Modified Barrier Functions in Linear Programming

Roman Polyak

Department of Mathematical Sciences
IBM T.J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598 USA

March 1992

Abstract:

In this paper we developed Modified Barrier Function (MBF's) theory for nondegenerate dual pair Linear Programming (LP) problems and methods, that are based on this theory, for their simultaneous solution.

The MBF methods have up to superlinear rate of convergence and they converge linearly with a given ratio when the penalty parameter is fixed. The numerical realization of these methods leads to primal and dual Newton MBF Methods (NMBM's).

The main difference of NMBM's from the Interior Point Methods that are based on Classical Barrier Functions (CBF's), results from the difference between MBF and CBF trajectories. The NMBM's follow the MBF trajectories that along with "warm start" possess a "hot start".

Beginning at the "hot start" both primal and dual variables became and thereafter remain "well defined". So by updating the vector of dual variables instead of the penalty parameter, it is possible to decrease substantially (from $0(\sqrt{n})$ to $0(\ln n)$) the number of Newton steps required to reduce the gap between primal and dual objective functions by a given factor.

As it turns out, the crucial parameters that determines the "hot start", rate of convergence of MBF methods as well as the complexity of the NMBM's, are the condition numbers of the primal and dual LP.

We introduce and numerically characterize the condition numbers in terms of the key parameters of dual pair LP.

The main results of this paper remain true if only one of the dual pair LP has a unique solution.

The practical implementation of the Newton MBF Method, that was undertaken recently at the IBM T.J. Watson Research Center, allows us to observe the "hot start" phenomenon practically for every LP.

Keywords: Modified Barrier Functions, dual problems, non-degeneracy, condition number

1. Introduction

In this paper we introduce Modified Barrier Functions (MBF's) for primal and dual Linear Programming (LP) problems. We develop the MBF theory for nondegenerate dual pairs of LP and methods for their simultaneous solution that are based on this theory.

MBF's are Classical Lagrangians for an equivalent problem that one can obtain from the original problem by a monotone transformation of the constraints. However, in contrast to the Classical Lagrangian for the original LP problem, the MBF's have a unique extrenum under the fixed optimal Lagrange multipliers. This extrenum coincides with the LP solution and one can find it as a minimum of a smooth and strongly convex function.

MBF's have some common features with Augmented Lagrangians (see [3], [11], [15], [26], [28]), and one could consider the MBF's as Interior Augmented Lagrangian's. However, in contrast to the Augmented Lagrangian approach to LP (see [2], [22]), instead of solving at every step a subproblem with inequality constraints that remains the combinatorial nature of the original problem, the MBF's methods require solving at every step a smooth and strongly convex optimization problem with linear equality constraints or finding the minimum of a smooth and strongly convex function. It allows us to eliminate the combinatorial nature of the subproblems and at the same time not only maintain the main Augmented Lagrangian's advantages, but to gain some new important properties.

In contrast to the Classical Barrier Functions (CBF's), (see [4], [5], [8], [13]) the MBF's are smooth of any order in a neighborhood of primal and dual solutions and the condition number of their Hessians are stable when the process approaches the solution.

Moreover, being Classical Lagrangians for an equivalent problem, MBF's combine the global CBF self-concordant properties (see [20]), that guarantee the polynomial complexity of the interior point methods (see [1], [7], [12], [14] [27]), with excellent local properties, which allow us to speed up essentially the process in the second stage, and to make the process potentially more stable.

The numerical realization of the MBF methods leads to the Newton Modified Barrier Methods (NMBM's). As it turns out, for any nondegenerate dual pair LP along with the "warm start" (see [1], [7], [12], [14], [27]), there exists a "hot start" (see [23]-[25]). From this point on both the primal and dual variables become and thereafter remain "well defined". The primal variable is "well defined" in terms of S. Smale's theorem (see [29]) whereas the dual is "well defined" in terms of the basic MBF theorem (see [24]).

As a result, the number of Newton steps required to reduce the gap between primal and dual objective functions by a given factor, beginning at the "hot start", is significantly less than that required by the Interior Point Methods, that are based on CBF's.

The moment when the process reaches the "hot start" is crucial for the complexity of NMBM. The "hot start" essentially depends on the *condition numbers of the primal and dual LP* which is introduced and numerically characterized in terms of the key parameters of the dual pair LP.

Some of the results in this paper were obtained in 1986 and part of it was mentioned in [24].

2. MBF for the Primal LP Problem

In this section, we transform the primal linear program into an equivalent problem using the logarithmic barrier functions. The Classical Lagrangian associated with the equivalent problem leads to primal MBF.

The following notation and assumptions will be used throughout the paper. Let $A = (N, B)$ be an $m \times n$ matrix ($n > m$), with columns $a_i = \begin{pmatrix} a_{1i} \\ \cdots \\ a_{mi} \end{pmatrix}$, $i = 1, \dots, n$, $N = (a_i, i = 1, \dots, n-m)$ be an $m \times (n-m)$ matrix and $B = (a_i, i = n-m+1, \dots, n)$ be an $m \times m$ matrix, $p \in \mathbf{R}^n$, $q \in \mathbf{R}^m$ and suppose that there exists an optimal solution x^* for the primal problem $P(A, p, q)$

$$(1) \qquad x^* \in Argmin \{(p, x) \mid Ax = q, \quad x \geq 0\},$$

then there exists an optimal solution for the dual problem $D(A, p, q)$

$$(2) \qquad \mathrm{v}^* \in Argmax \{(q, \mathrm{v}) \mid \mathrm{v}A \leq p\}.$$

We assume that the dual pair (1) and (2) are *nondegenerate*, i.e. rank $A =$ rank $B = m$, and the complementary slackness conditions hold in strict form with the additional proviso that

$$(3) \quad u^*_i > 0 \text{ and } x^*_i = 0,\ j = 1, \dots, n-m;\ x^*_i > 0 \text{ and } u^*_i = 0,\ i = n-m+1, \dots, n,$$

where $u = p - \mathrm{v}A$ and $u^* = p - \mathrm{v}^*A$. (Under the nondegeneracy assumption the optimal solutions x^* and v^* are unique, i.e. (1) and (2) are equalities.)

Let $p_N = (p_1, \dots, p_{n-m})$, $p_B = (p_{n-m+1}, \dots, p_n)$, $u_N = (u_1, \dots, u_{n-m})$, $u_B = (u_{n-m+1}, \dots, u_n)$, so $u^*_N = p_N - \mathrm{v}^*N > 0$, $u^*_B = p_B - \mathrm{v}^*B = 0$.

Consider now a class of equivalent transformations of the problem (1) parameterized by a constant $k > 0$ and based on a monotone-increasing, twice-continuouslydifferentiable concave function $\psi(t)$ on $R^1 \to R^1$ that satisfies: $\psi(0) = 0$, $\psi'(0) = 1$, and $\psi''(0) < 0$. Clearly problem (1) is equivalent in finding

$$(4) \qquad x^* = argmin\{(p, x) \mid Ax = q,\ k^{-1}\psi(kx_i) \geq 0,\ \ j = 1, \dots, n\}.$$

Removing the linear manifold

$$Q = \{x \colon Ax = q\}$$

from the list of constraints, we define now the Lagrangian associated with (4) for the rest of the constraints:

$$L(x, u, k) = (p, x) - k^{-1} \sum_{j=1}^{n} u_i \psi(kx_i).$$

The MBF for the primal LP problem (1) we define by formula

$$F(x, u, k) = \begin{cases} L(x, u, k), & x \in \operatorname{int} \Omega_k \cap Q \\ \infty & , x \notin \operatorname{int} \Omega_k \cap Q \end{cases}$$

where $\Omega_k = \{x \in \mathbb{R}^n : x_i \geq -k^{-1}, i = 1, \dots, n\}$. We start with a few useful observations concerning the properties of this Lagrangian. Taking (3) into account and recalling that $\psi(0) = 0$, we have

$$F(x^*, u^*, k) = (p, x^*), \quad \forall\ k > 0.$$

Next, from (3) and $\psi'(0) = 1$ and taking into account $u^*_B = 0$, we obtain

$$F'_x(x, u^*, k)\big|_{x = x^*} = p - [diag\psi'(kx^*_i)]_{i=1}^{n} u^* = p - u^*,$$

where $F'_x(x, u, k)$ is the gradient of $F(x, u, k)$ with respect to $x \in \mathbb{R}^n$. Now, since $p - u^* = v^*A$ it follows that the optimal solution x^* to (4) also satisfies

$$x^* = argmin\{F(x, u^*, k) \mid x \in Q\}, \quad \forall\ k > 0.$$

Furthermore, in view of the uniqueness of x^* and the smoothness of $F(x, u, k)$, the vector

$$\tilde{x} = argmin\{F(x, \tilde{u}, k) \mid x \in Q\}$$

can be forced to lie as close to x^* as one desires simply by taking $\tilde{u}$ close enough to u^*.

