Continuous Trajectories for Primal-Dual Potential-Reduction Methods

Reha H. Tütüncü *

September 7, 2001

Abstract

This article considers continuous trajectories of the vector fields induced by two different primal-dual potential-reduction algorithms for solving linear programming problems. For both algorithms, it is shown that the associated continuous trajectories include the central path and the duality gap converges to zero along all these trajectories. For the algorithm of Kojima, Mizuno, and Yoshise, there is a a surprisingly simple characterization of the associated trajectories. Using this characterization, it is shown that all associated trajectories converge to the analytic center of the primal-dual optimal face. Depending on the value of the potential function parameter, this convergence may be tangential to the central path, tangential to the optimal face, or in between.

Key words: linear programming, potential functions, potential-reduction methods, central path, continuous trajectories for linear programming.

AMS subject classification: 90C05

1 Introduction

During the past two decades, interior-point methods (IPMs) emerged as very efficient and reliable techniques for the solution of linear programming problems. The development of IPMs and their theoretical convergence analyses often rely on certain continuous trajectories associated with the given linear program. The best known example of such trajectories is the *central path*–the set of minimizers of the parametrized logarithmic barrier function in the interior of the feasible region.

Primal-dual variants of IPMs not only solve the given linear program but also its dual. These variants have been very successful in practical implementations and form the focus of this article. If both the given LP and its dual have strictly feasible solutions the primal-dual central path starts from the analytic center of the primal-dual feasible set and converges to the analytic center of the optimal solution set. This property of the central path led to the development of *path following* IPMs: algorithms that try to reach an optimal solution by

*Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA (`reha+@andrew.cmu.edu`). Research supported in part by NSF through grant CCR-9875559.

generating a sequence of points that are "close" to a corresponding sequence of points on the central path that converge to its limit point.

Weighted centers are generalizations of the central path that are obtained as minimizers of weighted logarithmic barrier functions. They share many properties of the central path, including convergence to an optimal point; see, e.g., [11]. Weighted centers and the central path can be characterized in alternative ways. For example, these trajectories are obtained as unique solutions of certain differential equations. Using this perspective, Adler and Monteiro analyzed the limiting behavior of continuous trajectories associated with primal-only affine-scaling algorithm [1]. Later, Monteiro analyzed continuous trajectories associated with projective-scaling and potential-reduction algorithms [8, 9], once again in the primal-only setting. In this article, we study trajectories of vector fields induced by primal-dual *potential-reduction* algorithms.

Potential-reduction algorithms use the following strategy: First, one defines a *potential function* that measures the quality (or potential) of any trial solution of the given problem combining measures of proximity to the set of optimal solutions, proximity to the feasible set in the case of infeasible-interior-points, and a measure of centrality within the feasible region. Potential functions are often chosen such that one approaches an optimal solution of the underlying problem by reducing the potential function. Then, the search for an optimal solution can be performed by progressive reduction of the potential function, leading to a potential-reduction algorithm. We refer the reader to two excellent surveys for further details on potential-reduction algorithms [2, 13].

Often, implementations of potential-reduction interior-point algorithms exhibit behavior that is similar to that of path-following algorithms. For example, they take about the same number of iterations as path-following algorithms and they tend to converge to the analytic center of the optimal face, just like most path-following variants. Since potential-reduction methods do not generally make an effort to follow the central path, this behavior is surprising. In an effort to better understand the limiting behavior of primal-dual potential-reduction algorithms for linear programs this paper studies continuous trajectories associated with two such methods. The first one is the algorithm proposed by Kojima, Mizuno, and Yoshise (KMY) [7], which uses scaled and projected steepest descent directions for the Tanabe-Todd-Ye (TTY) primal-dual potential function [12, 14]. The second algorithm we consider is a primal-dual variant of the Iri-Imai algorithm that uses Newton search directions for the multiplicative analogue of the TTY potential function [15, 16]. For both algorithms, we show that each associated trajectory has the property that the duality gap converges to zero as one traces the trajectory to its bounded limit. We also show that the central path is a special potential-reduction trajectory associated with both methods.

Our analysis is more extensive for the KMY algorithm, since we are able to express the points on the associated trajectories as unique optimal solutions of certain parametrized weighted logarithmic barrier problems. This characterization is similar to that obtained by Monteiro [9] for primal-only algorithms, but our formulas are explicit and much simpler. We show that all trajectories of the vector field induced by the KMY search directions converge to the analytic center of the primal-dual optimal face.

We also analyze the direction of convergence for these trajectories and demonstrate that their asymptotic behavior depends on the potential function parameter. There is a threshold value of this parameter–the value that makes the TTY potential function homogeneous. When the parameter is below this threshold, the centering is too strong and the trajectories converge tangentially to the central path. When the parameter is above the threshold, convergence happens tangentially to the optimal face. At the threshold value, the behavior of the trajectories is in between these two extremes and depends on the initial point.

Following this introduction, Section 2 discusses the two primal-dual potential-reduction methods mentioned above and search directions used by these two methods. Section 3 establishes the existence of continuous trajectories associated with these methods using standard results from the theory of differential equations. We prove the convergence of the trajectories associated with the Kojima-Mizuno-Yoshise algorithm in Section 4 and analyze the limiting behavior of these trajectories in Section 5. Our notation is fairly standard: For an n-dimensional vector x, the corresponding capital letter X denotes the $n \times n$ diagonal matrix with $X_{ii} \equiv x_i$. We will use the letter e to denote a column vector with all entries equal to 1 and its dimension will be apparent from the context. We also denote the base of the natural logarithm with e and sometimes the vector e and the scalar e appear in the same expression, but no confusion should arise. For a given matrix A, we use $\mathcal{R}(A)$ and $\mathcal{N}(A)$ to denote its range(column) and null space. For a vector-valued differentiable function $x(t)$ of a scalar variable t, we use the notation $\dot{x}$ or $\dot{x}(t)$ to denote the vector of the derivatives of its components with respect to t. For n dimensional vectors x and s, we write $x \circ s$ to denote their Hadamard (componentwise) product. Also, for an n dimensional vector x, we write x^p to denote the vector $X^p e$, where p can be fractional if $x > 0$.

2 Primal-Dual Potential-Reduction Directions

We consider linear programs in the following standard form:

$$(LP) \qquad \begin{array}{rrcl} \min_x & c^T x & & \\ & Ax & = & b \\ & x & \geq & 0, \end{array} \tag{1}$$

where $A \in \Re^{m \times n}$, $b \in \Re^m$, $c \in \Re^n$ are given, and $x \in \Re^n$. Without loss of generality we assume that the constraints are linearly independent. Then, the matrix A has full row rank. Further, we assume that $0 < m < n$; $m = 0$ and $m = n$ correspond to trivial problems.

The dual of this (primal) problem is:

$$(LD) \qquad \begin{array}{rrcccl} \max_{y,s} & b^T y & & & & \\ & A^T y & + & s & = & c \\ & & & s & \geq & 0, \end{array} \tag{2}$$

where $y \in \Re^m$ and $s \in \Re^n$. We can rewrite the dual problem by eliminating the y variables in (2). This is achieved by considering G^T, a null-space basis matrix for A, that is, G is an $(n-m) \times n$ matrix with rank $n-m$ and it satisfies $AG^T = 0$, $GA^T = 0$. Note also that, A^T is a null-space basis matrix for G. Further, let $d \in \Re^n$ be a vector satisfying $Ad = b$. Then, (2) is equivalent to the following problem which has a high degree of symmetry with (1):

$$(LD') \qquad \begin{array}{rrcl} \min_s & d^T s & & \\ & Gs & = & Gc \\ & s & \geq & 0. \end{array} \tag{3}$$

Let $\mathcal{F}$ and $\mathcal{F}^0$ denote the primal-dual feasible region and its relative interior:

$$\begin{aligned}\mathcal{F} &:= \{(x,s) : Ax = b, Gs = Gc, (x,s) \geq 0\},\\ \mathcal{F}^0 &:= \{(x,s) : Ax = b, Gs = Gc, (x,s) > 0\}.\end{aligned}$$

We assume that $\mathcal{F}^0$ is non-empty. This assumption has the important consequence that the primal-dual optimal solution set Ω defined below is nonempty and bounded:

$$\Omega := \{(x,s) \in \mathcal{F} : x^T s = 0\}. \tag{4}$$

We also define the optimal partition $\mathcal{B} \cup \mathcal{N} = \{1, \ldots, n\}$ for future reference:

$$\begin{aligned}\mathcal{B} &:= \{j : x_j > 0 \text{ for some } (x,s) \in \Omega\},\\ \mathcal{N} &:= \{j : s_j > 0 \text{ for some } (x,s) \in \Omega\}.\end{aligned}$$

The fact that $\mathcal{B}$ and $\mathcal{N}$ is a partition of $\{1, \ldots, n\}$ is a classical result of Goldman and Tucker.

Primal-dual feasible interior-point algorithms start with a point $(x^0, s^0) \in \mathcal{F}^0$ and move through $\mathcal{F}^0$ by generating search directions in the null space of the constraint equations which can be represented implicitly or explicitly, using null-space basis matrices:

$$\begin{aligned}\mathcal{N}\left(\begin{bmatrix} A & 0 \\ 0 & G \end{bmatrix}\right) &:= \{(\Delta x, \Delta s) : A\Delta x = 0, G\Delta s = 0\}\\ &= \{(\Delta x, \Delta s) : \Delta x = G^T \underline{\Delta x}, \Delta s = A^T \underline{\Delta s}, (\underline{\Delta x}, \underline{\Delta s}) \in \Re^{n-m} \times \Re^m\}.\end{aligned}$$

Given a search direction $(\Delta x, \Delta s)$, we will call the corresponding $(\underline{\Delta x}, \underline{\Delta s})$ a *reduced* search direction. A feasible interior-point algorithm that starts at $(x^0, s^0) \in \mathcal{F}^0$ can also be expressed completely in terms of the reduced iterates $(\underline{x}^k, \underline{s}^k)$ rather than the usual iterates (x^k, s^k) it generates. These two iterate sequences correspond to each other through the equations $x^k = x^0 + G^T \underline{x}^k$ and $s^k = s^0 + A^T \underline{s}^k$. Since G^T and A^T are basis matrices for the null-spaces of A and G respectively, $(\underline{x}^k, \underline{s}^k)$ are uniquely determined given (x^k, s^k) and the initial feasible iterate $(x^0, s^0) \in \mathcal{F}^0$: $\underline{x}^k = (GG^T)^{-1} G(x^k - x^0)$ and $\underline{s}^k = (AA^T)^{-1} A(s^k - s^0)$. The *reduced* feasible set and its interior can be defined as follows:

$$\begin{aligned}\mathcal{F}_R(x^0, s^0) &:= \{(\underline{x}, \underline{s}) : (x^0 + G^T \underline{x}, s^0 + A^T \underline{s}) \geq 0\},\\ \mathcal{F}_R^0(x^0, s^0) &:= \{(\underline{x}, \underline{s}) : (x^0 + G^T \underline{x}, s^0 + A^T \underline{s}) > 0\}.\end{aligned}$$

We will suppress the dependence of these sets on (x^0, s^0) in our notation. Note that $\mathcal{F}_R^0$ is a full-dimensional open subset of $\Re^n$. This simple alternative view of feasible interior-point methods is useful for our purposes since it allows us use standard results from the theory of ordinary differential equations later in our analysis.

