The second of these paths is denoted by $PCP^*_{(b,c,\delta c)}$ and corresponds to minimizing $\delta c x$ over the optimal face. Hence, $PCP^*_{(b,c,\delta c)}$ is defined by the linear program

$$\min\{\delta c_B x_B : A_B x_B = b, x_B \ge 0, x_N = 0\}.$$

The elements of $PCP^*_{(b,c,\delta c)}$ have the form of $(z_B(\eta, \delta c), 0)$, and hence $PCP^*_{(b,c,\delta c)}$ is equipotent to $\{z_B(\eta, b, \delta c) : \eta > 0\}$. The closure of $PCP_{(b,c)}$ is $\overline{PCP}_{(b,c)}$ and is either $PCP_{(b,c)} \cup \{x^*(b,c)\} \cup \{\bar{x}(b)\}$ or $PCP_{(b,c)} \cup \{x^*(b,c)\}$, depending on whether or not the feasible region is bounded. The closure of $PCP^*_{(b,c,\delta c)}$ is $\overline{PCP^*}_{(b,c,\delta c)} = PCP^*_{(b,c,\delta c)} \cup \{(z^*_B(b,c), 0)\} \cup \{(\bar{z}_B(b,c), 0)\}$.

Theorem 4.1 *If $\tau^k \downarrow 0$, we have that $\overline{PCP}_{(b^k, c(\tau^k))} \to \overline{PCP}_{(b,c)} \cup \overline{PCP^*}_{(b,c,\delta c)}$.*

Proof: We begin by establishing that

$$PCP_{(b^k, c(\tau^k))} \to \overline{PCP}_{(b,c)} \bigcup \overline{PCP^*}_{(b,c,\delta c)}.$$

Let $x^k \in PCP_{(b^k, c(\tau^k))}$ be such that $x^k \to \hat{x}$. Then, for each k there is a μ^k such that $x^k = x(\mu^k, b^k, c(\tau^k))$. Let μ^{k_i} be a convergent subsequence of μ^k (remember that ∞ is a possible cluster point). We consider three cases to

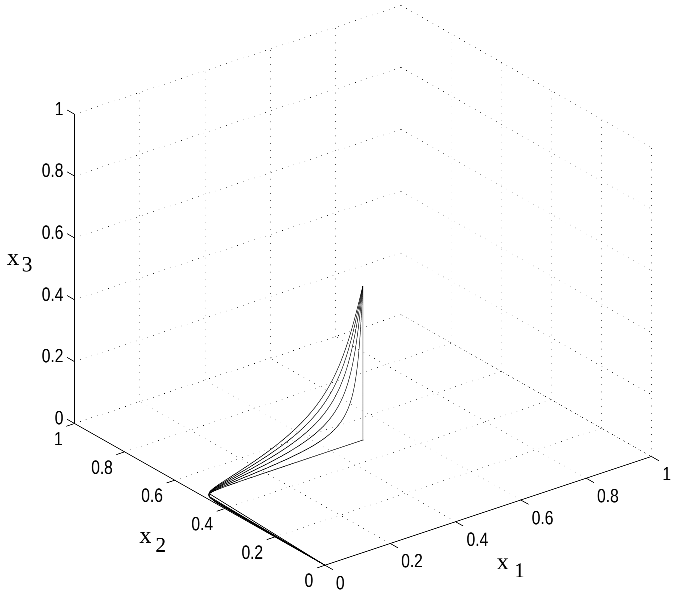

Figure 2: The central paths of the perturbed data converge to the union of two central paths.

show that $\hat{x} \in \overline{PCP}_{(b,c)} \bigcup \overline{PCP^*}_{(b,c,\delta c)}$.

Case 1: If $\mu^{k_i} \to \hat{\mu} > 0$, we have from (6) that

$$x(\mu^{k_i}, b^{k_i}, c(\tau^{k_i})) \to x(\hat{\mu}, b, c) = \hat{x} \in \overline{PCP}_{(b,c)}.$$

Case 2: Suppose that $\mu^{k_i} \downarrow 0$. If $\delta c \in \mathcal{H}^2$, Theorem 3.2 shows that

$$x(\mu^{k_i}, b^{k_i}, c(\tau^{k_i})) \to x^*(b, c) = \hat{x} \in \overline{PCP}_{(b,c)}.$$

Otherwise, $\delta c \notin \mathcal{H}^2$, and Theorem 3.2 shows that μ^{k_i}/τ^{k_i} must converge. From Theorem 3.1 we have that

$$x(\mu^{k_i}, b^{k_i}, c(\tau^{k_i})) \to \hat{x} = \begin{cases} x^*(b,c) & \text{if} \quad \mu^{k_i}/\tau^{k_i} \to \infty \\ (z_B(\eta, b, \delta c), 0) & \text{if} \quad \mu^{k_i}/\tau^{k_i} \to \eta > 0 \\ (z_B^*(b, \delta c), 0) & \text{if} \quad \mu^{k_i}/\tau^{k_i} \to 0. \end{cases}$$

Since $x^*(b,c) \in \overline{PCP}_{(b,c)}$, and both $(z_B(\eta, b, \delta c), 0)$ and $(z_B^*(b, \delta c), 0)$ are in $\overline{PCP^*}_{(b,c,\delta c)}$, we have that $\hat{x}$ is in $\overline{PCP}_{(b,c)} \cup \overline{PCP^*}_{(b,c,\delta c)}$.

Case 3: Suppose that $\mu^{k_i} \to \infty$. Then, $c(\tau^{k_i})/\mu^{k_i} \to 0 \in \text{row}(A)$. If we knew that $\mathcal{P}_b$ were bounded, we would have from Theorem 2.4 that

$$x(\mu^{k_i}, b^{k_i}, c(\tau^{k_i})) \to \hat{x} = \bar{x}(b) \in \overline{PCP}_{(b,c)}.$$

So, our goal in this case becomes to use the fact that $x(\mu^{k_i}, b^{k_i}, c(\tau^{k_i}))$ converges as $\mu^{k_i} \to \infty$ to show that $\mathcal{P}_b$ is bounded. Let $x^i = x(\mu^{k_i}, b^{k_i}, c(\tau^{k_i}))$, $y^i = y(\mu^{k_i}, b^{k_i}, c(\tau^{k_i}))$, and $s^i = s(\mu^{k_i}, b^{k_i}, c(\tau^{k_i}))$. From Gordon's Theorem of the alternative we have that $\mathcal{P}_b$ is bounded if, and only if, there is a row vector y such that $yA > 0$. For $j = 1, 2, \ldots, n$, we have that $x_j^i s_j^i = \mu^{k_i}$, $x_j^i \to \hat{x}_j$, and $\mu^{k_i} \to \infty$. Consequently, we have that $s_j^i \to \infty$. From the dual constraints we have that $(-y^i)A = s^i - c(\tau^{k_i}) \to \infty$, and hence the system $yA > 0$ is consistent. So, $\mathcal{P}_b$ is bounded.

At this point we have established that if $x^k \in PCP_{(b^k,c(\tau^k))}$ converges, then the limit of this sequence is in $\overline{PCP}_{(b,c)} \cup \overline{PCP^*}_{(b,c,\delta c)}$. We now show that any element in $\overline{PCP}_{(b,c)} \cup \overline{PCP^*}_{(b,c,\delta c)}$ is the limit of a sequence in $PCP_{(b^k,c(\tau^k))}$. Let x be in $\overline{PCP}_{(b,c)} \cup \overline{PCP^*}_{(b,c,\delta c)}$. Then, x is one of $\bar{x}(b)$ (if $\mathcal{P}_b$ is bounded), $x(\hat{\mu}, b, c)$ (for some positive $\hat{\mu}$), $x^*(b,c)$, $(z_B(\eta, b, \delta c), 0)$ (for some positive η), or $(z_B^*(b,\delta c), 0)$. From Theorem 2.4 and Lemma 2.3 we have for $\tau^k = 1/k$ that

$$\begin{aligned} &x(\hat{\mu}+1/k, b^k, c(\tau^k)) \to x(\hat{\mu}, b, c), \quad x(\sqrt{\tau^k}, b^k, c(\tau^k)) \to x^*(b,c), \\ &\quad x(\eta\tau^k, b^k, c(\tau^k)) \to (z_B(\eta, b, \delta c), 0), \quad x((\tau^k)^2, b^k, c(\tau^k)) \to (z_B^*(b, \delta c), 0), \text{ and} \\ &\quad x(k, b^k, c(\tau^k)) \to \bar{x}(b) \text{ (if } \mathcal{P}_b \text{ is bounded)..} \end{aligned}$$