These observations suggest that the solution to (1) may be found by using a rapidly converging method to minimize the smooth function $F(x, u, k)$ over Q and then recursively refining the initial approximation u to u^* based on the information obtained by minimizing $F(x, u, k)$ at x. In the first part of this paper we specify this procedure for the particular monotone transformation given by the logarithmic function $\psi(t) = \ell n(t+1)$ and for the Newton method for minimizing $F(x, u, k)$ on Q. So we have

$$F(x, u, k) = \begin{cases} (p, x) - k^{-1}\sum_{i=1}^{n} u_i \ell n(kx_i + 1), & \text{if } x \in \operatorname{int}\Omega_k \cap Q \\ \infty, & \text{if } x \notin \operatorname{int}\Omega_k \cap Q. \end{cases}$$

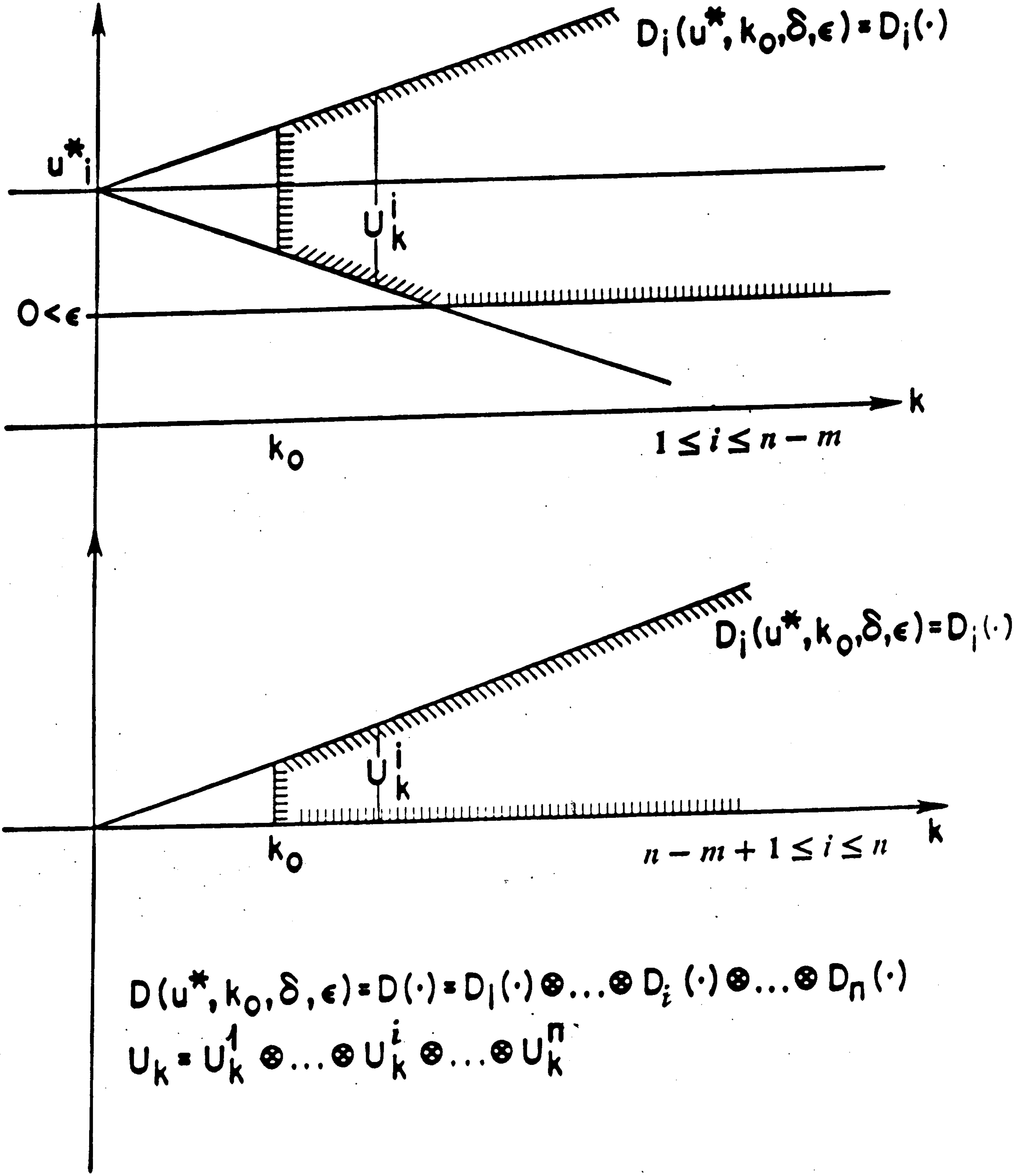
3. The Basic Theorem

In this section we present the main result on which the convergence analysis of the primal MBF method will be based. A version of this result appeared in [23] and [24] for nonlinear programming programs. The key role in the proof of the basic results in nonlinear programming play the second order optimality conditions. In case of a linear programming problem, these conditions make no sense. So to proof the corresponding results for linear programming, we have to use different techniques and arguments.

Let $\varepsilon > 0$ and $\delta > 0$ are small enough and a number $k_0 > 0$ is large enough. We will call the positive number

$$\sigma = \min\{\min\{x^*_i \mid i = n-m+1, \dots, n\}, \min\{u^*_i \mid i = 1, \dots, n-m\}\}$$

measure of the nondegeneracy of LP.



$D_i(u^*, k_0, \delta, \epsilon) = D_i(\cdot)$
u^*_i
U^i_k
$0 < \epsilon$
k
k_0
$1 \leq i \leq n-m$
$D_i(u^*, k_0, \delta, \epsilon) = D_i(\cdot)$
U^i_k
k
k_0
$n-m+1 \leq i \leq n$
$D(u^*, k_0, \delta, \epsilon) = D(\cdot) = D_1(\cdot) \otimes \ldots \otimes D_i(\cdot) \otimes \ldots \otimes D_n(\cdot)$
$U_k = U^1_k \otimes \ldots \otimes U^i_k \otimes \ldots \otimes U^n_k$

Let $D_i(\bullet) = D_i(u^*, k_0, \delta, \varepsilon) = \{u_i : u_i \geq \varepsilon, |u_i - u^*_i| \leq \delta k, k \geq k_0\}, i = 1, \dots, n-m$, $D_i(u^*, k_0, \delta, \varepsilon) = \{u_i : 0 \leq u_i \leq \delta k, k \geq k_0\}, i = n-m+1, \dots, n$, $D(u^*, k_0, \delta, \varepsilon) = D(\bullet) = D_1(\bullet) \otimes \dots \otimes D_i(\bullet) \otimes \dots \otimes D_n(\bullet)$ Also for any fixed $k \geq k_0$ define sets $U_k^i = \{u_i : \max\{\varepsilon, u^*_i - \delta k\} \leq u_i \leq u^*_i + \delta k\}, i = 1, \dots, n-m$, and $U_k^i = \{u_i;\ 0 \leq u_i \leq \delta k\}, i = n-m+1, \dots, n$. Let $U_k = U_k^1 \otimes \dots \otimes U_k^i \otimes \dots \otimes U_k^n$ then $D(\bullet) = \{(u, k) : u \in U_k, k \geq k_0\}$ (see Figure 1). Let $\|x\| = \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$. We will call a vector $u \in R^n_+$ *well defined* for a parameter $k \geq k_0$ if $u \in U_k$.

Theorem 1. Let condition (3) hold i.e. $\sigma > 0$ and rank $A =$ rank $B = m$, then there exists $k_0 > 0$ and small enough $\delta > 0$ such that for any $0 < \varepsilon < \min\{u^*_i \mid i = 1, \dots, n-m\}$ and for any $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$ the following statements hold true.