Primal-dual potential-reduction algorithms for linear programming are derived using *potential functions*, i.e., functions that measure the quality (or potential) of trial solutions for the primal-dual pair of problems. The most frequently used primal-dual potential function for linear programming problems is the Tanabe-Todd-Ye (TTY) potential function [12, 14]:

$$\Phi_\rho(x,s) := \rho \ln(x^T s) - \sum_{i=1}^n \ln(x_i s_i), \text{ for every } (x,s) > 0. \tag{5}$$

As most other potential functions used in optimization, the TTY potential function has the property that it diverges to $-\infty$ along a feasible sequence $\{(x^k, s^k)\}$ if and only if this sequence is converging to a primal-dual optimal pair of solutions, as long as $\rho > n$. Therefore, the primal-dual pair of LP problems can be solved by minimizing the TTY potential-function. We now evaluate the gradient of this function for future reference:

$$\nabla \Phi_\rho(x,s) \;=\; \begin{bmatrix} \frac{\rho}{x^Ts} s - x^{-1} \\ \frac{\rho}{x^Ts} x - s^{-1} \end{bmatrix}. \tag{6}$$

There are many variants of primal-dual potential-reduction algorithms that use the TTY potential function. They differ in the way they generate the potential-reduction search directions and in the line search strategies they use along these directions. For example, Todd and Ye use projective scaling to determine the search directions and use line search to keep the iterates approximately centered [14]. Kojima, Mizuno, and Yoshise's algorithm uses a steepest-descent search direction after a primal-dual scaling of the LP problem and does not require the centrality of the iterates [7]. Tütüncü [15, 16] uses modified Newton directions to reduce the TTY function. All the algorithms mentioned have been proven to have polynomial time worst-case complexities by demonstrating that a sufficient decrease in Φ_ρ can be obtained in every iteration of these algorithms.

In the remainder of this article, we will study continuous trajectories that are naturally associated with primal-dual potential-reduction algorithms of Kojima-Mizuno-Yoshise [7] and Tütüncü [15]. For this purpose, we first present the search directions used by these algorithms.

Let (x, s) be a given iterate and let $(\underline{x}, \underline{s})$ be the corresponding reduced iterate. The search direction proposed by Kojima, Mizuno, and Yoshise is the solution of the following system:

$$\begin{array}{rcrcrcl} & & A\Delta x & & & = & 0 \\ A^T\Delta y & & & + & \Delta s & = & 0 \\ & & S\Delta x & + & X\Delta s & = & \frac{x^Ts}{\rho} e - x \circ s, \end{array} \tag{7}$$

where $X = \mathrm{diag}(x)$, $S = \mathrm{diag}(s)$, and e is a vector of ones of appropriate dimension. When we discuss the search direction given by (7) and associated trajectories, we will assume that $\rho > n$. The solution to the system (7) can be regarded as a scaled and projected steepest descent direction for the TTY potential function Φ_ρ, using a primal-dual scaling matrix.

The system (7) can be reduced to the single block equation

$$SG^T\underline{\Delta x} + XA^T\underline{\Delta s} = \frac{x^Ts}{\rho} e - x \circ s$$

and its solution can be presented explicitly using orthogonal projection matrices. First, let

$$D = X^{\frac{1}{2}} S^{-\frac{1}{2}}, \quad \text{and} \quad V = X^{\frac{1}{2}} S^{\frac{1}{2}}. \tag{8}$$

Now define the primal-dual projection matrix Ψ as follows:

$$\begin{aligned} \Psi = \Psi(x,s) \;&=\; I - DA^T(AD^2A^T)^{-1}AD \\ &=\; D^{-1}G^T(GD^{-2}G^T)^{-1}GD^{-1}. \end{aligned} \tag{9}$$

This is the orthogonal projection matrix into the null space of AD which is the same as the range space of $D^{-1}G^T$. Finally, let $\overline{\Psi} = I - \Psi$, $v := Ve$, and $v^{-1} = V^{-1}e$. Then, the solution of (7) is:

$$\begin{bmatrix} \Delta x \\ \Delta s \end{bmatrix} = - \begin{bmatrix} D\Psi \\ D^{-1}\overline{\Psi} \end{bmatrix} \left(v - \frac{v^T v}{\rho} v^{-1} \right) =: g_1(x,s). \tag{10}$$

In the reduced space we have the following solution:

$$\begin{bmatrix} \underline{\Delta x} \\ \underline{\Delta s} \end{bmatrix} = - \begin{bmatrix} (GD^{-2}G^T)^{-1}GD^{-1} \\ (AD^2A^T)^{-1}AD \end{bmatrix} \left(v - \frac{v^T v}{\rho} v^{-1} \right) =: f_1(\underline{x}, \underline{s}). \tag{11}$$

All the terms on the right-hand-side are expressed in terms of x and s, which are linear functions of $\underline{x}$ and $\underline{s}$. Using this correspondence, we express the reduced directions as a function of the reduced iterates.

The search direction used in [15, 16] is a Newton direction for the multiplicative analogue of the TTY potential-function. It can also be seen as a modified Newton direction for the TTY potential-function itself. In [15], it is observed that the multiplicative analogue of the TTY potential-function is a convex function for $\rho \geq 2n$ which motivates the use of Newton directions to reduce this function. Because of this property, we assume that $\rho \geq 2n$ when we discuss the search directions (19) below and trajectories associated with them. Its development is discussed in detail in [15] and like the KMY direction, it can be represented explicitly using orthogonal projection matrices. Let

$$\begin{aligned} \Xi = \Xi(x) &:= X^{-1}G^T(GX^{-2}G^T)^{-1}GX^{-1} \\ &= I - XA^T(AX^2A^T)^{-1}AX, \end{aligned} \tag{12}$$

and

$$\begin{aligned} \Sigma = \Sigma(s) &:= S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1} \\ &= I - SG^T(GS^2G^T)^{-1}GS. \end{aligned} \tag{13}$$

We note that Ξ and Σ are orthogonal projection matrices into the range spaces of $X^{-1}G^T$ and $S^{-1}A^T$, respectively. These range spaces are the same as the null spaces of AX and GS, respectively. To simplify the notation, we will suppress the dependence of Ψ, Ξ, and Σ to (x,s) in our notation. We will denote the normalized complementarity vector with ν:

$$\nu := \frac{x \circ s}{x^T s}. \tag{14}$$

The following quantities appear in the description of the search direction used in [15]:

$$\beta_1 := \nu^T (\Xi + \Sigma) \nu, \tag{15}$$
$$\beta_2 := \nu^T (\Xi + \Sigma) e, \tag{16}$$
$$\beta_3 := e^T (\Xi + \Sigma) e, \tag{17}$$
$$\Delta := (\rho\beta_1 - 1)(\rho - \beta_3 - 1) + \rho(1 - \beta_2)^2. \tag{18}$$

We are now ready to reproduce the search directions introduced in [15]:

$$\begin{bmatrix} \Delta x \\ \Delta s \end{bmatrix} = -\frac{1}{\Delta}\begin{bmatrix} X\Sigma \\ S\Xi \end{bmatrix}(\rho(1-\beta_2)\nu + (\rho\beta_1 - 1)e) =: g_2(x,s). \tag{19}$$

Once again, there is a corresponding description for the reduced search directions:

$$\begin{bmatrix} \underline{\Delta x} \\ \underline{\Delta s} \end{bmatrix} = -\frac{1}{\Delta}\begin{bmatrix} (GX^{-2}G^T)^{-1}GX^{-1} \\ (AS^{-2}A^T)^{-1}AS^{-1} \end{bmatrix}(\rho(1-\beta_2)\nu + (\rho\beta_1 - 1)e) =: f_2(\underline{x},\underline{s}). \tag{20}$$

3 Potential-Reduction Trajectories

Let $(x^0, s^0) \in \mathcal{F}^0$ be any given point in the interior of the feasible region. We will study continuous trajectories of the vector fields $f_i(\underline{x},\underline{s}) : \mathcal{F}_R^0(x^0,s^0) \to \Re^n$ for $i = 1,2$. That is, we will analyze the trajectories determined by the maximal solutions of the following autonomous ordinary differential equations:

$$\begin{bmatrix} \dot{\underline{x}} \\ \dot{\underline{s}} \end{bmatrix} = f_1(\underline{x},\underline{s}), \qquad (\underline{x}(0),\underline{s}(0)) = (0,0), \tag{21}$$

$$\begin{bmatrix} \dot{\underline{x}} \\ \dot{\underline{s}} \end{bmatrix} = f_2(\underline{x},\underline{s}), \qquad (\underline{x}(0),\underline{s}(0)) = (0,0). \tag{22}$$

Note that $(0,0) \in \mathcal{F}_R^0(x^0,s^0)$ for any given $(x^0,s^0) \in \mathcal{F}^0$.

Our objective in this section is to demonstrate that there exist unique solutions to the ODEs defined above and that their trajectories have limit points on the primal-dual optimal face. Monteiro [9] performs a similar study of trajectories based on primal-only potential-reduction algorithms. We will use some of his results and develop primal-dual versions of some others.

For simplicity, we use z and $\underline{z}$ to denote the primal and dual pair of variables (x,s) in the original space and $(\underline{x},\underline{s})$ in the reduced space. Expressions $\underline{z}(t)$ and $\dot{\underline{z}}(t)$ denote $\begin{bmatrix} \underline{x}^T(t) & \underline{s}^T(t) \end{bmatrix}^T$ and $\begin{bmatrix} \dot{\underline{x}}^T(t) & \dot{\underline{s}}^T(t) \end{bmatrix}^T$, respectively.