Since all four of these sequences are in $PCP_{(b^k,c(\tau^k))}$, we have that

$$PCP_{(b^k,c(\tau))} \to \overline{PCP}_{(b,c)} \cup \overline{PCP^*}_{(b,c,\delta c)}.$$

What remains to be shown is that if the sequence $x^k \in \overline{PCP}_{(b^k,c(\tau^k))}$ converges and contains either $x^*(b^k, c(\tau^k))$, or in the case that $\mathcal{P}_b$ is bounded, $\bar{x}(b^k)$ infinitely many times, the limit of this sequence is in $\overline{PCP}_{(b,c)} \cup \overline{PCP^*}_{(b,c,\delta c)}$. If $\mathcal{P}_b$ is bounded, we have from Lemma 2.4 that $\mathcal{P}_{b^k}$ is bounded for sufficiently large k. Furthermore, Lemma 3.3 shows that $\bar{x}(b)$ is a continuous function of b. So, if $x^k \in \overline{PCP}_{(b^k,c(\tau^k))}$ contains $\bar{x}(b^k)$ infinitely many times and converges to $\hat{x}$, we have that $\hat{x} = \bar{x}(b) \in \overline{PCP}_{(b,c)}$. Suppose that $x^k \in \overline{PCP}_{(b^k,c(\tau^k))}$ converges to $\hat{x}$, and that this sequence contains $x^*(b^k, c(\tau^k))$ infinitely many times. Without loss in generality we assume that $x^k = x^*(b^k, c(\tau^k))$. First, because $(B|N)$ need not be the same as $(B(b^k, c(\tau^k))|N(b^k, c(\tau^k)))$, we do not automatically know that $x_N^k = x_N^*(b^k, c(\tau^k)) \to 0$ (Lemma 2.3 does not apply). However, x_N^k does converge to 0 as the following argument shows. Let $\varepsilon > 0$. For each k we have that

$$x_N^k = x_N^*(b^k, c(\tau^k)) = \lim_{\mu \downarrow 0} x_N(\mu, b^k, c(\tau^k)).$$

So, there is a $\hat{\mu}^k > 0$ such that $\mu \in (0, \hat{\mu}^k)$ implies that $\|x_N(\mu, b^k, c(\tau^k)) - x_N^*(b^k, c(\tau^k))\| < \varepsilon/2$. Choose $\mu^k \in (0, \hat{\mu}^k)$ so that $\mu^k \downarrow 0$. From Lemma 2.3 we have that $x_N(\mu^k, b^k, c(\tau^k)) \to 0$. Hence, there exists a natural number K, such that for $k \geq K$ we have that $\|x_N(\mu^k, b^k, c(\tau^k))\| < \varepsilon/2$. Hence, for $k \geq K$,

$$\|x_N(b^k, c(\tau^k))\| \leq \|x_N(b^k, c(\tau^k)) - x_N(\mu^k, b^k, c(\tau^k))\| + \|x_N(\mu^k, b^k, c(\tau^k))\| < \varepsilon.$$

So, $x_N^k = x_N^*(b^k, c(\tau^k)) \to 0$. Using this fact, Lemma 3.3 to establish the 5^{th} equality, and Lemma 3.4 to establish the 4^{th} equality, we have that

$$\begin{aligned} \hat{x}_B = \lim_{k\to\infty} x_B^*(b^k, c(\tau^k)) &= \lim_{k\to\infty} \left(\lim_{\mu\downarrow 0} x_B(\mu, b^k, c(\tau^k)) \right) \\ &= \lim_{k\to\infty} \left(\lim_{\mu\downarrow 0} z_B(\mu/\tau^k, b^k - A_N x_N(\mu, b^k, c(\tau^k)), \delta c) \right) \\ &= \lim_{k\to\infty} z_B^*(b^k - A_N x_N^*(b^k, c(\tau^k)), \delta c) \\ &= z_B^*(b, \delta c). \end{aligned}$$

Hence, we have that $x^k = x^*(b^k, c(\tau^k)) \to (z_B^*(b, \delta c), 0) \in \overline{PCP^*}_{(b,c,\delta c)}$, which completes the proof. ■

A corollary to Theorem 4.1 is that the perturbed central path is continuous over $\mathcal{H}^2(b,c)$, meaning that so long as $\delta c \in \mathcal{H}^2(b,c)$, $\overline{PCP}_{(b^k,c(\tau^k))} \to \overline{PCP}_{(b,c)}$. This follows because if $\delta c \in \mathcal{H}^2(b,c)$, we have from Lemma 3.6 that

$$\overline{PCP^*}_{(b,c,\delta c)} = \{(\bar{z}_B(b, \delta c), 0)\} = \{x^*(b,c)\} \subset \overline{PCP}_{(b,c)}.$$

This result is stated in the following corollary.

Corollary 4.1 *We have that if $\delta c \in \mathcal{H}^2(b,c)$ and $\tau^k \downarrow 0$, then $\overline{PCP}_{(b^k,c(\tau^k))} \to \overline{PCP}_{(b,c)}$.*

We conclude this section by showing why our results are stated from the primal perspective. This is because it is possible for b^k, $c(\tau^k)$, and $x(\mu^k, b^k, c(\tau^k))$ to converge, while the dual elements diverge. For example, suppose that $c \in \text{row}(A)$, which implies that

- $\mathcal{P}_b$ is bounded,
- $(B|N) = (\{1, 2, \ldots, n\}|\emptyset)$,
- $x(\mu^k, b^k, c(\tau^k)) = x_B(\mu^k, b^k, c(\tau^k)) = z_B(\mu^k/\tau^k, b^k, \delta c)$, and
- $x^*(b, c) = \bar{x}(b) = \bar{z}_B(b, \delta c)$.

Let $\tau^k \downarrow 0$ and μ^k be the sequence $1, 2, 1, 2, 1, 2, \ldots$. Then, $\mu^k/\tau^k \to \infty$, and we have from Corollary 2.1 that $x(\mu^k, b^k, c(\tau^k)) = z_B(\mu^k/\tau^k, b^k, \delta c) \to \bar{z}_B(b, \delta c) = x^*(b, c) = \bar{x}(b)$. However, Theorem 2.3 implies that the corresponding dual sequence $s(\mu^k, b^k, c(\tau^k))$ has the two cluster points of $s(1, b, c)$ and $s(2, b, c) = 2s(1, b, c)$. The problem here is that $s_i(\mu^k, b^k, c(\tau^k)) = \mu^k/x_i(\mu^k, b^k, c(\tau^k))$, and we see that the dual elements fail to converge because the sequence μ^k diverges. To guarantee the convergence of $s(\mu^k, b^k, c(\tau^k))$ one needs to guarantee the convergence of $\mu^k/x_i(\mu^k, b^k, c(\tau^k))$, $i = 1, 2, \ldots, n$ (which is not implied by the convergence of μ^k and $x(\mu^k, b^k, c(\tau^k))$). While Theorem 4.2 does not completely resolve this issue, it does show when the convergence of μ^k is guaranteed.

Theorem 4.2 *Let $\tau^k \downarrow 0$. Then, the convergence of $x(\mu^k, b^k, c(\tau^k))$ implies the convergence of μ^k if, and only if, $c \notin \text{row}(A)$.*

Proof: Assume that $c \in \text{row}(A)$. Then, as discussed on page 6, $\mathcal{P}_b$ is bounded. Let $\mu^k = 1, 2, 1, 2, \ldots$ and $\tau^k = 1/k$. Then, $\mu^k/\tau^k \to \infty$, and as just discussed, $x(\mu^k, b^k, c(\tau^k)) \to \bar{x}(b)$. Hence, the convergence of $x(\mu^k, b^k, \bar{c}_{\tau^k})$ cannot guarantee the convergence of μ^k.