1) there exists a unique pair $\hat{x} \equiv \hat{x}(u, k) = argmin\{F(x, u, k) \mid x \in \text{int}\, \Omega_k \cap Q\}$ and $\hat{v} \equiv \hat{v}(u, k)$ such that $F'_x(\hat{x}, u, k) - \hat{v}A = 0$

2) for the triple $\hat{x} \equiv \hat{x}(u, k)$, $\hat{u} \equiv \hat{u}(u, k) = [diag(k\hat{x}_i + 1)^{-1}]_{i=1}^n u$, $\hat{v} = \hat{v}(u, k) = (p - \hat{u})A^T(AA^T)^{-1}$ the estimate

$$\max\{\|\hat{x} - x^*\|, \|\hat{u} - u^*\|, \|\hat{v} - v^*\|\} \leq ck^{-1}\|u - u^*\| \tag{5}$$

holds and $c > 0$ is independent of $k \geq k_0$,

3) for any fixed $k > 0$ we have $x^* = \hat{x}(u^*, k)$, $u^* = \hat{u}(u^*, k)$, $v^* = \hat{v}(u^*, k)$ i.e. $u^* \in \mathbb{R}^n_+$ is a fixed point of the mapping $u \to \hat{u}(u, k): \mathbb{R}^n_+ \to \mathbb{R}^n_+$

4) the restriction of $F(x, u, k)$ to Q is a strongly convex function in $x \in \mathbb{R}^{n-m}$.

Proof

1) Let $t_i = (u_i - u^*_i)k^{-1}$, $i = \overline{1, n}$, $t = (t_1, \dots, t_n)$ $S(0, \delta) = \{t = (t_1, \dots, t_n) \mid |t_i| \leq \delta, i = 1, \dots, n\}$ $\hat{u}_N = (\hat{u}_i, i = 1, \dots, n-m)$, $\hat{u}_i(x, t, k) = kt_i(kx_i + 1)^{-1}$, $i = n-m+1, \dots, n$

$\hat{u}_B(x, t, k) = (\hat{u}_i(x, t, k), i = n - m + 1, \ldots, n)$. So, from (3) we have $u^*_N > 0^{n-m}$. By definition $\hat{u}^*_B = \hat{u}_B(x^*, 0, k) = 0^m$ for any $k > 0$. Now we consider the map $\Phi(x, \hat{u}_N, v, t, k)$: $\mathbf{R}^{3n+1} \rightarrow \mathbf{R}^{2n}$

$$\Phi(x, \hat{u}_N, v, t, k) \equiv ((p - (\hat{u}_N, \hat{u}_B(x, t, k)) - vA)^T,$$
$$t_i(kx_i + 1)^{-1} + k^{-1} u^*_i(kx_i + 1)^{-1} - k^{-1}\hat{u}_i,\ i = 1, \ldots, n - m;\ -Ax + q).$$

For any $k > 0$ we have $\Phi(x^*, u^*_N, v^*, 0^n, k) = 0^{2n}$. The first step of the proof is to apply the second implicit function theorem (see [3] p. 12) to the map $\Phi(x, \hat{u}_N, v, t, k)$. So let us proceed to verify the conditions of that theorem.

Let $0^{p,q}$ the $p \times q$ matrix of zeros, I^q the $q \times q$ identity matrix , $U^*_N = [\text{diag}\, u^*_i]_{i=1}^{n-m}$; $T_N = [\text{diag}\, t_i]_{i=1}^{n-m}$, $T_B = [\text{diag}\, t_i]_{i=1}^{n-m+1}$, $\Delta_N(x, k) = [\text{diag}\,(k\,x_i + 1)]_{i=1}^{n-m}$ $\Delta_B(x, k) = [\text{diag}\,(k\,x_i + 1)]_{i=n-m+1}^{n}$, $\Delta(x, k) = [\text{diag}\,(k\,x_i + 1)]_{i=1}^{n}$.

We consider the vector functions $g_1(x, \hat{u}_N, v, t, k) = (p - (\hat{u}_N, \hat{u}_B(x, t, k)) - v\,A)^T$, $g_2(x, \hat{u}_N, v, t, k) = (t_i(k\,x_i + 1)^{-1} + k^{-1} u^*_i(k\,x_i + 1)^{-1} - k^{-1}\hat{u}_i, i = 1, \ldots, n - m)$, $g_3(x, \hat{u}_N, v, t, k) = -A\,x + q$ and their Jacobians $J_{x\,\hat{u}_N v}(g_i(\bullet)) = (J_x(g_i(\bullet))$; $J_{\hat{u}_N}(g_i(\bullet)); J_v(g_i(\bullet))) = (g'_{ix}(\bullet); g'_{i\,\hat{u}_N}(\bullet); g'_{i\,v}(\bullet)), i = 1, 2, 3$ for $x = x^*$, $\hat{u}_N = u^*_N$, $v = v^*$, $t = 0^n$ and any $k > 0$. We obtain

$$J_x(g_1(x^*, u^*_N, v^*, 0, k)) = g'_{1\,x}(\bullet) = \begin{bmatrix} 0^{n-m,n-m} & 0^{n-m,m} \\ 0^{m,n-m} & k^2\, T_B\, \Delta_B^{-2}(x, k) \end{bmatrix}_{x = x^*,\, t = 0} = 0^{n,n},$$

$$J_{\hat{u}_N}(g_1(\bullet)) = g'_{1\,\hat{u}_N}(\bullet) = \begin{bmatrix} -I^{n-m} \\ 0^{m,n-m} \end{bmatrix},\quad J_v(g_1(\bullet)) = g'_{1\,v}(\bullet) = \begin{bmatrix} -N^T \\ -B^T \end{bmatrix}.$$

$$g'_{2\,x}(\bullet) = [-k\,T_N\,\Delta_N^{-2}(x, k) - U^*_N\,\Delta_N^{-2}(x, k), 0^{n-m,m}]_{x = x^*,\, t = 0} = [-U^*_N;\ 0^{n-m,m}]$$

$$g'_{2\,\hat{u}_N}(\bullet) = -k^{-1}\, I^{n-m},\quad g_{2\,v}(\bullet) = 0^{n-m,m}$$

and

$$g'_{3x}(\bullet) = [-N, -B],\ g'_{3\hat{u}_N}(\bullet) = 0^{m,n-m},\ g_{3v}(\bullet) = 0^{m,n}.$$

Therefore for the matrix $\Phi_k \equiv \Phi'_{x\hat{u}_N v}(x^*, u^*_N, v^*, 0^n, k) : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ we have

$$\Phi_{(k)} = \begin{array}{c} \begin{array}{cccc} y_1 & y_2 & u_N & v \end{array} \\ \left[\begin{array}{cccc} \multicolumn{2}{c}{0^{n,n}} & -I^{n-m} & -N^T \\ & & 0^{m,n-m} & -B^T \\ -U^*_N & 0^{n-m,m} & -k^{-1} I^{n-m} & 0^{n-m,m} \\ -N & -B & 0^{m,n-m} & 0^{m,m} \end{array}\right]. \end{array}$$

First of all we shall prove that $\Phi_{(k)}$ is nonsingular. Let $w = (y_1, y_2, u_N, v)$ $y_1 \in R^{n-m}, y_2 = R^m, u_N \in R^{n-m}, v \in R^m$. To prove the nonsingularity $\Phi_{(k)}$ we have to show that $\Phi_{(k)} w = 0^{2n} \Rightarrow w = 0^{2n}$ for any $w \in \mathbb{R}^{2n}$. We can rewrite $\Phi_{(k)} w = 0^{2n}$ in the next form

$$u_N + N^T v = 0^{n-m}, B^T v = 0^m, U^* y_1 - k^{-1} u_N = 0^{n-m},\ N y_1 + B y_2 = 0^m.$$

Due to nonsingularity B from $B^T v = 0^m$ we obtain $v = 0^m$, then from $u_N + N^T v = 0^{n-m}$ and $v = 0^m$ we have $u_N = 0^{n-m}$. Taking into account $u^*_N = (u^*_1, \ldots, u^*_{n-m}) > 0^{n-m}$ and $u_N = 0^{n-m}$ from $U^* y_1 - k^{-1} u_N = 0^{n-m}$ we have $y_1 = 0^{n-m}$ for any $k > 0$ that together with $N y_1 + B y_2 = 0^m$ and nonsingularity B gives $y_2 = 0^m$. So for any $w \in \mathbb{R}^{2n}$ we have $\Phi_{(k)} w = 0^{2n} \Rightarrow w = 0^{2n}$, therefore $\Phi_{(k)}$ is nonsingular for any $k > 0$.