A differentiable path $\underline{z}(t) : (l,u) \to \mathcal{F}^0$ is a solution to the ODE (21) (or, to the ODE (22)), if $\dot{\underline{z}}(t) = f_1(\underline{z})$ ($\dot{\underline{z}}(t) = f_2(\underline{z})$) for all $t \in (l,u)$ for some scalars l and u such that $l < 0 < u$. This solution is called a *maximal solution* if there is no solution $\hat{z}(t)$ that satisfies the corresponding ODE over a set of t's that properly contain (l,u). The vector fields $f_1(\underline{z})$ and $f_2(\underline{z})$ have the domain $\mathcal{F}_R^0$, an open subset of $\Re^n$ and the image set is $\Re^n$. Note that, both $f_1(\underline{z})$ and $f_2(\underline{z})$ are continuous (in fact, C^∞) functions on $\mathcal{F}^0 \subset \Re^{2n}_{++}$. Given these properties, the following result follows immediately from the standard theory of ordinary differential equations; see, e.g., Theorem 1 on p. 162 and Lemma on p. 171 of the textbook by Hirsch and Smale [5]:

Theorem 3.1 *For each $(\underline{x}^0,\underline{s}^0) \in \mathcal{F}_R^0$ there is a maximum open interval (l_1,u_1) ((l_2,u_2)) containing 0 on which there is a unique solution $\underline{z}_1(t)$ of the ODE (21) ($\underline{z}_2(t)$ of the ODE (22)).*

We define corresponding trajectories in the original problem space as follows:

$$z_1(t) := (x_1(t), s_1(t)) \quad := \quad (x^0, s^0) + (G^T \underline{x}_1(t), A^T \underline{s}_1(t)), \tag{23}$$
$$z_2(t) := (x_2(t), s_2(t)) \quad := \quad (x^0, s^0) + (G^T \underline{x}_2(t), A^T \underline{s}_2(t)). \tag{24}$$

By differentiation, we observe that these trajectories satisfy the following differential equations, respectively:

$$\begin{bmatrix} \dot{x} \\ \dot{s} \end{bmatrix} = g_1(x, s), \qquad (x(0), s(0)) = (x^0, s^0), \tag{25}$$
$$\begin{bmatrix} \dot{x} \\ \dot{s} \end{bmatrix} = g_2(x, s), \qquad (x(0), s(0)) = (x^0, s^0). \tag{26}$$

Now that we established the existence and uniqueness of *potential-reduction trajectories*, we turn our attention to the analysis of their convergence. The following result holds the key to this analysis:

Proposition 3.1 (Proposition 2.2, [9]) *Let U be an open subset of $\Re^n$ and f be a continuously differentiable function from U to $\Re^n$. Assume that there exists a function $h : U \to \Re$ such that for any solution $z(t)$ of the ODE $\dot{z} = f(z)$, the composite function $h(z(t))$ is a strictly decreasing function of t. Then, any maximal solution $z : (l, u) \to U$ of $\dot{z} = f(z)$ satisfies the following property: Given any compact subset $\mathcal{S}$ of U there exists a number $\beta \in (l, u)$ such that $z(t) \notin \mathcal{S}$ for all $t \in [\beta, u)$.*

The proposition says that, if the solution of the ODE can be made monotonic through composition with another function, then the trajectory can not remain in any compact subset of the domain and must tend to the boundary of the domain as t tends to its upper limit. This is a variant of a standard ODE result (see the theorem on p. 171 of [5]). This standard result requires the finiteness of the upper end u of the t domain–Monteiro removes this restriction by placing the condition about the existence of the function h with the stated monotonicity property [9].

We first show that there is indeed a function h satisfying the condition in Proposition 3.1. This function is defined as $h(\underline{x}, \underline{s}) = \Phi_\rho\left((x^0, s^0) + (G^T \underline{x}, A^T \underline{s})\right)$. In other words, we show that the TTY potential function is decreasing along the trajectories of the ODEs (25) and (26):

Lemma 3.1 $\phi_1(t) := \Phi_\rho\left((x^0, s^0) + (G^T \underline{x}_1(t), A^T \underline{s}_1(t))\right) = \Phi_\rho\left(x_1(t), s_1(t)\right)$ *and* $\phi_2(t) := \Phi_\rho\left(x_2(t), s_2(t)\right)$ *are both decreasing functions of t.*

Proof:
Let us start with $\phi_1(t)$. Assume that $(\underline{x}_1(t), \underline{s}_1(t))$ satisfies (21) for all $t \in (l, u)$. Then, $(x_1(t), s_1(t))$ satisfies (25) for all $t \in (l, u)$ and for any such t,

$$\begin{aligned} \phi_1'(t) &= \nabla \Phi_\rho^T\left(x_1(t), s_1(t)\right) \begin{bmatrix} \dot{x}_1(t) \\ \dot{s}_1(t) \end{bmatrix} \\ &= -\begin{bmatrix} \frac{\rho}{x^T s} s - x^{-1} \\ \frac{\rho}{x^T s} x - s^{-1} \end{bmatrix}^T \begin{bmatrix} D\Psi \\ D^{-1}\overline{\Psi} \end{bmatrix} \left(v - \frac{v^T v}{\rho} v^{-1}\right), \end{aligned}$$

where we wrote x and s in place of $x_1(t)$ and $s_1(t)$ to simplify the notation. Let $w = \frac{\rho}{x^T s} xs - e$. Then,

$$\begin{aligned} \phi_1'(t) &= -\frac{x^T s}{\rho} w^T \left(X^{-1} D \Psi V^{-1} \right) w - \frac{x^T s}{\rho} w^T \left(S^{-1} D^{-1} \overline{\Psi} V^{-1} \right) w \\ &= -\frac{x^T s}{\rho} (V^{-1} w)^T \left(\Psi + \overline{\Psi} \right) (V^{-1} w) \\ &= -\frac{x^T s}{\rho} \| V^{-1} w \|^2, \end{aligned}$$

where we used $X^{-1} D = S^{-1} D^{-1} = V^{-1}$. Further, note that $w = 0$ would require $\frac{\rho}{x^T s} xs = e$ which implies $\rho = n$. Since we assumed that $\rho > n$ for the KMY direction, we have that $w \neq 0$ and $V^{-1} w \neq 0$ and that $\phi_1'(t) < 0$. Therefore, $\phi_1(t)$ is a decreasing function of t.

We proceed similarly for $\phi_2(t)$:

$$\phi_2'(t) = \nabla \Phi_\rho^T (x_2(t), s_2(t)) \begin{bmatrix} \dot{x}_2(t) \\ \dot{s}_2(t) \end{bmatrix}.$$

By construction, the direction vector $g_2(x, s)$ introduced in [15] satisfies $g_2(x, s) = -\left[\nabla^2 F_\rho(x, s) \right]^{-1} \nabla F_\rho(x, s)$ where $F_\rho(x, s) = \exp\{\Phi_\rho(x, s)\}$. Further, $\nabla F_\rho(x, s) = F_\rho(x, s) \nabla \Phi_\rho(x, s)$. The function $F_\rho(x, s)$ is strictly convex on $\mathcal{F}^0$ if $\rho \geq 2n$ (see [15]), which we assumed, and therefore, its Hessian is positive definite. Thus,

$$\phi_2'(t) = -\frac{1}{F_\rho(x, s)} \nabla F_\rho^T (x, s) \left[\nabla^2 F_\rho(x, s) \right]^{-1} \nabla F_\rho(x, s) < 0.$$

Above, we again wrote x and s in place of $x_2(t)$ and $s_2(t)$ to simplify the notation. Note also that $\nabla \Phi_\rho^T(x, s) \neq 0$ because $\nabla \Phi_\rho^T(x, s) = 0$ would require $\frac{\rho}{x^T s} xs = e$ which, as above, is not possible since we assume $\rho \geq 2n$. Thus, $\phi_2(t)$ is also a decreasing function of t. ∎

We prove two technical results before presenting the main theorem of this section:

Lemma 3.2 *Given $\rho > n$, $\beta > 0$ and M, the set*

$$\mathcal{S} := \{(x, s) \in \mathcal{F}^0 : x^T s \geq \beta, \Phi_\rho(x, s) \leq M\} \tag{27}$$

is compact.

Proof:
We first show that $\mathcal{S}$ is a bounded subset of $\mathcal{F}^0$. Given $(x, s) \in \mathcal{F}^0$ and $\rho > n$, the following inequality is well known and follows from the arithmetic mean-geometric mean inequality (see, e.g., [17, Lemma 4.2]):

$$\Phi_\rho(x, s) \geq (\rho - n) \ln(x^T s) + n \ln n.$$

Therefore, $\Phi_\rho(x, s) \leq M$ implies that

$$x^T s \leq \exp\left(\frac{M - n \ln n}{\rho - n} \right). \tag{28}$$

Fix $(\hat{x}, \hat{s}) \in \mathcal{F}^0$. This is possible since we assumed that $\mathcal{F}^0$ is nonempty. Since $(x-\hat{x})^T(s-\hat{s}) = 0$, we have that $x^T s = \hat{x}^T s + \hat{s}^T x - \hat{x}^T \hat{s}$. Combining this identity with (28) we obtain

$$\hat{x}^T s + \hat{s}^T x = \sum_{j=1}^{n} \hat{x}_j s_j + \sum_{j=1}^{n} \hat{s}_j x_j \quad \leq \quad \hat{x}^T \hat{s} + \exp\left(\frac{M - n \ln n}{\rho - n}\right) := M_1$$

for every $(x, s) \in \mathcal{S}$. Since all terms on the left in the inequality above are positive, we have that

$$x_j \leq \frac{M_1}{\hat{s}_j}, \quad \forall j, \qquad s_j \leq \frac{M_1}{\hat{x}_j}, \quad \forall j.$$

Thus, we proved that x and s are bounded above if they are in $\mathcal{S}$. Let $U = \max_j\{\max(\frac{M_1}{\hat{s}_j}, \frac{M_1}{\hat{x}_j})\}$. Now,

$$\begin{aligned} \ln x_j &= \rho \ln(x^T s) - \sum_{i \neq j}^{n} \ln(x_i s_i) - \ln s_j - \Phi_\rho(x, s) \\ &\geq \rho \ln \beta - (2n-1) \ln U - M =: M_2, \end{aligned}$$

if $(x, s) \in \mathcal{S}$. Therefore,

$$x_j \geq e^{M_2}, \quad \forall j, \qquad s_j \geq e^{M_2}, \quad \forall j,$$

where the inequality for s_j follows from identical arguments. Thus, all components of the vectors $(x, s) \in \mathcal{S}$ are bounded below by a positive constant. This proves that $\mathcal{S}$ is a bounded subset of $\mathcal{F}^0$. Since $\mathcal{S}$ is clearly closed with respect to $\mathcal{F}^0$ we must have that $\mathcal{S}$ is a compact subset of $\mathcal{F}^0$. □

Corollary 3.1 *Given $\rho > n$, $\beta > 0$ and M, the set*

$$\mathcal{S}_R := \{(\underline{x}, \underline{s}) \in \mathcal{F}_R^0 : (x^0 + G^T \underline{x})^T (s^0 + A^T \underline{s}) \geq \beta, \Phi_\rho(x^0 + G^T \underline{x}, s^0 + A^T \underline{s}) \leq M\}$$

is compact.

Proof:
This is immediate since $\mathcal{S}_R$ is obtained by a linear transformation of the compact set $\mathcal{S}$. □

Now, we are ready to establish that the duality gap in our potential-reduction trajectories converges to zero:

Theorem 3.2 *Let $\underline{z}_1(t) : (l_1, u_1) \to \mathcal{F}_R^0$ and $\underline{z}_2(t) : (l_2, u_2) \to \mathcal{F}_R^0$ be maximal solutions of the ODEs (21) and (22). Let $(x_1(t), s_1(t))$ and $(x_2(t), s_2(t))$ be as in equations (23)-(24). Then,*

$$\lim_{t \to u_i} x_i(t)^T s_i(t) = 0$$

for $i = 1, 2$.