Assume that $c \notin \text{row}(A)$, and suppose for the sake of attaining a contradiction that μ^k does not converge. Then there are subsequences, μ^{k_i} and μ^{k_j}, such that

$$0 \le \lim_{i \to \infty} \mu^{k_i} < \lim_{j \to \infty} \mu^{k_j} \le \infty.$$

If $\mu^{k_i} \to \mu^1 > 0$, we have from (6) that $x(\mu^{k_i}, b^{k_i}, c(\tau^{k_i})) \to x(\mu^1, b, c)$. From (6) and Corollary 2.1 we have that

$$x(\mu^{k_j}, b^{k_j}, c(\tau^{k_j})) \to \begin{cases} x(\mu^2, b, c) & \text{if} \quad \mu^{k_j} \to \mu^2 < \infty \\ \bar{x}(b, c) & \text{if} \quad \mu^{k_j} \to \infty. \end{cases}$$

However, Theorem 2.2 shows that $cx(\mu^1, b, c) < cx(\mu^2, b, c) < c\bar{x}(b, c)$, where the last inequality is included only when $\bar{x}$ exists. This is a contradiction since this implies that

$$\lim_{i \to \infty} x(\mu^{k_i}, b^{k_i}, c(\tau^{k_i})) \ne \lim_{j \to \infty} x(\mu^{k_j}, b^{k_j}, c(\tau^{k_j})).$$

The only situation left is when $\mu^{k_i} \downarrow 0$. However, if $\mu^{k_i} \downarrow 0$, we have the contradiction from Lemma 2.3 that

$$0 = \lim_{i \to \infty} x_N(\mu^{k_i}, b^{k_i}, c(k_i)) \ne \lim_{j \to \infty} x_N(\mu^{k_j}, b^{k_j}, c(\tau^{k_j})) > 0.$$

■

In this section, we have shown that while the limit of a central path is not continuous in b and c, the perturbed central paths are well behaved if viewed as a set. Moreover, from Corollary 4.1 we have that the central path is continuous over $\mathcal{H}^2(b, c)$.

5 Independent, Nonlinear Perturbations

In this section we remove the restriction that the perturbation in c be linear. The analysis increases in difficulty, and characterizing the convergence of the perturbed central path under arbitrary, simultaneous, and independent perturbations in b and c remains an open question. We provide sufficient conditions to guarantee the convergence of $x(\mu^k, b^k, c^k)$ and develop a process to find the limit. An example illustrates the difficulties of establishing exactly when $x(\mu^k, b^k, c^k)$ converges.

The sufficient conditions require that $\mathcal{G}^2$ be partitioned into equivalence classes. For any $b \in \mathcal{G}_b$, we say that c^1 and c^2 in $\mathcal{G}^2$ are "A-similar", denoted by $c^1 \overset{A}{\sim} c^2$, if $PCP_{(b,c^1)} \cap PCP_{(b,c^2)} \neq \emptyset$. The first goal of this section is to show that $\overset{A}{\sim}$ is an equivalence relation on $\mathcal{G}^2$. We begin by showing that central paths may not intersect unless they are equal. The first lemma provides sufficient conditions for two primal central paths to be equivalent.

Lemma 5.1 *Let $c_0^1 = proj_{\text{null}(A)} c^1$ and $c_0^2 = proj_{\text{null}(A)} c^2$. Then, $PCP_{(b,c_0^1)} = PCP_{(b,c^1)}$ and $PCP_{(b,c_0^2)} = PCP_{(b,c^2)}$. Moreover, if $c_0^1 = \alpha c_0^2$ for some $\alpha > 0$, $PCP_{(b,c^1)} = PCP_{(b,c^2)}$.*

Proof: Let $c_R^1 = \text{proj}_{\text{row}(A)} c^1$ and $c_R^2 = \text{proj}_{\text{row}(A)} c^2$, so that $c^1 = c_0^1 + c_R^1$ and $c^2 = c_0^2 + c_R^2$. Let $\alpha > 0$ be such that $c_0^1 = \alpha c_0^2$. Since c_R^1 and c_R^2 are in row(A), we have from Theorem 2.3 that $c_R^1 x$ and $c_R^2 x$ are constant on $\mathcal{P}_b$. This means that $x(\mu, b, c^1)$ and $x(\mu, b, c^2)$ are respectively the unique solutions to

$$\min \left\{ c_0^1 x - \mu \sum_{i=1}^{n} \ln(x_i) : x \in \mathcal{P}_b^o \right\} \quad \text{and} \quad \min \left\{ c_0^2 x - \mu \sum_{i=1}^{n} \ln(x_i) : x \in \mathcal{P}_b^o \right\}.$$

Hence, $PCP_{(b,c_0^1)} = PCP_{(b,c^1)}$ and $PCP_{(b,c_0^2)} = PCP_{(b,c^2)}$. Multiplying the objective function of the first math program by α shows that $x(\mu, b, c^1) = x(\alpha\mu, b, c^2)$, which implies that $PCP_{(b,c^1)} = PCP_{(b,c^2)}$. ∎

The following corollary is stated for future reference.

Corollary 5.1 *If $proj_{\text{null}(A)} c^1 = \alpha proj_{\text{null}(A)} c^2$, for some $\alpha > 0$, then*

$$x(\mu, b, c^1) = x(\mu, b, proj_{\text{null}(A)} c^1) = x(\alpha\mu, b, proj_{\text{null}(A)} c^2) = x(\alpha\mu, b, c^2).$$

Proof: The result is immediate from the proof of Lemma 5.1. ∎

The next theorem establishes that the central paths within a polyhedron are either the same or disjoint. Since $PCP_{(b,c)}$ contains only those elements that correspond to a positive μ, this does not say that two different central paths may not terminate at the same point. However, it does say that two different central paths may not cross en-route to either $x^*(b, c)$ or $\bar{x}(b)$.

Theorem 5.1 *If $PCP_{(b,c^1)} \cap PCP_{(b,c^2)} \neq \emptyset$, $PCP_{(b,c^1)} = PCP_{(b,c^2)}$.*

Proof: From Corollary 5.1 we know that there is no loss of generality by assuming that c^1 and c^2 are in null(A). Let μ^1 and μ^2 be positive such that $x(\mu^1, b, c^1) = x(\mu^2, b, c^2)$. Since $s(\mu^1, b, c^1)X(\mu^1, b, c^1) = \mu^1 e^T$ and $s(\mu^2, b, c^1)X(\mu^2, b, c^1) = \mu^2 e^T$, we have that $s(\mu^1, b, c^1) = \mu^1 e^T X^{-1}(\mu^1, b, c^1)$ and $s(\mu^2, b, c^1) = \mu^2 e^T X^{-1}(\mu^2, b, c^1)$. From the dual feasibility constraints we have that

$$c^1 - \mu^1 e^T X^{-1}(\mu^1, b, c^1) - y(\mu^1, b, c^1)A = 0 \quad \text{and} \quad c^2 - \mu^2 e^T X^{-1}(\mu^2, b, c^2) - y(\mu^2, b, c^2)A = 0.$$

Multiplying the first equation by $1/\mu^1$, the second equation by $1/\mu^2$, and subtracting yields

$$(1/\mu^1)c^1 - (1/\mu^2)c^2 = ((1/\mu^1)y(\mu^1, b, c^1) - (1/\mu^2)y(\mu^2, b, c^2))A.$$

Since the left-hand side is in the null(A) and the right-hand side is in the row(A), both must be zero. Hence, $c^1 = (\mu^1/\mu^2)c^2$, and from Lemma 5.1 we have that $PCP_{(b,c^1)} = PCP_{(b,c^2)}$. ∎

Two important corollaries follow.

Corollary 5.2 *If $c^1 \overset{A}{\sim} c^2$, $PCP_{(b,c^1)} = PCP_{(b,c^2)}$.*

Corollary 5.3 *We have that $proj_{\text{null}(A)} c^1 = \alpha proj_{\text{null}(A)} c^2$, for some positive α if, and only if, $PCP_{(b,c^1)} = PCP_{(b,c^2)}$*

Proof: The sufficiency is established by Lemma 5.1. The necessity follows because if $PCP_{(b,c^1)} = PCP_{(b,c^2)}$, then there is a positive μ^1 and μ^2 such that $x(\mu^1, b, c^1) = x(\mu^2, b, c^2)$, and from the proof of Theorem 5.1 we have that $\text{proj}_{\text{null}(A)} c^1 = \alpha \text{proj}_{\text{null}(A)} c^2$ for some positive α. ■

Theorem 5.2 states that $\overset{A}{\sim}$ is indeed an equivalence relation.