Since $\Phi(x^*, u^*_N, v^*, 0^n, k) = 0^{2n}$, and the matrix $\Phi_{(k)}$ is nonsingular for any $k \in [k_0, k_1]$ it follows from the second implicit function theorem (see [3] p. 12) that there exist $\varepsilon > 0, \delta > 0$ and smooth vector-functions $x(t,k) = (x_1(t,k), \ldots, x_n(t,k)) \equiv x(\bullet)$, $\hat{u}_N(t,k) = (\hat{u}_1(t,k), \ldots, \hat{u}_{n-m}(t,k))) = \hat{u}_N(\bullet)$, $v(t,k) = (v_1(t,k), \ldots, v_m(t,k)) = v(\bullet)$ defined uniquely in a neighborhood of $S(K, \delta) = \{(t,k) \mid |t_i| \le \delta, i = 1, \ldots, n, k \in [k_0, k_1]\}$ that

$$(6)\qquad \Phi(x(t,k), \hat{u}_N(t,k), v(t,k), t,k) \equiv \Phi(\bullet) \equiv 0^{2n} \quad \forall (t,k) \in S(K,\delta)$$

and $x(0,k) = x^*$, $\hat{u}_N(0,k) = u^*_N = (u^*_1, \dots, u^*_{n-m})$, $v(0,k) = v^*$ for any $k \in [k_0, k_1]$. Let $\hat{u}_i(t,k) = \hat{u}_i(\hat{x}(\bullet), \bullet) = k\, t_i(k\,\hat{x}(\bullet)+1)^{-1}$, $i = n-m+1, \dots, n$; $\hat{u}_B(t,k) = \hat{u}_B(\bullet) = (\hat{u}_i(\hat{x}(\bullet), \bullet), i = n-m+1, \dots, n)$, $\hat{u}(t,k) = \hat{u}(\bullet) = (\hat{u}_N(\bullet), \hat{u}_B(\bullet))$.

From (6) for any $(t,k) \in S(K,\delta)$ we have

$$(7)\qquad p^T - \begin{pmatrix} \hat{u}_N(\bullet) \\ \hat{u}_B(\bullet) \end{pmatrix} - A^T v(\bullet) = 0^n$$

$$(8)\qquad k^{-1}\hat{u}_i(\bullet) = t_i(kx_i(\bullet)+1)^{-1} + k^{-1}u^*_i(kx_i(\bullet)+1)^{-1}, \quad i = 1, \dots, n-m$$

$$(9)\qquad Ax(\bullet) = q.$$

We also define

$$(10)\qquad \hat{u}_i(\bullet) = t_i k(kx_i(\bullet)+1)^{-1}, \quad i = n-m+1, \dots, n.$$

From (8) and (10) we obtain $\hat{u}(\bullet) = \Delta^{-1}(x(\bullet), k)u$ and $F'_x(x(\bullet), u, k) = p - \hat{u}(\bullet)$. Taking into account (7) we have $F'_x(x(\bullet), u, k) = A^T v(\bullet)$. In view of (9), convexity $F(x,u,k)$ in x and $x \to \Omega_k \Rightarrow F(x,u,k) \to \infty$ we conclude that

$$x(\bullet) = \operatorname{argmin}\{F(x,u,k) \mid x \in \operatorname{int} \Omega_k \cap Q\}.$$

So we proved the first part of Theorem 1.

2) From (10) we have

$$\hat{u}_i = k\, t_i(k\, x_i(\bullet)+1)^{-1} = t_i(x_i(\bullet)+k^{-1})^{-1} = \frac{u_i - u^*_i}{k}(x_i(\bullet)+k^{-1})^{-1}, i = n-m+1, \dots, n.$$

For a small enough $\delta > 0$ and $(t,k) \in S(K,\delta)$ we have $x_i(\bullet) \geq \sigma/2$ so $\hat{u}_i(\bullet) \leq \frac{2}{k\sigma}(u_i - u^*_i)$, $i > n-m$ and $\|\hat{u}_B(\bullet) - u^*_B\| \leq 2\sigma^{-1}k^{-1}\|u_B - u^*_B\|$.

Now we show that estimate (5) holds for $x(t, k)$, $\hat{u}_N(t, k)$, $v(t, k)$. To this end we differentiate the identities (7)-(9) with respect to t.

Let $x(\bullet) = \begin{pmatrix} x_N(\bullet) \\ x_B(\bullet) \end{pmatrix}$, $x_N(\bullet) = \begin{pmatrix} x_1(\bullet) \\ \cdots \\ x_{n-m}(\bullet) \end{pmatrix}$ $x_B(\bullet) = \begin{pmatrix} x_{n-m+1}(\bullet) \\ \cdots \\ x_n(\bullet) \end{pmatrix}$,

$J_t(x_N(\bullet)) = x'_{N,t}(\bullet) : \mathbb{R}^n \to \mathbb{R}^{n-m}$, $J_t(x_B(\bullet)) = x'_{B,t}(\bullet) : \mathbb{R}^n \to \mathbb{R}^m$, $J_t(\hat{u}_N(\bullet)) = \hat{u}'_{N,t}(\bullet) : \mathbb{R}^n \to \mathbb{R}^{n-m}$, $J_t(\hat{u}_B(\bullet)) = \hat{u}'_{B,t}(\bullet) : \mathbb{R}^n \to \mathbb{R}^m$, $J_t(v(\bullet)) = v'_t(\bullet) : \mathbb{R}^n \to \mathbb{R}^m$ - Jacobians of the vector functions $x_N(\bullet)$, $x_B(\bullet)$, $\hat{u}_N(\bullet)$, $\hat{u}_B(\bullet)$ and $v(\bullet)$.

Differentiating (7) we obtain

$$\begin{bmatrix} \hat{u}'_{N,t}(\bullet) \\ \hat{u}'_{B,t}(\bullet) \end{bmatrix} - A^T v'_t(\bullet) = 0^{n,n}.$$

The last system we can rewrite as follows

$$\begin{bmatrix} \hat{u}'_{N,t}(\bullet) \\ 0^{m,n} \end{bmatrix} - A^T v'_t(\bullet) = \begin{bmatrix} 0^{n-m,n} \\ \hat{u}'_{B,t}(\bullet) \end{bmatrix}. \tag{11}$$

Differentiating (8) with respect to t we obtain

$$-[diag\,(k\,t_i + u^*_i)]_{i=1}^{n-m} \Delta_N^{-2}(\bullet) x'_{N,t}(\bullet) - k^{-1} \hat{u}'_{N,t}(\bullet) = -[\Delta_N^{-1}(\bullet), 0^{n-m,m}]. \tag{12}$$

Then differentiating (9) with respect to t we obtain

$$Ax'_t(\bullet) = Nx'_{N,t}(\bullet) + Bx'_{B,t}(\bullet) = 0. \tag{13}$$

Multiplying both sides of the systems (12) to $\Delta_N^2(\bullet)$ we obtain

$$\begin{aligned} &-[diag(kt_i + u^*_i)]_1^{n-m} x'_{N,t}(\bullet) - k^{-1} \Delta_N^2(\bullet) \bullet \hat{u}'_{N,t}(\bullet) \\ &= -[\Delta_N(\bullet),\ 0^{n-m,m}] \end{aligned} \tag{14}$$

Let

$$\Phi'_{x\hat{u}_N v}(\bullet) = \begin{bmatrix} 0^{n,n} & & I^{n-m} & -N^T \\ & & 0^{m,n-m} & -B^T \\ -[diag(kt_i + u^*_i)]_1^{n-m} & 0^{n-m,m} & -k^{-1}\Delta_N^2(\bullet) & 0^{n-m,m} \\ -N & -B & 0^{m,n-m} & 0^{m,m} \end{bmatrix}.$$