Proof:
We prove the theorem for $i = 1$; the proof for $i = 2$ is identical. Since $\underline{z}_1(t) \in \mathcal{F}_R^0$, $(x_1(t), s_1(t))$ remain in $\mathcal{F}^0$ for all $t \in (l_1, u_1)$. Therefore, $x_1(t)^T s_1(t) > 0$ for all $t \in (l_1, u_1)$. Then, if $\lim_{t\to u_1} x_1(t)^T s_1(t) \neq 0$ we must have an infinite sequence (t_k) such that $\lim_{k\to\infty} t_k = u_1$ and that $x_1(t_k)^T s_1(t_k)$ remains bounded away from zero. That is, there exists a positive β such that $x_1(t_k)^T s_1(t_k) \geq \beta$ for all k. Since we already established that $\Phi_\rho(x_1(t), s_1(t))$ is a decreasing function of t, by choosing $M = \Phi_\rho(x_1(t_0), s_1(t_0))$ we have that $(x_1(t_k), s_1(t_k)) \in \mathcal{S}$ where $\mathcal{S}$ is as in (27). Using Lemma 3.2 and its corollary we conclude that the assumption $\lim_{t\to u_1} x_1(t)^T s_1(t) \neq 0$ implies the existence of a compact subset of $\Re^n$ to which the maximal solution $\underline{z}_1(t)$ of the ODE (21) always returns to as t tends to u_1. But this contradicts Proposition 3.1. Therefore, the assumption must be wrong and we have that

$$\lim_{t\to u_1} x_1(t)^T s_1(t) = 0.$$

This completes the proof. ∎

4 Convergence of the Trajectories

The last result of the previous section demonstrates the convergence of the duality gap to zero along continuous trajectories we associated with two different primal-dual potential-reduction algorithms. This theorem, however, does not guarantee the convergence of these trajectories to an optimal solution. All accumulation points of these trajectories will be optimal solutions and there will be a unique accumulation point if the optimal solution is unique. However, we need a more careful analysis to establish the convergence of these trajectories when the primal-dual optimal solution set is not a singleton and this is the purpose of the current section.

Not surprisingly, the difficulty mentioned in the previous paragraph regarding the proof of convergence of trajectories is similar to the difficulties encountered in convergence proofs for potential-reduction interior-point algorithms. To overcome this difficulty, we want to find an implicit description of the trajectories, perhaps similar to the description of the *central path* or *weighted centers*. We present such a description for trajectories associated with the Kojima-Mizuno-Yoshise algorithm. The complicated vector field associated with the algorithm in [15, 16] appears difficult to analyze in a similar fashion and we leave that as a topic of future research. Before proceeding with the analysis of the Kojima-Mizuno-Yoshise trajectories, we establish that the central path belongs to both families of potential-reduction trajectories we defined in the previous section.

4.1 The Central Path

The central path $\mathcal{C}$ of the primal-dual feasible set $\mathcal{F}$ is the set of points on which the componentwise product of the primal and dual variables is constant:

$$\mathcal{C} \; := \; \{(x(\mu), s(\mu)) \in \mathcal{F}^0 : x(\mu) \circ s(\mu) = \mu e, \text{ for some } \mu > 0\}. \tag{29}$$

The points on the central path are obtained as unique minimizers of certain barrier problems associated with the primal and dual LPs and they converge to the analytic center of the primal-dual optimal face; see, e.g., [17]. The central path or its subsets can be parametrized in different ways. For example, let $\phi(t) : (l,u) \to \Re$ be a positive valued function. Then, we can define the following parameterization of the corresponding subset of $\mathcal{C}$:

$$\mathcal{C}_\phi \quad := \quad \{(x(t),s(t)) \in \mathcal{F}^0 : x(t) \circ s(t) = \phi(t)e, \text{ for some } t \in (l,u)\}. \tag{30}$$

As long as the function ϕ used in the parameterization of the central path is differentiable, the central path and corresponding subsets of it are differentiable with respect to the underlying parameter. One can easily verify that, we have the following identities for $(x(t),s(t)) \in \mathcal{C}_\phi$:

$$A\dot{x}(t) = 0, \quad G\dot{s}(t) = 0, \tag{31}$$

$$\dot{x}(t) \circ s(t) + x(t) \circ \dot{s}(t) = \phi'(t)e, \tag{32}$$

for $t \in (l,u)$. From (31)-(32) it follows that

$$\begin{bmatrix} \dot{x}(t) \\ \dot{s}(t) \end{bmatrix} = \begin{bmatrix} D(t)\Psi(t) \\ D(t)^{-1}\overline{\Psi(t)} \end{bmatrix} \phi'(t)v(t)^{-1} \tag{33}$$

$$= \begin{bmatrix} D(t)\Psi(t) \\ D(t)^{-1}\overline{\Psi(t)} \end{bmatrix} \frac{\phi'(t)}{\sqrt{\phi(t)}} e \tag{34}$$

where $D(t), v(t)^{-1}, \Psi(t)$, and $\overline{\Psi(t)}$ are defined identically to D, v^{-1}, Ψ, and $\overline{\Psi}$ in Section 2, except that x and s are replaced by $x(t)$ and $s(t)$. Next, we make the following observation:

$$\begin{aligned} g_1(x(t),s(t)) &= -\begin{bmatrix} D(t)\Psi(t) \\ D(t)^{-1}\overline{\Psi(t)} \end{bmatrix} \left(v(t) - \frac{v(t)^T v(t)}{\rho} v(t)^{-1} \right) \\ &= -\begin{bmatrix} D(t)\Psi(t) \\ D(t)^{-1}\overline{\Psi(t)} \end{bmatrix} \sqrt{\phi(t)} \left(1 - \frac{n}{\rho}\right) e. \end{aligned}$$

Therefore, if we choose $\phi(t)$ such that

$$\frac{\phi'(t)}{\sqrt{\phi(t)}} = -\sqrt{\phi(t)} \left(1 - \frac{n}{\rho}\right) \tag{35}$$

we will have the subset $\mathcal{C}_\phi$ of the central path $\mathcal{C}$ satisfy the ODE (25). Let us denote the solution of the differential equation (35) with $\phi_1(t)$, which is easily computed:

$$\phi_1(t) = \phi_1(0) \cdot \exp\{-\left(1 - \frac{n}{\rho}\right) t\}. \tag{36}$$

So, for any given (x^0, s^0) on the central path $\mathcal{C}$ with $x^0 \circ s^0 = \mu e$, we can choose $\phi_1(0) = \mu$ and the unique maximal solution of the ODE (25) is obtained as the trajectory

$$\mathcal{C}_{\phi_1} = \{(x(t),s(t)) \in \mathcal{F}^0 : x(t) \circ s(t) = \mu \exp\{-(1 - \frac{n}{\rho})t\}e, \text{ for } t \in (l_1, \infty)\}$$

for some $l_1 < 0$. The fact that the upper bound on the range of t is ∞ follows from Theorem 3.2—as t tends to the upper bound, $x(t)^T s(t) = n\mu \exp\{-\left(1 - \frac{n}{\rho}\right) t\}$ must approach zero.

We can use the same approach to show that subsets of the central path are solutions to the ODE (26). We first evaluate $g_2(x(t), s(t))$ for $(x(t), s(t)) \in \mathcal{C}_\phi$. We make the following observations using $x(t) \circ s(t) = \phi(t)e$:

$$\begin{gathered}\nu = \tfrac{1}{n}e,\ \Xi = \Psi(t),\ \Sigma = \overline{\Psi(t)},\ \Xi + \Sigma = I,\\ \beta_1 = \tfrac{1}{n},\ \beta_2 = 1,\ \beta_3 = n,\ \Delta = (\tfrac{\rho}{n} - 1)(\rho - n - 1),\\ X = \sqrt{\phi(t)}D(t),\ S = \sqrt{\phi(t)}D^{-1}(t).\end{gathered}$$

Therefore,

$$\begin{aligned} g_2(x(t), s(t)) &= -\frac{1}{(\frac{\rho}{n} - 1)(\rho - n - 1)}\sqrt{\phi(t)}\begin{bmatrix} D(t)\Psi(t) \\ D(t)^{-1}\overline{\Psi(t)} \end{bmatrix}\left(\frac{\rho}{n} - 1\right)e \\ &= -\begin{bmatrix} D(t)\Psi(t) \\ D(t)^{-1}\overline{\Psi(t)} \end{bmatrix}\frac{\sqrt{\phi(t)}}{\rho - n - 1}e. \end{aligned}$$

This time we need to choose $\phi(t)$ such that

$$\frac{\phi'(t)}{\sqrt{\phi(t)}} = -\frac{\sqrt{\phi(t)}}{\rho - n - 1} \tag{37}$$

to have the subset $\mathcal{C}_\phi$ of the central path $\mathcal{C}$ satisfy the ODE (26). The solution $\phi_2(t)$ of (37) is:

$$\phi_2(t) = \phi_2(0) \cdot \exp\{-\frac{t}{\rho - n - 1}\}. \tag{38}$$

As above, for any given (x^0, s^0) on the central path $\mathcal{C}$ with $x^0 \circ s^0 = \mu e$, we can choose $\phi_2(0) = \mu$ and the unique maximal solution of the ODE (26) is obtained as the trajectory

$$\mathcal{C}_{\phi_2} = \{(x(t), s(t)) \in \mathcal{F}^0 : x(t) \circ s(t) = \mu \exp\{-\frac{t}{\rho - n - 1}\}e,\ \text{for } t \in (l_2, \infty)\}$$

for some $l_2 < 0$. Thus, we proved the following result:

Theorem 4.1 *For any given (x^0, s^0) on the central path $\mathcal{C}$, the solution of both the ODEs (25) and (26) with initial condition $(x(0), s(0)) = (x^0, s^0)$ is a trajectory that is a subset of $\mathcal{C}$.* □

Theorem 4.1 provides a theoretical basis for the practical observation that central path-following search directions are often very good potential-reduction directions as well. Another related result is by Nesterov [10] who observes that the neighborhood of the central path is the region of fastest decrease for a homogeneous potential function. (Φ_ρ is homogeneous when $\rho = 2n$).