Theorem 5.2 *$\overset{A}{\sim}$ is an equivalence relation on $\mathcal{G}^2$. Furthermore, the equivalence class of c^1 is,*

$$[c^1]_A = \{c : \textit{proj}_{\text{null}(A)} c^1 = \alpha \textit{proj}_{\text{null}(A)} c, \textit{ for some positive } \alpha\}.$$

Proof: Clearly $c^1 \overset{A}{\sim} c^1$, and if $c^1 \overset{A}{\sim} c^2$, then $c^2 \overset{A}{\sim} c^1$. So $\overset{A}{\sim}$ is reflexive and symmetric. From Corollary 5.2 we have that if $c^1 \overset{A}{\sim} c^2$ and $c^2 \overset{A}{\sim} c^3$, then $PCP_{(b,c^1)} = PCP_{(b,c^2)} = PCP_{(b,c^3)}$, which implies that $c^1 \overset{A}{\sim} c^3$. Hence, $\overset{A}{\sim}$ is transitive and an equivalence relation. From Theorem 5.1 and Corollary 5.3 we have that the equivalence classes are as stated. ■

Our conditions that guarantee the convergence of $x(\mu^k, b^k, c^k)$ rely on two new types of convergence. For a sequence x^k, we let $\mathbf{C}(x^k)$ be the set of cluster points of x^k. Furthermore, for any sequence c^k, we set $d^k = c^k/\|c^k\|$ so long as $c^k \neq 0$, and we define $\mathcal{F}(c^k)$ to be

$$\mathcal{F}(c^k) = \mathbf{C}(c^k) \cup \mathbf{C}(d^k).$$

In addition to the cluster points of c^k, the set $\mathcal{F}$ contains the "limiting directions" of the cost vectors. For example, if c^k is $(1/k, 1/k)$ for k even and (k, k^2) for k odd, $\mathbf{C}(c^k) = \{(0,0)\}$ and $\mathbf{C}(d^k) = \{(1/\sqrt{2}, 1/\sqrt{2}), (0,1)\}$. We say that c^k is *class convergent* if the cluster points of c^k and the limiting directions of c^k are contained in the same equivalence class.

Definition 5.1 *The sequence c^k is class convergent to $[c]_A$ if $\mathcal{F}(c^k) \subseteq [c]_A$.*

Definition 5.2 *The sequence (μ^k, c^k) is proportionately convergent if for any two subsequences, say c^{k_i} and c^{k_j}, having the property that*

$$\lim_{i\to\infty} \textit{proj}_{\text{null}(A)} c^{k_i}/\|c^{k_i}\| = \alpha \lim_{j\to\infty} \textit{proj}_{\text{null}(A)} c^{k_j}/\|c^{k_j}\|$$

we subsequently have that

$$\lim_{i\to\infty} \mu^{k_i}/\|c^{k_i}\| = \alpha \lim_{j\to\infty} \mu^{k_j}/\|c^{k_j}\|.$$

We point out that a proportionately convergent sequence may have the property that c^k contains a subsequence of zeros, it is just that this subsequence is not a candidate for either c^{k_i} or c^{k_j}. Proportional convergence imposes an interesting property on subsequences of c^k that converge to elements in $\mathbf{C}(c^k/\|c^k\|) \cap \text{row}(A)$. Suppose that (μ^k, c^k) is proportionately convergent and that c is in both $\mathbf{C}(c^k/\|c^k\|)$ and $\text{row}(A)$. Then $\text{proj}_{\text{null}(A)} c = 0$, and we have that $\text{proj}_{\text{null}(A)} c = \alpha \text{proj}_{\text{null}(A)} c$ for any positive α. Consequently, if $c^{k_i} \to c$, we have that $\mu^{k_i}/\|c^{k_i}\| \to 0$, which establishes the following result.

Theorem 5.3 *If (μ^k, c^k) is proportionately convergent and $c^{k_i} \to c \in \text{row}(A)$, then $\mu^{k_i} = o(\|c^{k_i}\|)$*

The next theorem provides sufficient conditions for $x(\mu^k, b^k, c^k)$ to converge to an element of a central path. The sequence c^k is not required to converge, but is instead required to be class convergent. As Example 5.1 demonstrates, this weaker condition on c^k is still too restrictive for necessity.

Theorem 5.4 *We have that $x(\mu^k, b^k, c^k)$ converges to an element of $PCP_{(b,c)}$ provided that*

1. *c^k is class convergent to $[c]_A$,*
2. *(μ^k, c^k) is proportionately convergent,*
3. *$c^k \neq 0$ for $k = 1, 2, 3, \ldots,$ and*
4. *$\mu^k/\|c^k\| = \Theta(1)$.*

Proof: Since $\mu^k/\|c^k\| = \Theta(1)$ and $x(\mu^k, b^k, c^k) = x(\mu^k/\|c^k\|, b^k, c^k/\|c^k\|)$, we have from Lemma 2.2 that $x(\mu^k, b^k, c^k)$ is bounded. The result is established by showing that all cluster points of $x(\mu^k, b^k, c^k)$ are equal. Consider the subsequences

$$x(\mu^{k_i}, b^{k_i}, c^{k_i}) \to \hat{x}^1, \quad x(\mu^{k_j}, b^{k_j}, c^{k_j}) \to \hat{x}^2, \quad c^{k_i}/\|c^{k_i}\| \to \hat{c}^1, \text{ and } \quad c^{k_j}/\|c^{k_j}\| \to \hat{c}^2.$$

From the class convergence we have that there is a positive α^1 and α^2 such that,

$$\begin{aligned} \lim_{i\to\infty} \alpha^1 \text{proj}_{\text{null}(A)} c^{k_i}/\|c^{k_i}\| &= \alpha^1 \text{proj}_{\text{null}(A)} \hat{c}^1 \\ &= \text{proj}_{\text{null}(A)} c \\ &= \alpha^2 \text{proj}_{\text{null}(A)} \hat{c}^2 \\ &= \lim_{j\to\infty} \alpha^2 \text{proj}_{\text{null}(A)} c^{k_j}/\|c^{k_j}\|. \end{aligned}$$

From the proportional convergence of (μ^k, c^k) and the assumption that $\mu^k/\|c^k\|$ is bounded away from zero, we have that

$$0 < \hat{\mu} = \lim_{i\to\infty} \alpha^1 \mu^{k_i}/\|c^{k_i}\| = \lim_{j\to\infty} \alpha^2 \mu^{k_j}/\|c^{k_j}\|.$$

From Corollary 5.1 we see that

$$\begin{aligned} & x(\mu^{k_i}, b^{k_i}, c^{k_i}) = x(\alpha^1 \mu^{k_i}/\|c^{k_i}\|, b^{k_i}, \alpha^1 \text{proj}_{\text{null}(A)} c^{k_i}/\|c^{k_i}\|) \\ \text{and} \quad & x(\mu^{k_j}, b^{k_j}, c^{k_j}) = x(\alpha^2 \mu^{k_j}/\|c^{k_j}\|, b^{k_j}, \alpha^2 \text{proj}_{\text{null}(A)} c^{k_j}/\|c^{k_j}\|). \end{aligned}$$

We now have from (6) that

$$\begin{aligned} \hat{x}^1 &= \lim_{i\to\infty} x(\mu^{k_i}, b^{k_i}, c^{k_i}) \\ &= \lim_{i\to\infty} x(\alpha^1 \mu^{k_i}/\|c^{k_i}\|, b^{k_i}, \alpha^1 \text{proj}_{\text{null}(A)} c^{k_i}/\|c^{k_i}\|) \\ &= x(\hat{\mu}, b, \text{proj}_{\text{null}(A)} c) \\ &= \lim_{j\to\infty} x(\alpha^2 \mu^{k_j}/\|c^{k_j}\|, b^{k_j}, \alpha^2 \text{proj}_{\text{null}(A)} c^{k_j}/\|c^{k_j}\|) \\ &= \lim_{j\to\infty} x(\mu^{k_j}, b^{k_j}, c^{k_j}) \\ &= \hat{x}^2. \end{aligned}$$

Hence, $x(\mu^k, b^k, c^k)$ converges to an element in $PCP_{(b,c)}$. ∎

We point out that Theorem 5.4 only guarantees convergence to an element of $PCP_{(b,c)}$, and hence every component of the limit is positive. This is guaranteed in the proof by the condition that $\mu^k/\|c^k\| = \Theta(1)$. The situation is more complicated if $\mu^k/\|c^k\| \downarrow 0$, and we illustrate the increased complication in the following example. This example has the desirable property that c^k converges, but even with this property the convergence of $x(\mu^k, b^k, c^k)$ requires the analysis of several nested linear programs. The example shows how we construct the *induced* sequences of (μ^k, b^k, c^k).