Let $\Phi'(\bullet) \equiv \Phi'_{x\hat{u}_N v(\bullet)}$, then combining (11), (13), (14) we obtain

$$(15) \qquad \begin{bmatrix} x'_t(\bullet) \\ \hat{u}'_{N,t}(\bullet) \\ v'(\bullet) \end{bmatrix} = \Phi'^{-1}(\bullet) \times \begin{bmatrix} 0^{n-m,n} \\ \hat{u}'_{B,t}(\bullet) \\ [\Delta_N(\bullet); 0^{n-m,m}] \\ 0^{m,n} \end{bmatrix} = \Phi'^{-1}(\bullet) \times R(\bullet).$$

In order to estimate the norm of the $2n \times 2n$ matrix $\Phi'^{-1}(\bullet)$ and the norm of the $2n \times n$ matrix $R(\bullet)$ we consider the $m \times n$ matrix $\hat{u}'_{B,t}(\bullet)$ and the $(n-m) \times (n-m)$ matrix $\Delta_N(\bullet)$ in more detail. For the matrix

$$\hat{u}'_{B,t}(\bullet) = J_t(\hat{u}_B(\bullet)) = \begin{bmatrix} \hat{u}'_{n-m+1,t}(\bullet) \\ \cdots \\ \hat{u}'_{n,t}(\bullet) \end{bmatrix}$$

we have

$$\hat{u}'_{B,t}(\bullet) = \left[0^{m,n-m}; k\,\Delta_B^{-1}(x(\bullet), k)\right] - k^2\, T_B\, \Delta_B^{-1}(x(\bullet), k) x'_{B,t}(\bullet).$$

Now we consider system (15) for $t = 0^n$. Taking into account $x_i(0,k) = x^*_i = 0, i = 1, \ldots, n-m$, we obtain $\hat{u}'_{B,t}(0, k) = [0^{m,n-m};\ k\,\Delta_B^{-1}(x^*, k)]$ and $\Delta_N(0, k) = I^{n-m}$ for any $k > 0$. Further we consider two matrices

$$\Phi'_{x\hat{u}_N v}(0, k) = \Phi_k = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \text{ and } R(0, k) = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$$

where

$$P_{11} = 0^{n,n}, \; P_{12} = \begin{bmatrix} -I^{n-m} & -N^T \\ 0^{m,n-m} & -B^T \end{bmatrix},$$

$$P_{21} = \begin{bmatrix} -U^*_N & 0^{n-m,m} \\ -N & -B \end{bmatrix}, \; P_{22} = \begin{bmatrix} -k^{-1} I^{n-m} & 0^{n-m,m} \\ 0^{m,n-m} & 0^{m,m} \end{bmatrix},$$

$$R_1 = \begin{bmatrix} 0^{n-m,n-m} & 0^{n-m,m} \\ 0^{m,n-m} & k\,\Delta_B^{-1}(x^*, k) \end{bmatrix} \; R_2 = \begin{bmatrix} I^{n-m} & 0^{n-m,m} \\ 0^{m,n-m} & 0^{m,m} \end{bmatrix}.$$

From (15) we obtain

$$\begin{bmatrix} x'_t(0,k) \\ \hat{u}'_{N,t}(0,k) \\ v'_t(0,k) \end{bmatrix} = \Phi_k^{-1} \bullet R(0,k).$$

To estimate the norms $\|x'_t(0,k)\|$, $\|\hat{u}'_{N,t}(0,k)\|$ and $\|v'_t(0,k)\|$ we have to estimate the norm of the matrix

$$\Pi(A,p,q,k) = \Phi_k^{-1} R(0,k) : \mathbb{R}^n \to \mathbb{R}^{2n}.$$

First of all note that P_{12}^{-1} and P_{21}^{-1} exist and

$$P_{12}^{-1} = \begin{bmatrix} -I^{n-m} & N^T B^{-T} \\ 0^{m,n-m} & -B^{-T} \end{bmatrix}, \; P_{21}^{-1} = \begin{bmatrix} -U^{*-1}_N & 0^{n-m,m} \\ B^{-1} N U^{*-1} & -B^{-1} \end{bmatrix}.$$

The existence of P_{12}^{-1} and P_{21}^{-1} guarantee the existence of the inverse matrix

$$P^{-1} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}^{-1}.$$

Moreover, we have

(16) $$P^{-1} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} = Q$$

where $Q_{11} = -P_{21} P_{22} (P_{12} - P_{11} P_{21}^{-1} P_{22})^{-1}$, $Q_{22} = (P_{11} P_{21}^{-1} P_{22} - P_{12})^{-1} P_{11} P_{21}^{-1}$ $Q_{12} = P_{21}^{-1} (I^n - P_{22} (P_{11} P_{21}^{-1} P_{22} - P_{12})^{-1} P_{11} P_{21}^{-1})$, $Q_{21} = (P_{12} - P_{11} P_{21}^{-1} P_{22})^{-1}$.

Formula (16) can be verified directly : $P\,Q = I^{2n}$. Then

$$Q_{11} = k^{-1} \begin{bmatrix} U^{*-1}_N & -U^{*-1}_N N^T B^{-T} \\ -B^{-1} N U^{*-1}_N & B^{-1} N U^{*-1}_N N^T B^{-T} \end{bmatrix}, Q_{12} = P_{21}^{-1} = \begin{bmatrix} -U^{*-1}_N & 0^{n-m,m} \\ B^{-1} N U^{*-1}_N & -B^{-1} \end{bmatrix}$$

$$Q_{21} = P_{12}^{-1} = \begin{bmatrix} -I^{n-m} & N^T B^{-T} \\ 0^{m,n-m} & -B^{-T} \end{bmatrix}, \quad Q_{22} = O^{n,n}.$$

Further

$$\Pi\,(A, p, q, k) = \begin{bmatrix} Q_{11} R_1 + Q_{12} R_2 \\ Q_{21} R_1 + Q_{22} R_2 \end{bmatrix} = \begin{bmatrix} \Pi_1\,(A, p, q, k) \\ \Pi_2\,(A, p, q, k) \end{bmatrix}$$

and for the norm $\|\Pi(A, p, q, k)\|$, which subordinate with the vector norm $\|x\|_\infty = \max_{1 \le i \le n} |x_i|$, we obtain

$$\|\Pi(A, p, q, k)\| = \max\{\|\Pi_1(A, p, q, k)\|, \|\Pi_2(A, p, q, k)\|\}.$$

Then

$$\Pi_1(A, p, q, k) = \begin{bmatrix} -U^{*-1}_N & -k^{-1} U^{*-1}_N N^T B^{-T} (k\,\Delta_B^{-1} (x^*, k)) \\ B^{-1} N U^{*-1}_N & k^{-1} B^{-1} N U^{*-1}_N B^{-T} (k\,\Delta_B^{-1} (x^*, k)) \end{bmatrix}$$

$$\Pi_2(A, p, q, k) = \begin{bmatrix} 0^{n-m,n-m} & N^T B^{-T} (k\,\Delta_B^{-1} (x^*, k)) \\ 0^{m,n-m} & -B^{-T} (k\,\Delta_B^{-1} (x^*, k)) \end{bmatrix}.$$

Further

$$\|\Pi_1(A,p,q,k)\| \le \max\{\|U^{*-1}_N\| + \|k^{-1} U^{*-1}_N N^T B^{-T}(k\,\Delta_B^{-1}(x^*,k))\|,$$
$$\|B^{-1} N U^{*-1}_N\| + \|k^{-1} B^{-1} N U^{*-1}_N B^{-T}(k\Delta_B^{-1}(x^*,k))\|\}$$

and

$$\|\Pi_2(A,p,q,k)\| \le \max\{\|N^T B^{-T}(k\,\Delta_B^{-1}(x^*,k))\|, \|B^{-T}(k\,\Delta_B^{-1}(x^*,k))\|\}.$$

Taking into account $\min_{1\le i\le n-m} u^*_i \ge \sigma > 0$ and $\min_{n-m+1\le i\le n} x^*_i \ge \sigma > 0$ we obtain $\|U^{*-1}_N\| \le \sigma^{-1}$ and $\|k\,\Delta_B^{-1}(x^*,k)\| = \|[\mathrm{diag}\,(x^*_i + k^{-1})^{-1}]_{i=n-m+1}^n\| \le \sigma^{-1}$. Hence

$$\|\Pi_1(A,p,q,k)\| \le \sigma^{-1}\max\{1 + k^{-1}\sigma^{-1}\|N^T B^{-T}\|, \|B^{-1}N\|(1 + k^{-1}\sigma^{-1}\|B^{-T}\|)\}.$$

So for any $k \ge k_0 = \sigma^{-1}\max\{\|B^{-T}\|, \|N^T B^{-T}\|\}$ we have

$$\|\Pi_1(A,p,q,k)\| \le 2\sigma^{-1}\max\{1, \|B^{-1}N\|\}.$$

Also

$$\|\Pi_2(A,p,q,k)\| \le \sigma^{-1}\max\{\|N^T B^{-T}\|, \|B^{-T}\|\}.$$

Hence, there exists an independent on $k \ge k_0 = \sigma^{-1}\max\{\|B^{-T}\|, \|N^T B^{-T}\|\}$ estimate for the norm $\|\Phi_k^{-1} R(0,k)\| = \|\Pi(A,p,q,k)\|$, i.e. the next inequality

$$\|\Phi_k^{-1} R(0,k)\| \le 2\,\sigma^{-1}\max\{1, \|B^{-1}N\|, 0.5\|B^{-T}\|, 0.5\|N^T B^{-T}\|\} = c_1.$$

holds true for any $k \ge k_0$.