4.2 A Characterization of the Trajectories

Now we turn our attention to the trajectories $(x_1(t), s_1(t))$ of the ODE (25) that do not start on the central path. From this point on, we will not consider the trajectories $(x_2(t), s_2(t))$ and therefore, drop the subscript 1 from expressions like $x_1(t)$, $s_1(t)$ for simplicity. Consider

$(x(t), s(t))$ for $t \in (l, u)$, the maximal solution of the ODE (25). We define the following function of t:

$$\gamma(t) \quad := \quad x(t) \circ s(t) - h(t)e \tag{39}$$

where $x(t) \circ s(t)$ denotes the Hadamard product of the vectors $x(t)$ and $s(t)$, e is an n-dimensional vector of ones, and $h(t)$ is a scalar valued function of t that we will determine later. The function $\gamma(t)$ is defined from (l, u) to $\Re^n$. Let $\phi(t) = x(t) \circ s(t) - \frac{x(t)^T s(t)}{\rho} e$. Differentiating (39) with respect to t, we obtain:

$$\begin{aligned}\dot{\gamma}(t) &= \dot{x}(t) \circ s(t) + \dot{s}(t) \circ x(t) - \dot{h}(t)e \\ &= -V\Psi V^{-1}\phi(t) - V\overline{\Psi}V^{-1}\phi(t) - \dot{h}(t)e \\ &= -\phi(t) - \dot{h}(t)e \\ &= -x(t) \circ s(t) + \frac{x(t)^T s(t)}{\rho} e - \dot{h}(t)e.\end{aligned}$$

Now, if we were to choose $h(t)$ such that

$$h(t) \quad = \quad \frac{x(t)^T s(t)}{\rho} - \dot{h}(t) \tag{40}$$

we would have that

$$\dot{\gamma}(t) \quad = \quad -x(t) \circ s(t) + h(t)e = -\gamma(t). \tag{41}$$

This last equation has an easy solution:

$$\gamma(t) \quad = \quad \gamma(0) \cdot e^{-t}.$$

Therefore, to get a better description of the solution trajectory $(x(t), s(t))$, it suffices to find a scalar valued function $h(t)$ that satisfies (40). For this purpose, we use the standard "variation of parameters" technique and look for a solution of the form $h(t) = h_1(t)e^{-t}$. Then,

$$\begin{aligned}\frac{x(t)^T s(t)}{\rho} &= h(t) + \dot{h}(t) \\ &= h_1(t)e^{-t} + \dot{h_1}(t)e^{-t} - h_1(t)e^{-t} \\ &= \dot{h_1}(t)e^{-t}.\end{aligned}$$

So, setting the constant that comes from integration arbitrarily to 0, we obtain

$$h_1(t) = \int_0^t e^\tau \frac{x(\tau)^T s(\tau)}{\rho} d\tau, \text{ and } h(t) = e^{-t} \int_0^t e^\tau \frac{x(\tau)^T s(\tau)}{\rho} d\tau. \tag{42}$$

Note that, $h(0) = 0$. Given (x^0, s^0), let $p := \gamma(0) = x^0 \circ s^0$. Then, $\gamma(t) = p \cdot e^{-t}$. The equation (42) describing $h(t)$ is not very useful to describe $(x(t), s(t))$ since it involves these terms in its definition. Fortunately, we can simplify (42). First, observe from (39) that $x(t)^T s(t) = (e^T p)e^{-t} + n \cdot h(t)$. Now, consider

$$\begin{aligned} h(t)e^t &= \int_0^t e^\tau \frac{x(\tau)^T s(\tau)}{\rho} d\tau = \int_0^t e^\tau \frac{(e^T p)e^{-\tau} + n \cdot h(\tau)}{\rho} d\tau \\ &= \frac{e^T p}{\rho} t + \frac{n}{\rho} \int_0^t h(\tau) e^\tau d\tau. \end{aligned}$$

This is an integral equation involving the function $h(t)e^t$. Applying the variation of parameters technique again and integrating by parts we obtain the solution of this equation:

$$\begin{aligned} h(t)e^t &= \frac{e^T p}{n} \left(e^{\frac{n}{\rho} t} - 1 \right), \text{ and,} \\ h(t) &= \frac{e^T p}{n} \left(\exp\left\{ -\left(1 - \frac{n}{\rho}\right) t \right\} - e^{-t} \right). \end{aligned} \tag{43}$$

One can easily verify that when (x^0, s^0) is on the central path $\mathcal{C}$ with $x^0 \circ s^0 = \mu e$, using (43) one recovers the following identity we derived in the previous subsection: $x(t) \circ s(t) = \mu \exp\{-\left(1 - \frac{n}{\rho}\right) t\} e$, for $t \in (l, \infty)$. Using (39) and (43) we obtain the following identity:

$$\begin{aligned} x(t)^T s(t) &= (e^T p)e^{-t} + nh(t) = (e^T p) \exp\left\{ -\left(1 - \frac{n}{\rho}\right) t \right\}, \\ &= x(0)^T s(0) \exp\left\{ -\left(1 - \frac{n}{\rho}\right) t \right\}. \end{aligned} \tag{44}$$

Proposition 4.1 *Consider $(x(t), s(t))$ for $t \in (l, u)$, the maximal solution of the ODE (25) and $h(t)$ defined in (43). Then,*

$$u = +\infty.$$

Furthermore, $(x(t), s(t))$ remains bounded as $t \to \infty$.

Proof:
First part of the proposition follows immediately from (44) using the conclusion of Theorem 3.2. For the boundedness result, observe that $\left(x(t) - x^0\right)^T \left(s(t) - s^0\right) = 0$. Therefore,

$$\begin{aligned} \sum_{j=1}^n \left(x_j^0 s(t)_j + s_j^0 x(t)_j \right) &= (x^0)^T s^0 + x^T(t) s(t) = e^T p \left(1 + \exp\left\{ -\left(1 - \frac{n}{\rho}\right) t \right\} \right) \\ &\leq 2e^T p. \end{aligned}$$

Since all the terms in the summation are nonnegative and $x^0 > 0$, $s^0 > 0$, the result follows. □

Note also that $\lim_{t\to\infty} h(t) = 0$. Let $w_j(t) = p_j e^{-t} + h(t)$ and $w(t) = [w_j(t)]$. Observe that $w_j(t) > 0 \ \forall\, j, t$. We now have that

$$Ax(t) = b, \ Gs(t) = Gc, \ x(t) \circ s(t) = w(t), \ \forall t \in (l, u), \tag{45}$$

from which the following theorem follows:

Theorem 4.2 *Consider $(x(t), s(t))$ for $t \in (l, u)$, the maximal solution of the ODE (25). For each $t \in (l, u)$, $x(t)$ is the unique solution of the following primal barrier problem:*

$$
(PBP) \qquad \begin{array}{rrcl} \min_x & c^T x - \sum_{j=1}^n w_j(t) \ln x_j & & \\ & Ax & = & b \\ & x & > & 0. \end{array} \tag{46}
$$

Similarly, for each $t \in (l, u)$, $s(t)$ is the unique solution of the following dual barrier problem:

$$
(LBP) \qquad \begin{array}{rrcl} \min_s & d^T s - \sum_{j=1}^n w_j(t) \ln s_j & & \\ & Gs & = & Gc \\ & s & > & 0. \end{array} \tag{47}
$$

Proof:
Optimality conditions of both barrier problems in the statement of the theorem are equivalent to (45); see, eg., [17]. Since both barrier problems have strictly convex objective functions, their optimal solutions are unique and satisfy (45). Therefore, $x(t)$ and $s(t)$ must be the optimal solutions of the respective problems. $\square$

Next, we observe that $\lim_{t\to\infty} \frac{w_j(t)}{h(t)} = 1$ for all j. In other words, the weights w_j in problems (PBP) and (DBP) are asymptotically uniform. Then, problems (PBP) and (DBP) resemble the barrier problems defining the central path, and one might expect that the solutions of these problems converge to the limit point of the central path: the analytic center of the optimal face. We will show below that this is indeed the case. First, we give relevant definitions: Recall the partition $\mathcal{B}$ and $\mathcal{N}$ of the index set of the variables that was given in Section 2. Let $(x^*, s^*) = ((x^*_{\mathcal{B}}, 0), (0, s^*_{\mathcal{N}}))$ denote the the analytic center of the primal-dual optimal face. That is, $x^*_{\mathcal{B}}$ and $s^*_{\mathcal{N}}$ are unique maximizers of the following problems:

$$
\begin{array}{rc} \max & \sum_{j\in\mathcal{B}} \ln x_j \\ & A_{\mathcal{B}} x_{\mathcal{B}} = b \\ & x_{\mathcal{B}} > 0, \end{array} \quad \text{and} \quad \begin{array}{rc} \max & \sum_{j\in\mathcal{N}} \ln s_j \\ & G_{\mathcal{N}} s_{\mathcal{N}} = Gc \\ & s_{\mathcal{N}} > 0. \end{array} \tag{48}
$$

The next lemma will be useful in the proof of the theorem that follows it and also in the next section:

Lemma 4.1 *Let $(x(t), s(t))$ for $t \in (l, \infty)$ denote the maximal solution of the ODE (25). Then $x_{\mathcal{B}}(t)$ and $s_{\mathcal{N}}(t)$ solve the following pair of problems:*

$$
\begin{array}{rc} \max & \sum_{j\in\mathcal{B}} w_j(t) \ln x_j \\ & A_{\mathcal{B}} x_{\mathcal{B}} = b - A_{\mathcal{N}} x_{\mathcal{N}}(t) \\ & x_{\mathcal{B}} > 0, \end{array} \quad \textit{and} \quad \begin{array}{rc} \max & \sum_{j\in\mathcal{N}} w_j(t) \ln s_j \\ & G_{\mathcal{N}} s_{\mathcal{N}} = Gc - G_{\mathcal{B}} s_{\mathcal{B}}(t) \\ & s_{\mathcal{N}} > 0. \end{array} \tag{49}
$$

Proof:
We prove the optimality of $x_{\mathcal{B}}(t)$ for the first problem in (49)–the corresponding result for $s_{\mathcal{N}}(t)$ can be proven similarly. $x_{\mathcal{B}}(t)$ is clearly feasible for the given problem. It is optimal if and only if there exists $y \in \Re^m$ such that

$$
w_{\mathcal{B}}(t) \circ x_{\mathcal{B}}^{-1}(t) = A_{\mathcal{B}}^T y.
$$

From (45) we obtain $w_{\mathcal{B}}(t) \circ x_{\mathcal{B}}^{-1}(t) = s_{\mathcal{B}}(t)$. Note that, for any s feasible for (LD') we have that $c - s \in \mathcal{R}(A^T)$ and therefore, $c_{\mathcal{B}} - s_{\mathcal{B}} \in \mathcal{R}(A_{\mathcal{B}}^T)$. Furthermore, since $s^* = (0, s_{\mathcal{N}}^*)$ is also feasible for (LD') we must have that $c_{\mathcal{B}} \in \mathcal{R}(A_{\mathcal{B}}^T)$ and that $s_{\mathcal{B}}(t) \in \mathcal{R}(A_{\mathcal{B}}^T)$. This is exactly what we needed. ∎

Now we are ready to present the main result of this section:

Theorem 4.3 *Let $(x(t), s(t))$ for $t \in (l, \infty)$ denote the maximal solution of the ODE (25). Then, $(x(t), s(t))$ converges to the analytic center of the primal-dual optimal face Ω.*

Proof:
Proposition 4.1 indicates that the trajectory $(x(t), s(t))$ is bounded as $t \to \infty$. Therefore, it must have at least one accumulation point, say, $(\hat{x}, \hat{s})$. Let (t^k) be a sequence such that $\left(x(t^k), s(t^k)\right) \to (\hat{x}, \hat{s})$ as $k \to \infty$ and let (x^k, s^k) denote $\left(x(t^k), s(t^k)\right)$.