Example 5.1 *Consider the linear program*

$$\min\{(1/k)x_1 + (1/\sqrt{k})x_2 + (1/\sqrt{k})x_3 : 0 \le x_1 \le 1,\ 0 \le x_2 \le 1,\ 0 \le x_3 \le 1\}.$$

Let $\mu^k = 1/k$, and let x_4, x_5, and x_6 be the slack vectors. We point out that Theorem 5.4 does not apply because $\mu^k/\|c^k\| = 1/\sqrt{1+2k} \downarrow 0$. We consider a sequence of linear programs to analyze the convergence of $x(\mu^k, b^k, c^k)$. The idea is to iteratively "reduce" the original problem by linearizing the cost-coefficient perturbations and then using the results from Section 3 to identify a collection of variables that must be zero. The data related to the j^{th} step of the procedure is indicated by the first superscript. For example, the initial sequences μ^k and b^k are the same as $\mu^{(0,k)}$ and $b^{(0,k)}$. We set $(B^0|N^0)$ to be $(\{1, 2, \ldots, n\}|\emptyset)$. The notational convention is slightly different *for the cost coefficients, where we set* c^k *to be* $\delta c_{B^0}^{(0,k)}$.

Let $\hat{c}_{B^0}^0$ *be the limit of* $\delta c_{B^0}^{(0,k)}$. *The 'root' problem is defined by this limit and is*

$$(LP^0) \quad \min\{\hat{c}_{B^0}^0 x_{B^0}^0 : A_{B^0} x_{B^0}^0 \leq b, x_{B^0}^0 \geq 0\},$$

where the superscript 0 *on the* x *indicates that these are the decision variables for the root problem. The optimal partition for* (LP^0) *is* $(B^1|N^1) = (\{1,2,3,4,5,6\}|\ \emptyset)$, *and we point out that* $(B^1|N^1)$ *partitions* B^0, *and hence* $B^1 \subseteq B^0$. *To define the first subproblem, we set*

$$\begin{aligned}
\tau^{(1,k)} &= \|c_{B^0}^{(0,k)} - \hat{c}_{B^0}^0\| = \sqrt{2k+1}/k, \\
\mu^{(1,k)} &= \mu^{(0,k)}/\tau^{(1,k)} = 1/\sqrt{2k+1}, \\
b^{(1,k)} &= b^{(0,k)} - A_{N^1} x_{N^1}^0(\mu^{(0,k)}, b^{(0,k)}, c^{(0,k)}) = (1,1,1)^T, \text{ and} \\
\delta c_{B^0}^{(1,k)} &= (1/\tau^{(1,k)})(c_{B^0}^{(0,k)} - \hat{c}_{B^0}^0) = (1/\sqrt{2k+1})(1, \sqrt{k}, \sqrt{k}).
\end{aligned}$$

The sequence $(\mu^{(1,k)}, b^{(1,k)}, \delta c_{B^0}^{(1,k)})$ *is the first* **induced** *sequence of* (μ^k, b^k, c^k), *and it is this sequence that defines the first subproblem. We "linearize"* $c_{B^0}^{(0,k)}$ *by rewriting it as* $\hat{c}_{B^0}^0 + \tau^{(1,k)} \delta c_{B^0}^{(1,k)}$, *from which we have that*

$$x_{B^0}(\mu^k, b^k, c^k) = x_{B^0}^0(\mu^{(0,k)}, b^{(0,k)}, \delta c_{B^0}^{(0,k)}) = x_{B^0}^0(\mu^{(0,k)}, b^{(0,k)}, \hat{c}_{B^0}^0 + \tau^{(1,k)} \delta c_{B^0}^{(1,k)}). \tag{15}$$

The constant term $\hat{c}_{B^0}^0$ *is used in the root problem to identify the optimal partition* (B^1, N^1), *and from Lemma 2.3 we know that the variables indexed by* N^1 *are zero for every cluster point of* $x(\mu^k, b^k, c^k)$. *Unfortunately, these may, or may not, be the only variables that are zero (and in this example none of the zero variables are indexed by* N^0 *because it is empty). These variables are essentially removed from the problem because they are moved to the right-hand side, which reduces the dimensionality of the problem and places us on a shifted optimal face of the root problem. If* $\delta c_{B^0}^{(1,k)}$ *had been constant, we could have established the limit of* $x(\mu^k, b^k, c^k)$ *from Theorems 3.1 and 3.2, and this limit would have been on the central path defined by minimizing* $\delta c_{B^0}^{(1,k)} x_{B^0}^0$ *over the optimal face of the root problem. However,* $\delta c_{B^0}^{(1,k)}$ *is not constant, and we repeat the process by linearizing the linear programs defined over the optimal face of the root problem. Notice that the sequence* $\delta c_{B^0}^{(1,k)}$ *does not converge to zero, but rather* $\delta c_{B^0}^{(1,k)} \to (0, 1/\sqrt{2}, 1\sqrt{2})$. *Furthermore, from Lemma 2.3 we have that* $b^{(1,k)} \to b$ *(this also follows in this case because* $N^0 = \emptyset$*). Defining* $\hat{c}_{B^1}^1$ *to be the limit of* $\delta c_{B^1}^{(1,k)}$, *we have that the first subproblem is*

$$(LP^1) \quad \min\{\hat{c}_{B^1}^1 x_{B^1}^1 : A_{B^1} x_{B^1}^1 \leq b, x_{B^1}^1 \geq 0\},$$

or equivalently

$$\min\{(1/\sqrt{2})x_2^1 + (1/\sqrt{2})x_3^1 : 0 \leq x_i^1 \leq 1, i = 1,2,3\}.$$

The relationship between LP^0 *and* LP^1 *is the same as the relationship between our standard linear program in (1) and the linear program in (12). So, from (14), (15), and the definition of* $\mu^{(1,k)}$, *we have that*

$$\left.\begin{aligned}
x_{B^1}(\mu^k, b^k, c^k) &= x_{B^1}^0(\mu^{(0,k)}, b^{(0,k)}, \delta c_{B^0}^{(0,k)}) \\
&= x_{B^1}^0(\mu^{(0,k)}, b^{(0,k)}, \hat{c}_{B^0}^0 + \tau^{(1,k)} \delta c_{B^0}^{(1,k)}) \\
&= x_{B^1}^1(\mu^{(0,k)}/\tau^{(1,k)}, b^{(1,k)}, \delta c_{B^1}^{(1,k)}) \\
&= x_{B^1}^1(\mu^{(1,k)}, b^{(1,k)}, \delta c_{B^1}^{(1,k)}).
\end{aligned}\right\} \tag{16}$$

The optimal partition for LP^1 *is* $(B^2|N^2) = (\{1,4,5,6\}|\{2,3\})$, *which partitions* B^1. *As before, we have from Lemma 2.3 that the components indexed by* N^2 *are zero in every cluster point of* $x_{B^1}^1(\mu^{(1,k)}, b^{(1,k)}, \delta c_{B^1}^{(1,k)})$, *and we move these variables to the right-hand side (this is the first 'reduction' for this example). The remaining components are indexed by* $B^2 \subseteq B^1$, *and we have from (16) that* $x_{B^2}(\mu^k, b^k, c^k) = x_{B^2}^1(\mu^{(1,k)}, b^{(1,k)}, \delta c_{B^1}^{(1,k)})$.