Taking into account (15) we obtain for small enough $\delta > 0$ and all $(t,k) \in S(K,\delta)$ that $\|\Phi'^{-1}_{x\hat{u}_N v}(x(\tau t,k), \hat{u}_N(\tau t,k), v(\tau t,u); \tau t, k)\mathbf{R}(x(\tau t,k);\ \tau t,k)\| = \|\Phi'^{-1}(\bullet)R(\bullet)\| \le 2\,c_1$ for $\forall \tau \in [0,1]$. Hence

$$\begin{bmatrix} x(t,k) - x(0,k) \\ \hat{u}_{N,t}(t,k) - \hat{u}_{N,t}(0,k) \\ v_t(t,k) - v_t(0,k) \end{bmatrix} = \begin{bmatrix} x(t,k) - x^* \\ \hat{u}_{N,t}(t,k) - u^*_N \\ v_t(t,k) - v^* \end{bmatrix} = \int_0^1 \Phi'^{-1}(\bullet)R(\bullet)[t]\,d\tau \le 2c_1\|t\|.$$

therefore

$$\max\{\|x(t,k) - x^*\|, \hat{u}_N(t,k) - u^*_N\|, \|v(t,k) - v^*)\|\} \le 2c_1\|t\| = 2c_1\|u - u^*\|.$$

Let $\hat{x}(u,k) = x(\frac{u-u^*}{k}, k)$, $\hat{u}(u,k) = (\hat{u}_N(\frac{u-u^*}{k}, k), \hat{u}_B(\frac{u-u^*}{k}, k))^T$, $\hat{v}(u,k) = v(\frac{u-u^*}{k}, k)$ and $c = 2\max\{c_1, \sigma\} = 2c_1$, then for any $(u,k) \in D(u^*, k_0, \delta, \varepsilon)$ we obtain

$$\max\{\|\hat{x}(u,k) - x^*\|, \|\hat{u}(u,k) - u^*\|, \|\hat{v}(u,k) - v^*\|\} \le ck^{-1}\|u - u^*\|.$$

i.e. estimate (5) takes place.

3) Statement 3) is a direct consequence of estimate (5).

4) To prove the fourth part of theorem 1 we partition the vector x into two vectors $x_N \in R^{n-m}$ and $x_B \in R^m$. For any vector $(u,k) \in D(u^*, u_0, \delta, \varepsilon)$ we have $\min_{1 \le j \le n-m} u_i \ge \varepsilon > 0$. To find the restriction of $F(x,u,k) = (p,x) - \sum_{i=1}^{n} u_i \ln(kx_i + 1) = (p,x) - k^{-1}\sum_{i=1}^{n-m} u_i \ln(kx_i + 1) - k^{-1}\sum_{i=n-m+1}^{n} u_i \ln(kx_i + 1)$ to Q we substitute the vector $x_B = -B^{-1}(q - Nx_N)$ into $F(x,u,k)$. Then

$$\varphi(x_N, u, k) = (p_N, x_N) - (p_B, B^{-1}(q - Nx_N)) - k^{-1}\sum_{i=1}^{n-m} u_i \ln(kx_i + 1) -$$

$$k^{-1}\sum_{i=n-m+1}^{n} u_i \ln(k(B^{-1}(q - Nx_N))_i + 1).$$

The strong convexity of $\varphi(x_N, u, k)$ in $\mathbf{R}^{n-m}$ is furnished in the neighborhood of $\hat{x}(u,k)$ for any $(u,k) \in D(u^*, k_0, \delta, \varepsilon)$ due to the correspondent properties of terms $-k^{-1}\sum_{i=1}^{n-m} u_i\, ln(k\,x_i + 1)$ since $\min_{1 \le i \le n-m} u_i \ge \varepsilon > 0$. We complete the proof of theorem 1.

Remark 1.

All facts which have been stated in the theorem 1 for the Modified Frisch function $F(x, u, k)$ hold true for the modified Carroll's function

$$C(x, u, k) = \begin{cases} (p, x) + k^{-1} \sum_{i=1}^{n} u_i((kx_i + 1)^{-1} - 1), & \text{if } x \in \operatorname{int} \Omega_k \cap Q \\ \infty, & \text{if } x \notin \operatorname{int} \Omega_k \cap Q \end{cases}.$$

Remark 2.

All results of the basic theorem remain true if instead of $F(x, u, k)$ one will consider the next MBF

$$F(x, u, K) = \begin{cases} (p, x) - \sum_{i=1}^{n} k_i^{-1} u_i \ln(k_i x_i + 1), & \text{if } x \in \operatorname{int} \Omega_k \cap Q \\ \infty, & \text{if } x \notin \operatorname{int} \Omega_k \cap Q \end{cases}.$$

with different penalty parameters $k_i, i = 1, \dots, n$ for different inequalities $x_i \geq 0, i = 1, \dots, n$, $K = (k_1, \dots, k_n)$, $\Omega_k = \{x = (x_1, \dots, x_n) : x_i \geq -k_i^{-1}, i = 1, \dots, n\}$ and $\min_{1 \leq i \leq n} k_i \geq k_0 > 0$ where $k_0 > 0$ is large enough. The same is true for the next version of MBF that corresponds to Carroll's function [4]

$$C(x, u, K) = \begin{cases} (p, x) + \sum_{i=1}^{n} k_i^{-1} u_i((k_i x_i + 1)^{-1} - 1), & \text{if } x \in \operatorname{int} \Omega_k \cap Q \\ \infty, & \text{if } x \notin \operatorname{int} \Omega_k \cap Q. \end{cases}$$

4. One-Parameter Shifted Barrier Function for the Primal LP

To use theorem 1 for solving dual pair of linear programming problems we should first find $(u,k) \in D(u^*, k_0, \varepsilon, \delta)$.

One can compute such an approximation by using the one-parameter Shifted Barrier Function

$$M(x,k) = F(x, e, k) = \begin{cases} (p,x) - k^{-1} \sum_{i=1}^{n} \ln(kx_i + 1), & \text{if } x \in \operatorname{int} \Omega_k \cap Q \\ \infty, & \text{if } x \overline{\in} \operatorname{int} \Omega_k \cap Q \end{cases}.$$

If $\Omega = \{x : Ax = q, x_i \geq 0, i = 1, \dots, n\}$ is a polytope, then $\Omega_k \cap Q$ is a polytope too for any $k > 0$.

The next theorem holds true.

Theorem 2. Let Ω be a polytope, then

1) for any $k > 0$ there exists unique vector

$$x(k) = argmin \{M(x,k)/x \in Q\}$$

and a vector $\mathrm{v}(k)$ such that

(20) $$M'_x(x(k), k) - \mathrm{v}(k)A = 0^n$$

2) if condition (3) is held and rank $A =$ rank $B = m$ i.e. x^* and v^* are unique solutions to problems (1) and (2), then there exists $k_0 > 0$ such that for the triple $x(k)$, $u(k) = [diag(kx_i(k) + 1)^{-1}]_{i=1}^{n} e$ and $\mathrm{v}(k) = (p - u(k))A^T(AA^T)^{-1}$ the estimate

(21) $$\max \{\|x(k) - u^*\|, \|u(k) - u^*\|, \|\mathrm{v}(k) - \mathrm{v}^*\|\} \leq ck^{-1}$$

holds and the constant $c > 0$ is independent of $k \geq k_0$.