Using Theorem 3.2, we have that $(\hat{x}, \hat{s}) \in \Omega$. We want to show that $(\hat{x}, \hat{s}) = (x^*, s^*)$. Let $(\Delta x, \Delta s) = (x^*, s^*) - (\hat{x}, \hat{s})$ and $(\tilde{x}^k, \tilde{s}^k) = (x^k, s^k) + (\Delta x, \Delta s)$. Note that $(\Delta x_{\mathcal{N}}, \Delta s_{\mathcal{B}}) = 0$ and therefore, $(\tilde{x}_{\mathcal{N}}^k, \tilde{s}_{\mathcal{B}}^k) = (x_{\mathcal{N}}^k, s_{\mathcal{B}}^k)$. Note also that, $(\tilde{x}^k, \tilde{s}^k) \to (x^*, s^*)$, and therefore, for k large enough $(\tilde{x}_{\mathcal{B}}^k, \tilde{s}_{\mathcal{N}}^k) > 0$. Thus, for such k, $\tilde{x}_{\mathcal{B}}^k$ is feasible for the first problem in (49) and using Lemma 4.1 we must have

$$\sum_{j \in \mathcal{B}} \frac{w_j(t)}{h(t)} \ln \tilde{x}_j^k \quad \leq \quad \sum_{j \in \mathcal{B}} \frac{w_j(t)}{h(t)} \ln x_j^k.$$

This inequality must hold in the limit as well. Recall that $\lim_{t \to \infty} \frac{w_j(t)}{h(t)} = 1$ for all j. Therefore, we must have that $\hat{x}_j > 0$ for all $j \in \mathcal{B}$ (otherwise the right-hand term would tend to $-\infty$ while the left-hand term does not), and

$$\sum_{j \in \mathcal{B}} \ln x_j^* \quad \leq \quad \sum_{j \in \mathcal{B}} \ln \hat{x}_j.$$

Since $x_{\mathcal{B}}^*$ is the unique maximizer of the first problem in (48), we conclude that $\hat{x}_{\mathcal{B}} = x_{\mathcal{B}}^*$. Similarly, $\hat{s}_{\mathcal{N}} = s_{\mathcal{N}}^*$. ∎

5 Asymptotic Behavior of the Trajectories

Our analysis in the previous section revealed that all primal-dual potential-reduction trajectories $(x(t), s(t))$ that solve the differential equation (25) converge to the analytic center (x^*, s^*) of the primal-dual optimal face Ω. In this section, we investigate the direction of convergence for these trajectories. That is, we want to analyze the limiting behavior of the normalized vectors $\left(\frac{\dot{x}(t)}{\|\dot{x}(t)\|}, \frac{\dot{s}(t)}{\|\dot{s}(t)\|}\right)$.

By definition, $(\dot{x}(t), \dot{s}(t))$ satisfy equation (25). However, it turns out to be easier to work with the equations (45) to analyze the limiting behavior of $(\hat{x}, \hat{s})$. Differentiating the identity

$$x(t) \circ s(t) \quad = \quad w(t) = e^{-t} p + h(t) e,$$

we obtain

$$\begin{aligned} x(t) \circ \dot{s}(t) + \dot{x}(t) \circ s(t) \quad &= \quad -e^{-t} p + \dot{h}(t) e \\ &= \quad -e^{-t} p - \frac{e^T p}{n} \left(1 - \frac{n}{\rho}\right) \exp\left\{-\left(1 - \frac{n}{\rho}\right) t\right\} e + \frac{e^T p}{n} e^{-t} e. \quad (50) \end{aligned}$$

For our asymptotic analysis, we express certain components of the vectors $\dot{x}(t)$ and $\dot{s}(t)$ in terms of the others. We first need to introduce some notation:

$$\begin{array}{ll}
\hat{w}_{\mathcal{B}}(t) = w_{\mathcal{B}}(t)\exp\{(1-\frac{n}{\rho})t\}, & \hat{w}_{\mathcal{N}}(t) = w_{\mathcal{N}}(t)\exp\{(1-\frac{n}{\rho})t\},\\
d_{\mathcal{B}}(t) = \hat{w}_{\mathcal{B}}^{\frac{1}{2}}(t)\circ x_{\mathcal{B}}^{-1}(t), & d_{\mathcal{N}}(t) = \hat{w}_{\mathcal{N}}^{\frac{1}{2}}(t)\circ s_{\mathcal{N}}^{-1}(t),\\
D_{\mathcal{B}}(t) = \mathrm{diag}\,(d_{\mathcal{B}}(t)), & D_{\mathcal{N}}(t) = \mathrm{diag}\,(d_{\mathcal{N}}(t)),\\
d_{\mathcal{B}}^{-1}(t) = D_{\mathcal{B}}^{-1}(t)e, & d_{\mathcal{N}}^{-1}(t) = D_{\mathcal{N}}^{-1}(t)e,\\
\tilde{A}_{\mathcal{B}}(t) = A_{\mathcal{B}}D_{\mathcal{B}}^{-1}(t), & \tilde{G}_{\mathcal{N}}(t) = G_{\mathcal{N}}D_{\mathcal{N}}^{-1}(t),\\
\tilde{x}_{\mathcal{B}}(t) = D_{\mathcal{B}}(t)\dot{x}_{\mathcal{B}}(t) = d_{\mathcal{B}}(t)\circ\dot{x}_{\mathcal{B}}(t), & \tilde{s}_{\mathcal{N}}(t) = D_{\mathcal{N}}(t)\dot{s}_{\mathcal{N}}(t) = d_{\mathcal{N}}(t)\circ\dot{s}_{\mathcal{N}}(t),\\
\tilde{p}_{\mathcal{B}}(t) = \hat{w}_{\mathcal{B}}^{-\frac{1}{2}}(t)\circ p_{\mathcal{B}}, & \tilde{p}_{\mathcal{N}}(t) = \hat{w}_{\mathcal{N}}^{-\frac{1}{2}}(t)\circ p_{\mathcal{N}}.
\end{array}$$

Lemma 5.1 *Consider $(\dot{x}(t),\dot{s}(t))$ where $(x(t),s(t))$ for $t\in(l,\infty)$ denote the maximal solution of the ODE (25). Then, the following equalities hold:*

$$D_{\mathcal{B}}(t)\dot{x}_{\mathcal{B}}(t) = -\tilde{A}_{\mathcal{B}}^{+}(t)A_{\mathcal{N}}\dot{x}_{\mathcal{N}}(t) - \frac{n\cdot e^{-\frac{n}{\rho}t}}{\rho\left(1-e^{-\frac{n}{\rho}t}\right)}\left(I-\tilde{A}_{\mathcal{B}}^{+}(t)\tilde{A}_{\mathcal{B}}(t)\right)\tilde{p}_{\mathcal{B}}, \tag{51}$$

$$D_{\mathcal{N}}(t)\dot{s}_{\mathcal{N}}(t) = -\tilde{G}_{\mathcal{N}}^{+}(t)G_{\mathcal{B}}\dot{s}_{\mathcal{B}}(t) - \frac{n\cdot e^{-\frac{n}{\rho}t}}{\rho\left(1-e^{-\frac{n}{\rho}t}\right)}\left(I-\tilde{G}_{\mathcal{N}}^{+}(t)\tilde{G}_{\mathcal{N}}(t)\right)\tilde{p}_{\mathcal{N}}. \tag{52}$$

Here, $\tilde{A}_{\mathcal{B}}^{+}(t)$ and $\tilde{G}_{\mathcal{N}}^{+}(t)$ denote the pseudo-inverse of $\tilde{A}_{\mathcal{B}}(t)$ and $\tilde{G}_{\mathcal{N}}(t)$, respectively.

Proof:
We only prove the first identity; the second one follows similarly. Recall from Lemma 4.1 that $x_{\mathcal{B}}(t)$ solves the first problem in (49). Therefore, as in the proof of Lemma 4.1, we must have that

$$w_{\mathcal{B}}(t)\circ x_{\mathcal{B}}^{-1}(t)\in\mathcal{R}(A_{\mathcal{B}}^T). \tag{53}$$

Differentiating with respect to t we obtain:

$$\begin{array}{c}
w_{\mathcal{B}}(t)\circ x_{\mathcal{B}}^{-2}(t)\circ\dot{x}_{\mathcal{B}}(t) - \dot{w}_{\mathcal{B}}(t)\circ x_{\mathcal{B}}^{-1}(t)\in\mathcal{R}(A_{\mathcal{B}}^T), \quad \text{or,}\\
w_{\mathcal{B}}(t)\circ x_{\mathcal{B}}^{-1}(t)\circ\dot{x}_{\mathcal{B}}(t) - \dot{w}_{\mathcal{B}}(t)\in\mathcal{R}(X_{\mathcal{B}}A_{\mathcal{B}}^T).
\end{array}$$

Observe that

$$\dot{w}_{\mathcal{B}}(t) = -p_{\mathcal{B}}e^{-t} + \dot{h}(t)e_{\mathcal{B}} = -w_{\mathcal{B}}(t) + \frac{e^Tp}{\rho}e^{-(1-\frac{n}{\rho})t}e_{\mathcal{B}}.$$

Therefore, letting $\hat{w}_{\mathcal{B}}(t) = w_{\mathcal{B}}(t)\exp\left\{\left(1-\frac{n}{\rho}\right)t\right\}$, we obtain

$$\hat{w}_{\mathcal{B}}(t)\circ x_{\mathcal{B}}^{-1}(t)\circ\dot{x}_{\mathcal{B}}(t) + \hat{w}_{\mathcal{B}}(t) - \frac{e^Tp}{\rho}e_{\mathcal{B}}\in\mathcal{R}(X_{\mathcal{B}}A_{\mathcal{B}}^T). \tag{54}$$

From (53) it also follows that $\hat{w}_{\mathcal{B}}(t)\in\mathcal{R}(X_{\mathcal{B}}A_{\mathcal{B}}^T)$. Note also that,

$$\frac{e^Tp}{\rho}e_{\mathcal{B}} = \frac{n}{\rho\left(1-e^{-\frac{n}{\rho}t}\right)}\hat{w}_{\mathcal{B}}(t) - \frac{n\cdot e^{-\frac{n}{\rho}t}}{\rho\left(1-e^{-\frac{n}{\rho}t}\right)}p_{\mathcal{B}}.$$

Combining these observations with (54) we get

$$\hat{w}_{\mathcal{B}}(t) \circ x_{\mathcal{B}}^{-1}(t) \circ \dot{x}_{\mathcal{B}}(t) + \frac{n \cdot e^{-\frac{n}{\rho}t}}{\rho\left(1 - e^{-\frac{n}{\rho}t}\right)} p_{\mathcal{B}} \in \mathcal{R}(X_{\mathcal{B}} A_{\mathcal{B}}^T). \tag{55}$$

Note also that

$$A_{\mathcal{B}} \dot{x}_{\mathcal{B}}(t) = -A_{\mathcal{N}} \dot{x}_{\mathcal{N}}(t). \tag{56}$$

Using the notation introduced before the statement of the lemma, (55) and (56) can be rewritten as follows:

$$\tilde{x}_{\mathcal{B}}(t) + \frac{n \cdot e^{-\frac{n}{\rho}t}}{\rho\left(1 - e^{-\frac{n}{\rho}t}\right)} \tilde{p}_{\mathcal{B}}(t) \in \mathcal{R}(\tilde{A}_{\mathcal{B}}^T), \tag{57}$$

$$\tilde{A}_{\mathcal{B}}(t) \tilde{x}_{\mathcal{B}}(t) = -A_{\mathcal{N}} \dot{x}_{\mathcal{N}}(t). \tag{58}$$