Similar to the first subproblem, the second subproblem relies on

$$\begin{aligned} \tau^{(2,k)} &= \|\delta c_{B^1}^{(1,k)} - \hat{c}_{B^1}^1\| \\ \mu^{(2,k)} &= \mu^{(1,k)}/\tau^{(2,k)} \\ b^{(2,k)} &= b^{(1,k)} - A_{N^2} x_{N^2}^1(\mu^{(1,k)}, b^{(1,k)}, \delta c_{B^1}^{(1,k)}), \text{ and} \\ \delta c_{B^1}^{(2,k)} &= (1/\tau^{(2,k)})(\delta c_{B^1}^{(1,k)} - \hat{c}_{B^1}^1). \end{aligned}$$

The second induced sequence of (μ^k, b^k, c^k) *is* $(\mu^{(2,k)}, b^{(2,k)}, \delta c_{B^1}^{(2,k)})$, *and as before, we have that* $\delta c_{B^1}^k = \hat{c}_{B^1}^1 + \tau^{(2,k)} \delta c_{B^1}^{(2,k)}$. *Hence,*

$$x_{B^1}^1(\mu^{(1,k)}, b^{(1,k)}, \delta c_{B^1}^{(1,k)}) = x_{B^1}^1(\mu^{(1,k)}, b^{(1,k)}, \hat{c}_{B^1}^1 + \tau^{(2,k)} \delta c_{B^1}^{(2,k)}).$$

It is easily checked that $\mu^{(2,k)} = \sqrt{2/(2 + 2/(\sqrt{2k} + \sqrt{2k+1})^2)} \to 1$, *which is important because* $\mu^{(2,k)}$ *does* **not** *converge to zero. The second subproblem requires only the* B^2 *components of the limit of* $\delta c_{B^1}^{(2,k)}$, *and it is easy to check that* $\delta c_{B^2}^{(2,k)} \to (1,0,0,0)$ *(the zero components follow because* x_4, x_5, *and* x_6 *are slack variables). Setting* $\hat{c}_{B^2}^2$ *to be the limit of* $\delta c_{B^2}^{(2,k)}$, *we have that the second subproblem is*

$$(LP^2) \quad \min\{\hat{c}_{B^2}^2 x_{B^2}^2 : A_{B^2} x_{B^2}^2 \le b, x_{B^2}^2 \ge 0\} = \min\{x_1^2 : 0 \le x_1^2 \le 1\}.$$

Similar to (16), we have that

$$\begin{aligned} x_{B^2}(\mu^k, b^k, c^k) &= x_{B^2}^0(\mu^{(0,k)}, b^{(0,k)}, \delta c_{B^0}^{(0,k)}) \\ &= x_{B^2}^0(\mu^{(0,k)}, b^{(0,k)}, \hat{c}_{B^0}^0 + \tau^{(1,k)} \delta c_{B^0}^{(1,k)}) \\ &= x_{B^2}^1(\mu^{(0,k)}/\tau^{(1,k)}, b^{(1,k)}, \delta c_{B^1}^{(1,k)}) \\ &= x_{B^2}^1(\mu^{(1,k)}, b^{(1,k)}, \delta c_{B^1}^{(1,k)}) \\ &= x_{B^2}^1(\mu^{(1,k)}, b^{(1,k)}, \hat{c}_{B^1}^1 + \tau^{(2,k)} \delta c_{B^1}^{(2,k)}) \\ &= x_{B^2}^2(\mu^{(1,k)}/\tau^{(2,k)}, b^{(2,k)}, \delta c_{B^2}^{(2,k)}) \\ &= x_{B^2}^2(\mu^{(2,k)}, b^{(2,k)}, \delta c_{B^2}^{(2,k)}). \end{aligned}$$

Since $\mu^{(2,k)} \to 1$, *we have from (6) that* $x_{B^1}^2(\mu^{(2,k)}, b^{(2,k)}, \delta c_{B^1}^{(2,k)}) \to x_{B^1}^2(1, b, 1)$, *and a straight forward calculation shows that* $x_1^2(1, b, 1) = (3 - \sqrt{5})/2$. *We conclude that*

$$x(\mu^k, b^k, c^k) \to ((3 - \sqrt{5})/2, 0, 0, (\sqrt{5} - 1)/2, 1, 1)^T.$$

The technique used in Example 5.1 suggests an algorithmic manner to calculate the limit of $x(\mu^k, b^k, c^k)$. Instead of trying to calculate this limit directly, we instead calculated the limit of c^k and use this limit to form a root problem. The N set of the corresponding optimal partition indexes a collection of variables that must go to zero, and in fact, this is the entire collection of zero variables if $\mu^{(1,k)}$ has a positive limit. However, if $\mu^{(1,k)}$ decreases to zero, the variables whose value must be zero are moved to the right-hand side, and the limit of $\delta c_{B^1}^{(1,k)}$ is calculated to form the first subproblem. Again, we know that any variables listed in the corresponding N set of the optimal partition are zero in the limit. The process repeats until either all variables are found to be zero, or until $\mu^{(j,k)}$ does not converge to zero for some j.

Example 5.1 has the property that $\delta c_{B^j}^{(j,k)}$ converges for $j = 0, 1, 2$, but the proof of Theorem 5.4 shows that this need not be the case. Instead, at each step of the procedure we need for $\delta c_{B^j}^{(j,k)}$ to be class convergent. As long as this is true, we continue to form the induced sequences until we have a criteria that guarantees convergence. The process in Table 3 describes how to construct the induced sequences, and Theorem 5.5 shows that $x(\mu^k, b^k, c^k)$ converges if this process terminates with an exit code of 0.

Lemma 5.2 *Let* $(B|N)$ *be the optimal partition for* $\min\{\hat{c}x : Ax = b, x \ge 0\}$. *If* c^k *is a non-zero sequence that is class convergent to* $[c]_A$ *and* $\mu^k/\|c^k\| \downarrow 0$, *then* $x_N(\mu^k, b^k, c^k) \downarrow 0$.

Step 1 Set $j=0$, $(B^0|N^0) = (\{1,2,\ldots,n\}|\emptyset)$ and $(\mu^k, b^k, c^k) = (\mu^{(0,k)}, b^{(0,k)}, \delta c_{B^0}^{(0,k)})$.

Step 2 Stop with exit code 0 if any of the following are true,

- $B^j = \emptyset$,
- $(\mu^{(j,k)}, b^{(j,k)}, \delta c_{B^j}^{(j,k)})$ satisfies conditions $(1)-(4)$ of Theorem 5.4, or
- $j \geq 1$ and $\mu^{(j,k)}/\|\delta c_{B^j}^{(j,k)}\| \to \infty$,

Step 3 If we have that $\|\delta c_{B^j}^{(j,k)}\| \neq 0$, $\mu^{(j,k)}/\|\delta c_{B^j}^{(j,k)}\| \downarrow 0$, and that there exists a $\hat{c}_{B^j}^j$ such that $\delta c_{B^j}^{(j,k)}$ is class convergent to $[\hat{c}^j]_{A_{B^j}}$, then continue with Step 4. Otherwise, stop with exit code 1.

Step 4 Solve the linear program

$$(LP^j) \quad \min\{\hat{c}_{B^j}^j x_{B^j}^j : A_{B^j} x_{B^j}^j = b, x_{B^j}^j \geq 0\}$$

and let $(B^{j+1}|N^{j+1})$ be the optimal partition.

Step 5 Set

$$\begin{aligned}
\tau^{(j+1,k)} &= \|\delta c^{(j,k)} - \hat{c}^j\| \\
\mu^{(j+1,k)} &= \mu^{(j,k)}/\tau^{(j+1,k)} \\
b^{(j+1,k)} &= b^{(j,k)} - A_{N^{j+1}} x_{N^{j+1}}^j(\mu^{(j,k)}, b^{(j,k)}, \delta c_{B^j}^{(j,k)}), \text{ and} \\
\delta c_{B^j}^{(j+1,k)} &= (1/\tau^{(j+1,k)})(\delta c_{B^j}^{(j,k)} - \hat{c}_{B^j}^j).
\end{aligned}$$

Step 7 Let $j = j+1$ and go to step 2.

Table 3: The process to construct the induced sequences of (μ^k, b^k, c^k).