Proof. 1) The function $M(x, k)$ is strongly convex and continuous on int $\Omega_k \cap Q$. Taking into account that Ω is a polytope and $M(x,k) \to \infty$ as $x \to \partial\Omega_k$ we arrive at the conclusion that there exists a vector $x(k) \in \operatorname{int} \Omega_k \cap Q$, which satisfies (20). Along with $x(k)$ minimization of $M(x, k)$ over Q yield vectors $u(k)$ and $v(k)$ such that

$$M'_x(x(k), k) - v(k)A = p - u(k) - v(k)A = 0^n$$

where $u(k) = (u_i(k)) = ((kx_i(k) + 1)^{-1}, i = 1, \dots, n) = [diag(kx_i(k) + 1)^{-1}]_{i=1}^n e$, $e = (1, \dots, 1) \in \mathbf{R}^n$.

2) Taking into account condition (3) we have unique x^* and v^*, and if $k \to \infty$ then $x(k) \to x^*$, $u(k) \to u^*$, $v(k) = (p - u(k))A^T(AA^T)^{-1} \to v^*$ (see [5]).

Now we are going to estimate $\|\Delta x\| = \|x(k) - x^*\|$, $\|\Delta u\| = \|u(k) - u^*\|$, and $\|\Delta v\| = \|v(k) - v^*\|$.

We recall that the triple $(x(k), u(k), v(k))$ satisfies the following system.

$$p - u(k) - v(k)A = 0^n, \quad Ax(k) = q.$$

Since $x(k)$ converges to x^* and the primal problem is nondegenerate it follows that for a sufficiently large k_0 and any $k \geq k_0$ we have $x_i(k) \geq \sigma/2 > 0, i = n - m + 1, \dots, n$. Therefore $u_i(k) = 0(k^{-1})$, $i = n - m + 1, \dots, n$. Then $u^*_i = 0, i = n - m + 1, \dots, n$, so for the vector $\Delta u_B = (u_i(k) - u^*_i, i = n - m + 1, \dots, n)$ we have $\|\Delta u_B\| = 2\sigma^{-1}k^{-1}$. It follows from $p - u(k) - v(k)A = 0^n$ that for the vectors $p_N = (p_1, \dots, p_{n-m})$, $p_B = (p_{n-m+1}, \dots, p_n)$; $u_N(k) = (u_1(k), \dots, u_{n-m}(k))$, $u_B(k) = (u_{n-m+1}(k), \dots, u_n(k))$, $x_N(k) = (x_i(k), i = 1, \dots, n - m)$, $x_B(k) = (x_i(k), i = n - m + 1, \dots, n)$, $x(k) = (x_N(k), x_B(k))$ we have

$$p_N - u_N(k) - v(k)N = 0^{n-m}, \; p_B - u_B(k) - v(k)B = 0^m, \; Ax(k) = q.$$

For the primal x^* and dual v^* solutions we obtain

$$p_N - u^*_N - v^*N = 0^{n-m}, \; p_B - u^*_B - v^*B = 0^m, \; Ax^* = q.$$

Therefore, for $\Delta x = x(k) - x^*$, $\Delta u_N = u_N(k) - u^*_N$, $\Delta u_B = u_B(k) - u^*_B$, $\Delta v = v(k) - v^*$ we obtain

$$\Delta u_N = \Delta v N,\ \Delta u_B = \Delta v B,\ A\Delta x = 0^m.$$

From $u_i(k) = (kx_i(k) + 1)^{-1}$, $i = 1, \dots, n-m$ we have $x_i(k) = k^{-1}(1 - u_i(k))(u_i(k))^{-1}$.

So for large enough k_0 and any $k \geq k_0$ we have $|x_i(k)| \leq 2k^{-1}\,|1 - u^*_i|\,u^{*-1}_i$, $i = 1, \dots, n-m$. Taking into account the strictly complementary slackness condition (3), we have $\min\{u^*_i \mid i = 1, \dots, n-m\} = \sigma > 0$ and $x^*_N = 0^{n-m}$. Let $\max\{\,|1 - u^*_i|$ $\mid i = 1, \dots, n-m\} = \beta$ then for $k \geq k_0$ we obtain

$$\|\Delta x_N(k)\| = \|x_N(k) - x^*_N\| \leq 2\beta\sigma^{-1}k^{-1}.$$

From $N\Delta x_N + B\Delta x_B = 0^m$ we have $\Delta x_B = -B^{-1}N\Delta x_N$ and $\|\Delta x_B\| \leq \|B^{-1}\|\,\|N\|\,\|\Delta x_N\|$ $\leq 2\sigma^{-1}\beta k^{-1}\|B^{-1}\|\,\|N\|$. So for $c_0 = 2\sigma^{-1}\beta\max\{1, \|B^{-1}\| \bullet \|N\|\} > 0$, which is independent of $k \geq k_0$, we have $\|\Delta x\| \leq c_0 k^{-1}$.

Taking into account that $\|\Delta u_B\| \leq 2\sigma^{-1}k^{-1}$ we obtain

$$\|\Delta v\| = \|B^{-1}\Delta u_B\| \leq \|B^{-1}\|\,\|\Delta u_B\| \leq 2\,\sigma^{-1} \bullet k^{-1}\|B^{-1}\|$$

and

$$\|\Delta u_N\| \leq \|N\|\,\|\Delta v\| \leq 2\,\sigma^{-1}k^{-1}\|N\|\,\|B^{-1}\|,$$

so for $c = 2\max\{c_0,\ \sigma^{-1}, \sigma^{-1}\|B^{-1}\|,\ \sigma^{-1}\|B^{-1}\|\,\|N\|\}$, that is independent on $k \geq k_0$, we have

$$\max\{\|x(k) - x^*\|,\ \|u(k) - u^*\|,\ \|v(k) - v^*\|\} \leq ck^{-1}.$$

Theorem 2 has been proved.

Remark 3.

All facts of theorem 2 hold true for the Carroll's Shifted Barrier function

$$N(x,k)=\begin{cases}(p,x)+k^{-1}\sum_{i=1}^{n}((kx_i+1)^{-1}-1), & x\in \operatorname{int}\Omega_k\cap Q\\ \infty &, x\notin \operatorname{int}\Omega_k\cap Q.\end{cases}$$

Remark 4.

All results of Theorem 2 remain true if instead of One-Parameter Shifted Barrier Function $M(x, k)$, one considers the next Multi-Parameter Shifted Barrier Function

$$M(x,K)=\begin{cases}(p,x)-\sum k_i^{-1}\, ln(k_ix_i+1), & \text{if} x\in \operatorname{int}\Omega_k\cap Q\\ \infty &, \text{if} x\notin \operatorname{int}\Omega_k\cap Q.\end{cases}$$

or

$$N(x,K)=\begin{cases}(p,x)+\sum k_i^{-1}((k_ix_i+1)^{-1}-1), & x\in \operatorname{int}\Omega_k\cap Q\\ \infty &, x\notin \operatorname{int}\Omega_k\cap Q.\end{cases}$$

where $\Omega_k=\{x=(x_1,\dots,x_n):x_i\geq -k_i^{-1}, i=1,\dots,n\}$, $\min_{1\leq i\leq n} k_i\geq k_0>0$ and $k_0>0$ is large enough.

5. Primal MBF Methods

In this section we consider the primal MBF methods. Their convergence as well as the rate of convergence are based on the primal MBF properties.

First we consider the permanent parameter version of the Primal MBF method.

Let $0 < \gamma < 0.5$ and $k \geq k_0$ are fixed, $x^0 \in \operatorname{int} \Omega_k \cap Q$ and $u^0 = e = (1, \dots, 1) \in \mathbb{R}^n$.

The permanent parameter version of the Primal MBF method consists of finding the sequence $\{w^s = (x^s, u^s, v^s)\}_{s=0}^{\infty}$ by formulas

$$(22) \qquad \begin{cases} x^{s+1} = argmin\{F(x, u^s, k) / x \in Q\} \\ u^{s+1} = [diag(kx_i^{s+1} + 1)^{-1}]_{i=1}^{n} u^s, \ v^{s+1} = (p - u^{s+1})A^T(AA^T)^{-1}. \end{cases}$$

The next assertion is a consequence of the theorems 1 and 2.