Let $\tilde{A}_{\mathcal{B}}^{+}(t)$ denote the pseudo-inverse of $\tilde{A}_{\mathcal{B}}(t)$ [3]. For example, if rank$(\tilde{A}_{\mathcal{B}}(t))$=$m$, then $\tilde{A}_{\mathcal{B}}^{+}(t) = \tilde{A}_{\mathcal{B}}^T(t)\left(\tilde{A}_{\mathcal{B}}(t)\tilde{A}_{\mathcal{B}}^T(t)\right)^{-1}$. Then, $P_{\mathcal{R}(\tilde{A}_{\mathcal{B}}^T)} := \tilde{A}_{\mathcal{B}}^{+}(t)\tilde{A}_{\mathcal{B}}(t)$ is the orthogonal projection matrix onto $\mathcal{R}(\tilde{A}_{\mathcal{B}}^T)$ and $P_{\mathcal{N}(\tilde{A}_{\mathcal{B}})} := I - \tilde{A}_{\mathcal{B}}^{+}(t)\tilde{A}_{\mathcal{B}}(t)$ is the orthogonal projection matrix onto $\mathcal{N}(\tilde{A}_{\mathcal{B}})$ [3]. From (58) we obtain

$$P_{\mathcal{R}(\tilde{A}_{\mathcal{B}}^T)} \tilde{x}_{\mathcal{B}}(t) \quad = \quad \tilde{A}_{\mathcal{B}}^{+}(t)\tilde{A}_{\mathcal{B}}(t)\tilde{x}_{\mathcal{B}}(t) = -\tilde{A}_{\mathcal{B}}^{+}(t) A_{\mathcal{N}} \dot{x}_{\mathcal{N}}(t),$$

and from (57), using the fact that $\mathcal{R}(\tilde{A}_{\mathcal{B}}^T)$ and $\mathcal{N}(\tilde{A}_{\mathcal{B}})$ are orthogonal to each other, we get

$$P_{\mathcal{N}(\tilde{A}_{\mathcal{B}})} \tilde{x}_{\mathcal{B}}(t) \quad = \quad -\frac{n \cdot e^{-\frac{n}{\rho}t}}{\rho\left(1 - e^{-\frac{n}{\rho}t}\right)} \left(I - \tilde{A}_{\mathcal{B}}^{+}(t)\tilde{A}_{\mathcal{B}}(t)\right) \tilde{p}_{\mathcal{B}}(t).$$

Combining, we have

$$\begin{aligned} \tilde{x}_{\mathcal{B}}(t) &= P_{\mathcal{R}(\tilde{A}_{\mathcal{B}}^T)} \tilde{x}_{\mathcal{B}}(t) + P_{\mathcal{N}(\tilde{A}_{\mathcal{B}})} \tilde{x}_{\mathcal{B}}(t) \\ &= -\tilde{A}_{\mathcal{B}}^{+}(t) A_{\mathcal{N}} \dot{x}_{\mathcal{N}}(t) - \frac{n \cdot e^{-\frac{n}{\rho}t}}{\rho\left(1 - e^{-\frac{n}{\rho}t}\right)} \left(I - \tilde{A}_{\mathcal{B}}^{+}(t)\tilde{A}_{\mathcal{B}}(t)\right) \tilde{p}_{\mathcal{B}}(t), \end{aligned}$$

which gives (51). □

Next, we compute limits of some of the expressions that appear in equations (51) and (52):

$$\lim_{t\to\infty} \hat{w}_{\mathcal{B}}(t) = \frac{e^T p}{n} e_{\mathcal{B}}, \quad \lim_{t\to\infty} \hat{w}_{\mathcal{N}}(t) = \frac{e^T p}{n} e_{\mathcal{N}}, \tag{59}$$

$$\lim_{t\to\infty} D_{\mathcal{B}}(t) = \sqrt{\frac{e^T p}{n}} (X_{\mathcal{B}}^*)^{-1}, \quad \lim_{t\to\infty} D_{\mathcal{N}}(t) = \sqrt{\frac{e^T p}{n}} (S_{\mathcal{N}}^*)^{-1}, \tag{60}$$

$$\lim_{t\to\infty} \tilde{A}_{\mathcal{B}}(t) = \sqrt{\frac{n}{e^T p}} A_{\mathcal{B}} X_{\mathcal{B}}^*, \quad \lim_{t\to\infty} \tilde{G}_{\mathcal{N}}(t) = \sqrt{\frac{n}{e^T p}} G_{\mathcal{N}} S_{\mathcal{N}}^*, \tag{61}$$

$$\lim_{t\to\infty} \tilde{p}_{\mathcal{B}}(t) = \sqrt{\frac{n}{e^T p}} p_{\mathcal{B}}, \quad \lim_{t\to\infty} \tilde{p}_{\mathcal{N}}(t) = \sqrt{\frac{n}{e^T p}} p_{\mathcal{N}}. \tag{62}$$

Lemma 5.2

$$\lim_{t\to\infty} \tilde{A}_{\mathcal{B}}^{+}(t) = \sqrt{\frac{e^T p}{n}}\,(A_{\mathcal{B}} X_{\mathcal{B}}^{*})^{+} \tag{63}$$

$$\lim_{t\to\infty} \tilde{G}_{\mathcal{N}}^{+}(t) = \sqrt{\frac{e^T p}{n}}\,(G_{\mathcal{N}} S_{\mathcal{N}}^{*})^{+}. \tag{64}$$

Proof:
This result about the limiting properties of the pseudo-inverses is an immediate consequence of Lemma 2.3 in [4] and equations (61). ∎

Consider the normalized direction vectors $(\hat{x}, \hat{s})$ which are defined as follows:

$$\hat{x}(t) = \exp\left\{\left(1-\frac{n}{\rho}\right)t\right\}\dot{x}(t), \quad \text{and} \quad \hat{s}(t) = \exp\left\{\left(1-\frac{n}{\rho}\right)t\right\}\dot{s}(t). \tag{65}$$

From (50) it follows that

$$x(t)\circ\hat{s}(t) + \hat{x}(t)\circ s(t) = -\frac{e^T p}{n}\left(1-\frac{n}{\rho}\right)e + \exp\left\{-\frac{n}{\rho}t\right\}\left(\frac{e^T p}{n}e - p\right). \tag{66}$$

This last equation indicates that $x(t)\circ\hat{s}(t)$ and $\hat{x}(t)\circ s(t)$ remain bounded as $t\to\infty$. Since $x_{\mathcal{B}}(t)$ and $s_{\mathcal{N}}(t)$ converge to $x_{\mathcal{B}}^{*} > 0$ and $s_{\mathcal{N}}^{*} > 0$ respectively, and therefore, remain bounded away from zero, we must have that $\hat{x}_{\mathcal{N}}(t)$ and $\hat{s}_{\mathcal{B}}(t)$ remain bounded. This observation leads to the following conclusion:

Lemma 5.3 *Let $(\hat{x}(t), \hat{s}(t))$ be as in (65) and assume that $\rho \le 2n$. Then, $(\hat{x}(t), \hat{s}(t))$ remains bounded as t tends to ∞.*

Proof:
From (51) we have that

$$D_{\mathcal{B}}(t)\hat{x}_{\mathcal{B}}(t) = -\tilde{A}_{\mathcal{B}}^{+}(t)A_{\mathcal{N}}\hat{x}_{\mathcal{N}}(t) - \frac{n\cdot e^{\left(1-\frac{2n}{\rho}\right)t}}{\rho\left(1-e^{-\frac{n}{\rho}t}\right)}\left(I - \tilde{A}_{\mathcal{B}}^{+}(t)\tilde{A}_{\mathcal{B}}(t)\right)\tilde{p}_{\mathcal{B}}. \tag{67}$$

When, $\rho \le 2n$, the factor $\frac{n\cdot e^{\left(1-\frac{2n}{\rho}\right)t}}{\rho\left(1-e^{-\frac{n}{\rho}t}\right)}$ is convergent as t tends to ∞. Now, using Lemma 5.2 and the equations (61)-(62), we conclude that the second term in the right-hand-side of the equation above remains bounded. Combining this observation with the the fact that $\hat{x}_{\mathcal{N}}(t)$ remains bounded as t tends to ∞, we obtain that $D_{\mathcal{B}}(t)\hat{x}_{\mathcal{B}}(t)$ remains bounded. Using (60) we conclude that $\hat{x}_{\mathcal{B}}(t)$ is also bounded as t tends to ∞. The fact that $\hat{s}(t)$ is bounded follows similarly. ∎

Now, the following two results are easy to prove:

Theorem 5.1 *Let $(\hat{x}(t), \hat{s}(t))$ be as in (65) and assume that $\rho \le 2n$. Then, we have that $\lim_{t\to\infty}\hat{x}_{\mathcal{N}}(t)$ and $\lim_{t\to\infty}\hat{s}_{\mathcal{B}}(t)$ exist and satisfy the following equations:*

$$\lim_{t\to\infty}\hat{x}_{\mathcal{N}}(t) = -\frac{e^T p}{n}\left(1-\frac{n}{\rho}\right)(s_{\mathcal{N}}^{*})^{-1}, \tag{68}$$

$$\lim_{t\to\infty}\hat{s}_{\mathcal{B}}(t) = -\frac{e^T p}{n}\left(1-\frac{n}{\rho}\right)(x_{\mathcal{B}}^{*})^{-1}. \tag{69}$$

Proof:
From (66) we have that

$$x_{\mathcal{B}}(t) \circ \hat{s}_{\mathcal{B}}(t) + \hat{x}_{\mathcal{B}}(t) \circ s_{\mathcal{B}}(t) \quad = \quad -\frac{e^T p}{n}\left(1 - \frac{n}{\rho}\right) e_{\mathcal{B}} + \exp\left\{-\frac{n}{\rho}t\right\}\left(\frac{e^T p}{n} e_{\mathcal{B}} - p_{\mathcal{B}}\right).$$

Taking the limit on the right-hand-side as $t \to \infty$ we obtain $-\frac{e^T p}{n}\left(\frac{\rho - n}{\rho}\right) e_{\mathcal{B}}$. Since $s_{\mathcal{B}}(t) \to 0$ and $\hat{x}_{\mathcal{B}}(t)$ is bounded, we must then have that $x_{\mathcal{B}}(t) \circ \hat{s}_{\mathcal{B}}(t)$ converges to $-\frac{e^T p}{n}\left(\frac{\rho - n}{\rho}\right) e_{\mathcal{B}}$. Since $x_{\mathcal{B}}(t) \to x^*_{\mathcal{B}}$, it follows that $\lim_{t\to\infty} \hat{s}_{\mathcal{B}}(t)$ exists and satisfies (69). The corresponding result for $\hat{x}_{\mathcal{N}}(t)$ follows identically. ∎