Proof: We have from Theorem 5.2 that there is no loss in generality by assuming that c is in null(A). Since $x(\mu^k, b^k, c^k) = x(\mu^k/\|c^k\|, b^k, c^k/\|c^k\|)$ and $c^k/\|c^k\|$ is bounded, we have from Lemma 2.2 that $x(\mu^k, b^k, c^k)$ is bounded. So, there is a subsequence $(\mu^{k_i}, b^{k_i}, c^{k_i})$ such that

$$\begin{aligned}
x(\mu^{k_i}, b^{k_i}, c^{k_i}) = x(\mu^{k_i}/\|c^{k_i}\|, b^{k_i}, c^{k_i}/\|c^{k_i}\|) &\to \hat{x}, \\
y(\mu^{k_i}/\|c^{k_i}\|, b^{k_i}, c^{k_i}/\|c^{k_i}\|) &\to \hat{y}, \\
s(\mu^{k_i}/\|c^{k_i}\|, b^{k_i}, c^{k_i}/\|c^{k_i}\|) &\to \hat{s}, \text{ and} \\
c^{k_i}/\|c^{k_i}\| &\to \hat{c}.
\end{aligned}$$

For notational ease, we let

$$x^i = x(\mu^{k_i}/\|c^{k_i}\|, b^{k_i}, c^{k_i}/\|c^{k_i}\|), \quad y^i = y(\mu^{k_i}/\|c^{k_i}\|, b^{k_i}, c^{k_i}/\|c^{k_i}\|), \text{ and } s^i = s(\mu^{k_i}/\|c^{k_i}\|, b^{k_i}, c^{k_i}/\|c^{k_i}\|).$$

From the assumption that c^k is class convergent to $[c]_A$ we have that there is a positive α such that $\alpha \text{proj}_{\text{null}(A)}\hat{c} = c$. Since

$$\left.\begin{aligned}
Ax^i &= b^{k_i} \\
y^i A + s^i &= c^{k_i}/\|c^{k_i}\| \\
s^i x^i &= n\mu^{k_i}/\|c^{k_i}\|
\end{aligned}\right\} \Rightarrow \left\{\begin{aligned}
A\hat{x} &= b \\
\hat{y}A + \hat{s} &= \hat{c} \\
\hat{s}\hat{x} &= 0,
\end{aligned}\right.$$

we have that $\hat{x}$ is an optimal solution to $\min\{\hat{c}x : Ax = b, x \geq 0\}$. Let $\tilde{y}$ be such that $\tilde{y}A = \text{proj}_{\text{row}(A)}\hat{c}$, from which we have that $\hat{c} = \text{proj}_{\text{null}(A)}\hat{c} + \text{proj}_{\text{row}(A)}\hat{c} = \text{proj}_{\text{null}(A)}\hat{c} + \tilde{y}A$. Substituting this into $\hat{y}A + \hat{s} = \hat{c}$, we have that

$$A\hat{x} = b, \quad \hat{x} \geq 0, \quad \alpha(\hat{y} - \tilde{y})A + \alpha\hat{s} = \alpha\text{proj}_{\text{null}(A)}\hat{c} = c, \quad \alpha\hat{s} \geq 0, \text{ and } \quad \hat{s}\hat{x} = 0.$$

Hence, $\hat{x}$ is also an optimal solution to $\min\{cx : Ax = b, x \geq 0\}$, which implies that $x(\mu^k, b^k, c^k) \to \hat{x}_N = 0$. ∎

Theorem 5.5 *If the process in Table 3 stops with an exit code of* 0*, then* $x(\mu^k, b^k, c^k)$ *converges.*

Proof: If the process terminates with $j = 0$ and an exit code of 0, then we have that $(\mu^k, b^k, c^k) = (\mu^{(0,k)}, b^{(0,k)}, \delta c_{B^0}^{(0,k)})$ satisfies conditions $(1) - (4)$ of Theorem 5.4, which implies that $x(\mu^k, b^k, c^k)$ converges. Suppose that the process in Table 3 terminates with an exit code of 0 and that the induced sequences are $(\mu^{(j,k)}, b^{(j,k)}, \delta c_{B^{j-1}}^{(j,k)})$, for $j = 1, 2, \ldots, J$. The proof follows with a careful inspection of how the sequence $x(\mu^k, b^k, c^k)$ partitions itself as the process continues. From the definition of the first induced sequence we have that

$$x(\mu^k, b^k, c^k) = x_{B^0}^0(\mu^{(0,k)}, b^{(0,k)}, \delta c_{B^0}^{(0,k)}) = \left(\begin{array}{c} x_{B^1}^0(\mu^{(0,k)}, b^{(0,k)}, \hat{c}^0 + \tau^{(1,k)} \delta c_{B^0}^{(1,k)}) \\ \hline x_{N^1}^0(\mu^{(0,k)}, b^{(0,k)}, \delta c_{B^0}^{(0,k)}) \end{array} \right).$$

From (14) we have that

$$x_{B^1}^0(\mu^{(0,k)}, b^{(0,k)}, \hat{c}^0 + \tau^{(1,k)} \delta c_{B^0}^{(1,k)}) = x_{B^1}^1(\mu^{(1,k)}, b^{(1,k)}, \delta c_{B^1}^{(1,k)}).$$

Using the second induced sequence we have that

$$x_{B^1}^1(\mu^{(1,k)}, b^{(1,k)}, \delta c_{B^1}^{(1,k)}) = \left(\begin{array}{c} x_{B^2}^1(\mu^{(1,k)}, b^{(1,k)}, \hat{c}^1 + \tau^{(2,k)} \delta c_{B^1}^{(2,k)}) \\ \hline x_{N^2}^1(\mu^{(1,k)}, b^{(1,k)}, \delta c_{B^1}^{(1,k)}) \end{array} \right),$$

which implies that

$$x(\mu^k, b^k, c^k) = \left(\begin{array}{c} x_{B^2}^1(\mu^{(1,k)}, b^{(1,k)}, \hat{c}^1 + \tau^{(2,k)} \delta c_{B^1}^{(2,k)}) \\ \hline x_{N^2}^1(\mu^{(1,k)}, b^{(1,k)}, \delta c_{B^1}^{(1,k)}) \\ \hline x_{N^1}^0(\mu^{(0,k)}, b^{(0,k)}, \delta c_{B^0}^{(0,k)}) \end{array} \right).$$

Again, from (14) and the definition the third induced sequence we have that

$$x_{B^2}^1(\mu^{(1,k)}, b^{(1,k)}, \hat{c}^1 + \tau^{(2,k)} \delta c_{B^1}^{(2,k)}) = x_{B^2}^2(\mu^{(2,k)}, b^{(2,k)}, \delta c_{B^2}^{(2,k)}) = \left(\begin{array}{c} x_{B^3}^2(\mu^{(2,k)}, b^{(2,k)}, \hat{c}^2 + \tau^{(3,k)} \delta c_{B^2}^{(3,k)}) \\ \hline x_{N^3}^2(\mu^{(2,k)}, b^{(2,k)}, \delta c_{B^2}^{(2,k)}) \end{array} \right).$$

The process continues until

$$x(\mu^k, b^k, c^k) = \left(\begin{array}{c} x^J_{B^J}(\mu^{(J,k)}, b^{(J,k)}, \delta c^{(J,k)}_{B^J}) \\ \hline x^{J-1}_{N^J}(\mu^{(J-1,k)}, b^{(J-1,k)}, \delta c^{(J-1,k)}_{B^{J-1}}) \\ \hline \vdots \\ \hline x^1_{N^2}(\mu^{(1,k)}, b^{(1,k)}, \delta c^{(1,k)}_{B^1}) \\ \hline x^0_{N^1}(\mu^{(0,k)}, b^{(0,k)}, \delta c^{(0,k)}_{B^0}) \end{array} \right).$$

The fact that the first induced sequence was created implies that $\delta c^{(0,k)}_{B^0} \neq 0$, $\mu^{(0,k)}/\|\delta c_{B^0}\| \downarrow 0$, and $\delta c^{(0,k)}_{B^0}$ is class convergent to $[\hat{c}^0_{B^0}]_{A_{B^0}}$. By assumption we have that $b^{(0,k)} = b^k \to b$. So, from Lemma 5.2 we have that $x^0_{N^1}(\mu^{(0,k)}, b^{(0,k)}, \delta c^{(0,k)}_{B^0}) \downarrow 0$, which subsequently implies that $b^{(1,k)} = b^{(0,k)} - A_{N^1} x^0_{N^1}(\mu^{(0,k)}, b^{(0,k)}, \delta c^{(0,k)}_{B^0}) \to b$. By the same logic, and repeated applications of Lemma 5.2, we find that $b^{(j,k)} \to b$, for $j = 1, 2, \ldots, J$, because

$$\left(\begin{array}{c} x^{J-1}_{N^J}(\mu^{(J-1,k)}, b^{(J-1,k)}, \delta c^{(J-1,k)}_{B^{J-1}}) \\ \hline \vdots \\ \hline x^1_{N^2}(\mu^{(1,k)}, b^{(1,k)}, \delta c^{(1,k)}_{B^1}) \\ \hline x^0_{N^1}(\mu^{(0,k)}, b^{(0,k)}, \delta c^{(0,k)}_{B^0}) \end{array} \right) \downarrow 0.$$

At this point we have that if the process terminated because $B^J = \emptyset$, then $x(\mu^k, b^k, c^k) \downarrow 0$. Suppose that $(\mu^{(J,k)}, b^{(J,k)}, \delta c^{(J,k)}_{B^J})$ satisfies conditions $(1) - (4)$ of Theorem 5.4, then we have that $x^J_{B^J}(\mu^{(J,k)}, b^{(J,k)}, \delta c^{(J,k)}_{B^J})$ converges, and hence, so does $x(\mu^k, b^k, c^k)$.