Let

$$\xi_{\mathcal{B}} = X^*_{\mathcal{B}} \left(A_{\mathcal{B}} X^*_{\mathcal{B}}\right)^+ A_{\mathcal{N}} \left(s^*_{\mathcal{N}}\right)^{-1}, \qquad \sigma_{\mathcal{N}} = S^*_{\mathcal{N}} \left(G_{\mathcal{N}} S^*_{\mathcal{N}}\right)^+ G_{\mathcal{B}} \left(x^*_{\mathcal{B}}\right)^{-1}$$
$$\pi_{\mathcal{B}} = X^*_{\mathcal{B}} \left(I - \left(A_{\mathcal{B}} X^*_{\mathcal{B}}\right)^+ A_{\mathcal{B}} X^*_{\mathcal{B}}\right) p_{\mathcal{B}}, \qquad \pi_{\mathcal{N}} = S^*_{\mathcal{N}} \left(I - \left(G_{\mathcal{N}} S^*_{\mathcal{N}}\right)^+ G_{\mathcal{N}} S^*_{\mathcal{N}}\right) p_{\mathcal{N}}.$$

Theorem 5.2 *Let $(\hat{x}(t), \hat{s}(t))$ be as in (65) and assume that $\rho \le 2n$. Then, we have that $\lim_{t\to\infty} \hat{x}_{\mathcal{B}}(t)$ and $\lim_{t\to\infty} \hat{s}_{\mathcal{N}}(t)$ exist. When $\rho < 2n$ we have the following identities:*

$$\lim_{t\to\infty} \hat{x}_{\mathcal{B}}(t) \quad = \quad \frac{e^T p}{n}\left(1 - \frac{n}{\rho}\right) \xi_{\mathcal{B}}, \tag{70}$$
$$\lim_{t\to\infty} \hat{s}_{\mathcal{N}}(t) \quad = \quad \frac{e^T p}{n}\left(1 - \frac{n}{\rho}\right) \sigma_{\mathcal{N}}. \tag{71}$$

When $\rho = 2n$, the following equations hold:

$$\lim_{t\to\infty} \hat{x}_{\mathcal{B}}(t) \quad = \quad \frac{e^T p}{2n} \xi_{\mathcal{B}} - \frac{n}{2(e^T p)} \pi_{\mathcal{B}}, \tag{72}$$
$$\lim_{t\to\infty} \hat{s}_{\mathcal{N}}(t) \quad = \quad \frac{e^T p}{2n} \sigma_{\mathcal{N}} - \frac{n}{2(e^T p)} \pi_{\mathcal{N}}. \tag{73}$$

Proof:
Recall equation (67). When $\rho < 2n$, the second term on the right-hand-side converges to zero since $e^{\left(1 - \frac{2n}{\rho}\right)t}$ tends to zero and everything else is bounded. Thus, using (59) we have $\lim_{t\to\infty} \hat{x}_{\mathcal{B}}(t) = X^*_{\mathcal{B}} \left(A_{\mathcal{B}} X^*_{\mathcal{B}}\right)^+ A_{\mathcal{N}} \lim_{t\to\infty} \hat{x}_{\mathcal{N}}(t)$ and (70) is obtained using Theorem 5.1. Similarly, one obtains (71).

When $\rho = 2n$, the factor in front of the second term in (67) converges to the positive constant $\frac{n}{\rho} = \frac{1}{2}$. Therefore, using Theorem 5.1 and equations (59)–(62) we get (72) and (73). ∎

Limits of the normalized vectors $\left(\frac{\dot{x}(t)}{\|\dot{x}(t)\|}, \frac{\dot{s}(t)}{\|\dot{s}(t)\|}\right)$ are obtained immediately from Theorems 5.1 and 5.2:

Corollary 5.1 *Let $(x(t), s(t))$ for $t \in (l, \infty)$ denote the maximal solution of the ODE (25) for a given $(x^0, s^0) \in \mathcal{F}^0$ and assume that $\rho \le 2n$. All trajectories of this form satisfy the following equations:*

$$\lim_{t\to\infty} \frac{\dot{x}(t)}{\|\dot{x}(t)\|} = \frac{q_P}{\|q_P\|}, \qquad \lim_{t\to\infty} \frac{\dot{s}(t)}{\|\dot{s}(t)\|} = \frac{q_D}{\|q_D\|} \tag{74}$$

where

$$q_P = \begin{bmatrix} \xi_{\mathcal{B}} \\ (s^*_{\mathcal{N}})^{-1} \end{bmatrix}, \quad \textit{and} \quad q_D = \begin{bmatrix} (x^*_{\mathcal{B}})^{-1} \\ \sigma_{\mathcal{N}} \end{bmatrix}, \qquad \textit{if } \rho < 2n,$$

$$q_P = \begin{bmatrix} \xi_{\mathcal{B}} - \left(\frac{n}{e^T p}\right)^2 \pi_{\mathcal{B}} \\ (s^*_{\mathcal{N}})^{-1} \end{bmatrix}, \quad \textit{and} \quad q_D = \begin{bmatrix} (x^*_{\mathcal{B}})^{-1} - \left(\frac{n}{e^T p}\right)^2 \pi_{\mathcal{N}} \\ \sigma_{\mathcal{N}} \end{bmatrix}, \qquad \textit{if } \rho = 2n.$$

When $\rho = 2n$ the TTY potential-function $\Phi_\rho(x, y)$ is a homogeneous function and $\exp\{\Phi_\rho(x, y)\}$ is a convex function for all $\rho \geq 2n$. The value $2n$ also represents a threshold value for the convergence behavior of the KMY trajectories. When $\rho = 2n$ the direction of convergence depends on the initial point $(x^0, s^0) \in \mathcal{F}^0$ as indicated by the appearance of the $p = x^0 \circ s^0$ terms in the formulas. We note that, when $\rho < 2n$ the asymptotic direction of convergence does not depend on the initial point and is identical to that of the central path. Therefore, when $\rho < 2n$ all trajectories of the vector field given by the search direction of the Kojima-Mizuno-Yoshise's primal-dual potential-reduction algorithm converge to the analytic center of the optimal face tangentially to the central path. We show below that the asymptotic behavior of the trajectories is significantly different when $\rho > 2n$:

Theorem 5.3 *Consider $(\dot{x}(t), \dot{s}(t))$ where $(x(t), s(t))$ for $t \in (l, \infty)$ denote the maximal solution of the ODE (25) and assume that $\rho > 2n$. Define*

$$\bar{x}(t) = \exp\{\frac{n}{\rho}t\}\dot{x}(t), \quad \textit{and} \quad \bar{s}(t) = \exp\{\frac{n}{\rho}t\}\dot{s}(t). \tag{75}$$

Then, we have that $\lim_{t\to\infty} \bar{x}(t)$ and $\lim_{t\to\infty} \bar{s}(t)$ exist and satisfy the following equations:

$$\lim_{t\to\infty} \frac{\bar{x}(t)}{\|\bar{x}(t)\|} = \frac{q_P}{\|q_P\|}, \quad \textit{and} \quad \lim_{t\to\infty} \frac{\bar{s}(t)}{\|\bar{s}(t)\|} = \frac{q_D}{\|q_D\|}, \tag{76}$$

where

$$q_P = \begin{bmatrix} -\pi_{\mathcal{B}} \\ 0_{\mathcal{N}} \end{bmatrix}, \quad \textit{and} \quad q_D = \begin{bmatrix} 0_{\mathcal{B}} \\ -\pi_{\mathcal{N}} \end{bmatrix}.$$

Proof:
From (51) we have that

$$D_{\mathcal{B}}(t)\bar{x}_{\mathcal{B}}(t) \;=\; -\tilde{A}^+_{\mathcal{B}}(t)A_{\mathcal{N}}\bar{x}_{\mathcal{N}}(t) - \frac{n}{\rho\left(1 - e^{-\frac{n}{\rho}t}\right)}\left(I - \tilde{A}^+_{\mathcal{B}}(t)\tilde{A}_{\mathcal{B}}(t)\right)\tilde{p}_{\mathcal{B}}. \tag{77}$$

Note that, $\bar{x}_N(t) = e^{-\left(1-\frac{2n}{\rho}\right)t}\hat{x}_{\mathcal{N}}(t)$. Since $\hat{x}_{\mathcal{N}}(t)$ is bounded and $-\left(1 - \frac{2n}{\rho}\right) < 0$, we conclude that $\bar{x}_N(t) \to 0$. Now, the first equation in (76) follows from equations (77) and (59)–(62). The second identity in (76) is obtained similarly. ∎

This final theorem indicates that when $\rho > 2n$, the trajectories associated with the Kojima-Mizuno-Yoshise algorithm will converge to the analytic center of the optimal face tangentially to the optimal face.

References

[1] I. Adler and R. D. C. Monteiro, "Limiting behavior of the affine scaling continuous trajectories for linear programming problems," *Mathematical Programming*, 50 (1991) pp. 29–51.

[2] K. M. Anstreicher, "Potential reduction algorithms," in: T. Terlaky, ed., *Interior point methods of mathematical programming*, (Kluwer Academic Publishers, Dordrecht, Netherlands, 1996) pp. 125–158.

[3] G. H. Golub and C. F. Van Loan, *Matrix Computations,* 2nd ed., (The Johns Hopkins University Press, Baltimore, 1989).

[4] O. Güler, " Limiting behavior of weighted central paths in linear programming," *Mathematical Programming*, 65 (1994) pp. 347–363.

[5] M. W. Hirsch and S. Smale, *Differential equations, dynamical systems, and linear algebra* (Academic Press, New York, 1974).

[6] M. Iri and H. Imai, "A multiplicative barrier function method for linear programming," *Algorithmica*, 1 (1986) 455–82.

[7] M. Kojima, S. Mizuno, and A. Yoshise, "An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementarity problems," *Mathematical Programming*, 50 (1991) pp. 331–342.

[8] R. D. C. Monteiro, "Convergence and boundary behavior of the projective scaling trajectories for linear programming," *Mathematics of Operations Research*, 16 (1991) pp. 842–858.

[9] R. D. C. Monteiro, "On the continuous trajectories for potential reduction algorithms for linear programming," *Mathematics of Operations Research*, 17 (1992) pp. 225–253.

[10] Yu. Nesterov, "Long-step strategies in interior-point primal-dual methods," *Mathematical Programming*, 76 (1997) pp. 47–94.

[11] C. Roos, T. Terlaky, and J.-Ph. Vial, *Theory and Algorithms for Linear Optimization* (John Wiley and Sons, Chichester, 1997).

[12] K. Tanabe, "Centered Newton method for mathematical programming," in: M. Iri and K. Yajima, eds., *Lecture Notes in Control and Information Sciences,* (Springer-Verlag, Berlin, 1988) pp. 197–206.

[13] M. J. Todd, "Potential-reduction methods in mathematical programming," *Mathematical Programming,* 76 (1997) 3–45.

[14] M. J. Todd and Y. Ye, "A centered projective algorithm for linear programming," *Mathematics of Operations Research*, 15 (1990) 508–529.

[15] R. H. Tütüncü, "A primal-dual variant of the Iri-Imai Algorithm for Linear Programming", *Mathematics of Operations Research*, 25 (2000),195–213.

[16] R. H. Tütüncü, "Quadratic convergence of potential-reduction methods for degenerate problems," *Mathematical Programming*, 90 (2001) pp. 169–203.

[17] S. Wright, *Primal-dual interior-point methods* (SIAM, Philadelphia, 1997).