Suppose that $\mu^{(J,k)}/\|\delta c^{(J,k)}_{B^J}\| \to \infty$, which subsequently implies that $\delta c^{(J,k)}_{B^J}/\mu^{(J,k)} \to 0 \in \text{row}(A)$. Since $\{x_{B^{J-1}} : A_{B^j} x_{B^J} = b, x_{B^J} \geq 0, x_{N^J} = 0\}$ is the optimal set of (LP^{J-1}), we have from Lemma 2.2 that $\{x_{B^J} : A_{B^j} x_{B^J} = b, x_{B^J} \geq 0\}$ is bounded. So, from Theorem 2.4 we have that $x_{B^J}(\mu^{(J,k)}, b^{(J,k)}, \delta c^{(J,k)}_{B^J})$ converges. ■

We conclude by pointing out that $x(\mu^k, b^k, c^k)$ can converge if the process in Table 3 terminates with an exit code of 1. As an example, let $b^k = 1$, $\mu^k = 1/k$, $A = [1, 1]$, and c^k be $(1, 1)$ if k is even and $(1/k, 1/k)$ if k is odd. Then, for all k we have that $c^k \in \text{row}(A)$, and from Theorem 2.3 we know that $x(\mu^k, b^k, c^k) = \bar{x}(b) = (1/2, 1/2)^T$. However, $\mu^k/\|c^k\|$ is $1/k\sqrt{2}$ if k is even and $1/\sqrt{2}$ is k is odd. Hence, the sequence $\mu^k/\|c^k\|$ does not decrease to zero and is not $\Theta(1)$, and the process terminates with an exit code of 1.

6 Conclusions and Future Research

We have accomplished three main goals with the analysis developed in this paper. First, we have completely characterized the convergence of the perturbed central path followed by many infeasible-interior-point methods. This result is succinctly depicted in Table 2. Second, we have shown that the perturbed central path converges as set, so long as the cost vector perturbation is linear. In fact, the central path is continues over the set of cost directions for which the optimal partition is invariant. Third, we provided sufficient conditions for the perturbed central path to converge under arbitrary, simultaneous changes in b and c. These are the first results in the literature that deal with this complicated situation, and characterizing the convergence under such data perturbations remains an open question.

References

[1] J. Bonnans and F. Porta. Infeasible path following algorithms for linear complementarity problems. Raport de recherche 2445, Institur National De Recherche en Informatique Et En Automatique, Rocquencourt, France, 1994.

[2] R. Caron, H. Greenberg, and A. Holder. Analytic centers and repelling inequalities. Technical Report CCM 142, Center for Computational Mathematics, University of Colorado at Denver, 1999. To appear in European Journal of Operations Research.

[3] A. Fiacco and G. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques.* John Wiley & Sons, New York, 1968.

[4] K. Frisch. The logarithmic potential method of convex programming. Technical report, Univeristy Institute of Economics, Oslo, Norway, 1955. memorandum.

[5] H. Greenberg. *Mathematical Programming Glossary.* World Wide Web, `http://www-math.cudenver.edu/~hgreenbe/glossary/glossary.html`, 1996-2001.

[6] H. Greenberg. Simultaneous primal-dual right-hand-side sensitivity analysis from a strictly complementary solution of a linear program. *SIAM Journal on Optimization*, 10(2):427–442, 2000.

[7] H. Greenberg, A. Holder, C. Roos, and T. Terlaky. On the dimension of the set of rim perturbations for optimal partition invariance. Technical Report CCM No. 94, University of Colorado, at Denver, 1996. To appear in *SIAM Journal of Optimization.*

[8] B. Grunbaum. *Convex Polytopes.* Wiley-Interscience, New York, NY, 1967.

[9] M. Halická. Two simple proofs for analyticity of the central path in linear programming. *Operations Research Letters*, 28(1):9–19, 2001.

[10] A. Holder, J. Sturm, and S. Zhang. Marginal and parametric analysis of the central optimal solution. *Information Systems and Operational Research*, 39(4):394–415, 2001.

[11] P. Huard. Resolution of mathematical programming with nonlinear constraints by the method of centres. In J. Abadie, editor, *Nonlinear Programming*, chapter 8, pages 209–219. John Wiley & Sons, Inc., New York, NY, 1967.

[12] L. McLinden. An analogue of Moreau's approximation theorem, with applications to the nonlinear complementarity problem. *Pacific Journal of Mathematics*, 88(1):101–161, 1980.

[13] N. Megiddo. Pathways to the optimal set in linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming: Interior-Point Algorithms and Related Methods*, pages 131–158. Springer-Verlag, New York, 1989.

[14] S. Mizuno, M. Todd, and Y. Ye. A surface of analytic centers and primal-dual infeasible-interior point algorithms for linear programming. *Mathematics of Operations Research*, 20(1):135–162, 1995.

[15] R. Monteiro and T. Tsuchiya. Limiting behavior of the derivatives of certain trajectories associated with a monotone horizontal linear complementarity problem. *Mathemaitcs of Operations Research*, 21:793–814, 1996.

[16] C. Papadimitriou and K. Steiglitz. *Combinatorial optimization : algorithms and complexity.* Prentice Hall, 1982.

[17] C. Roos, T. Terlaky, and J.-Ph. Vial. *Theory and Algorithms for Linear Optimization: An Interior Point Approach.* John Wiley & Sons, New York, NY, 1997.

[18] G. Sonnevend. An "analytic centre" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In A. Prekopa, J. Szelezsan, and B. Strazicky, editors, *Lecture Notes in Control and Information Sciences*, volume 84, pages 866–875. Springer-Verlag, Heidelberg, Germany, 1986.

[19] G. Sonnevend. An implementation of the method of analytic centers. In A. Bensoussan and J. Lions, editors, *Lecture Notes in Control and Information Sciences*, volume 111, pages 297–308. Springer-Verlag, Heidelberg, Germany, 1988.

[20] G. Sonnevend. New algorithms in convex programming based on a notion of "Centre" (for systems of analytic inequalities) and on rational extrapolation. In K. Hoffman and et al, editors, *International Series of Numerical Mathematics*, volume 84, pages 311–326. Birkhauser Verlag Basel, Germany, 1988.

[21] G. Sonnevend. Applications of the notion of analytic center in approximation (estimation) problems. *Journal of Computational and Applied Mathematics*, 28:349–358, 1989.

[22] G. Sonnevend and J. Stoer. Global ellipsoidal approximations and homotopy methods for solving convex analytic programs. *Applied Mathematics and Optimization*, 21:139–165, 1990.

[23] G. Sonnevend, J. Stoer, and G. Zhao. On the complexity of following the central path of linear programs by linear extrapolation II. *Mathematical Programming*, 52:527–553, 1991.

[24] M. Wright. The interior-point revolution in constrained optimization. Technical Report 98-4-09, Bell Laboratories, Murray Hill, New Jersey, 1998.

[25] S. Wright. *Primal-Dual Interior-Point Methods.* SIAM, Philadelphia, PA, 1997.

[26] Y. Ye. *Interior Point Algorithms Theory and Analysis.* John Wiley & Sons, Inc., New York, NY, 1997.

[27] A. Yildrim and M. Todd. Sensitivity analysis in linear programming and semidefinate programming using interior-point methods. *Mathematical Programming*, 90(2):229–261, 2001.

[28] A. Yildrim and M. Todd. An interior-point approach to sensitivity analysis in degenerate linear programs. *SIAM Journal of Optimization*, 12(3):692–714, 2002.