<u>Assertion 1</u>. If the dual pair linear programming problems (1)-(2) are nondegenerate, then for a given $0 < \gamma < 0.5$ there exists such $k_0 > 0$, that for the sequence $\{w^s = (x^s, u^s, v^s)\}_{s=0}^{\infty}$ given by (22), the next estimate

$$(23) \qquad \max \{\|x^s - x^*\|, \|u^s - u^*\|, \|v^s - v^*\|\} \leq (c k^{-1})^s \|e - u^*\| \leq \gamma^s$$

holds true for any $k \geq k_0$ and $c > 0$ is independent on k.

So the sequence $\{(x^s, v^s)\}_{s=0}^{\infty}$ converges to the primal x^* and dual v^* solution with linear rate of convergence. By increasing the parameter k one can make the ratio $0 < \gamma < 0.5$ as small as one wants.

The next version of the primal MBF method provides the possibility of changing the parameter from step to step. Let $\{k_s > 0\}_{s=0}^{\infty} : k_s < k_{s+1}, k_s \to \infty, \gamma_s = c\, k_s^{-1}$, $u^0 = e = (1, \dots, 1) \in \mathbb{R}^n, x^0 \in \operatorname{int} \Omega_k \cap Q$.

The varying parameter version of the Primal MBF Method consists of finding the sequence $\{w^s = (x^s, u^s, v^s)\}_{s=0}^{\infty}$ by formulas

$$(24) \qquad \begin{cases} x^{s+1} = \operatorname{argmin} [F(x, u^s, k_s) \mid x \in Q\} \\ u^{s+1} = [\operatorname{diag} (k_s x_i^{s+1} + 1)^{-1}]_{i=1}^{n} u^s ; v^{s+1} = (p - u^{s+1})A^T(A A^T)^{-1} \end{cases}$$

The next assertion is also a consequence of theorems 1 and 2.

Assertion 2. If the dual pair linear programming problems (1) - (2) are nondegenerate and $k_s \to \infty$, then for the sequence $\{(x^s, u^s, v^s)\}_{s=0}^{\infty}$ defined by (24), the next estimate

$$\max \{\|x^s - x^*\|, \|u^s - u^*\|, \|v^s - v^*\|\} \le \gamma_0 \gamma_1 \dots \gamma_s \|e - u^*\|, \tag{25}$$

holds true and $\gamma_s \to 0$.

This means that by increasing the parameter $k \ge k_0$ from step to step, one can obtain a method for simultaneous solution primal and dual LP with superlinear rate of convergence.

To realize method (22) numerically we have to avoid solving the problem

$$\hat{x} = argmin \{F(x, u, k)/x \in Q\}$$

at every step and at the same time maintain estimate (23).

Let $\mu > 0$ and $k > 0$ is large enough and $\tilde{x}^\circ \in \Omega_k \cap Q$, $\tilde{u}^\circ = e = (1, \dots, 1) \in \mathbf{R}^n$. The sequence $\{\tilde{x}^s, \tilde{u}^s, \tilde{v}^s\}_{s=0}^{\infty}$ we are finding by formulas

$$\begin{cases} \tilde{x}^{s+1} : \|F'_x(\tilde{x}^{s+1}, \tilde{u}^s, k)\| \le \mu k^{-1} \|[diag(k\tilde{x}_i^{s+1} + 1)^{-1}]_{i=1}^n \tilde{u}^s - \tilde{u}^s\| \\ \tilde{u}^{s+1} = [diag(k\tilde{x}_i^{s+1} + 1)^{-1}]_{i=1}^n \tilde{u}^s ; \tilde{v}^{s+1} = (p - \tilde{u}^{s+1}) A^T (AA^T)^{-1}. \end{cases} \tag{26}$$

The next assertion could be proven similar to Theorem 5 [21] (see also Lemma 2 from [24]).

Assertion 3. If the dual pair LP problems (1)-(2) is nondegenerate, then there exist such $k_0 > 0$ and independent of $k \ge k_0$, constant $c > 0$, that for the sequence $(\tilde{x}^s, \tilde{u}^s, \tilde{v}^s)\}_{s=0}^{\infty}$, the estimate

(27) $$\max\{\|\tilde{x}^{s+1} - x^*\|, \|\tilde{u}^{s+1} - u^*\|, \|\tilde{v}^{s+1} - v^*\|\} \le c(1+\mu)k^{-1}\|\tilde{u}^s - u^*\|$$

holds true for any $k \ge k_0$.

Method (26) enables us to avoid the infinite procedure at every step. But inequality (26) does not give any answer how many Newton steps one has to perform in order to find $\tilde{x}^{s+1}$ and to remain estimate (27).

In the next section we consider the Newton Primal MBF method for solving dual pair LP problems. This method enables us to estimate the number of Newton method steps, which are sufficient in maintaining estimate (27).

The principal difference between the Newton Primal MBF Method and the Projected Newton Method [1], [6], [9], [10], [14] (see also [12], [18], [19], [27], [30], [31] and bibliography in it), consists of using MBF instead of CBF. Instead of changing at every step the penalty parameter $k > 0$ in the Interior Point Methods, that are based on CBF, we update the Lagrange multipliers (residuals), keeping fixed from some point the parameter $k \ge k_0$.

The function $F(x, u, k)$, as well as its derivatives, exist at the primal solution and the condition number of the MBF Hessians on Q are stable in the neighborhood of the (x^*, u^*) under any fixed $k \ge k_0$.

Both properties mentioned above allow us to get a better rate of convergence, for any nondegenerate dual pair LP, with less computational complexity at every step, when compared with Interior Point Methods, that are based on CBF.

6. Primal Newton MBF Method

In this section we will describe the Primal Newton MBF Method for simultaneous solution of primal and dual LP. Along with using the Newton Method for minimizing $F(x, u, k)$ in x and updating the vector $u = (u_1, \dots, u_n)$, we adopt the penalty parameter $k > 0$ at the proper level ($k \geq k_0$) to guarantee estimate (23).

First of all we consider one Newton step for minimizing $F(x, u, k)$ on Q. Let $x \in \text{int}\, \Omega_k \cap Q$ and suppose $(u,k) \in D(u^*, k_0, \delta, \varepsilon)$ are fixed. To find the Newton's direction for minimizing $F(x, u, k)$ on Q under the fixed vector $u > 0$ and fixed scalar $k > 0$, one has to solve the next problem

$$(28) \qquad \zeta(x, u, k) = argmin\{ \frac{1}{2} (F''_{xx}(x, u, k)\, \zeta, \zeta) + (F'_x(x, u, k), \zeta) / A\zeta = 0 \}.$$

Define

$$U = [diag\, u_i]_{i=1}^n : \mathbf{R}^n \to \mathbf{R}^n, \ \Delta(x,k) = [diag(kx_i + 1)]_{i=1}^n : \mathbf{R}^n \to \mathbf{R}^n.$$

Then $F'_x(x, u, k) = F'_x(\bullet) = p - U\Delta^{-1}(x, k)e = p - U\Delta^{-1}(x,k)e$ and $F''_{xx}(x, u, k) = F''_{xx}(\bullet) = kU\Delta^{-2}(x,k)$. So we rewrite problem (28) as follows

$$(29) \ \zeta(x, u, k) = \zeta(\bullet) = argmin\{ \frac{1}{2} k(U\Delta^{-2}(x,k)\zeta, \zeta) + ((p - U\Delta^{-1}(x,k)e), \zeta) / A\zeta = 0 \}.$$

Having introduced the Lagrange multiplier vector, $v = (v_1, \dots, v_m)$, that corresponds to the system $A\zeta = 0$, we will obtain the system

$$(30) \qquad \begin{cases} kU\Delta^{-2}(x,k)\zeta + p - U\Delta^{-1}(x,k)e = vA \\ A\zeta = 0 \end{cases}$$

which is equivalent to problem (28). Let $\Delta(x,k)r = kU\zeta$, then $\zeta = k^{-1}\Delta(x,k)U^{-1}r$. We can rewrite system (28) as

$$(31)\quad \begin{cases} \Delta^{-1}(x,k)r = A^T v - p + U\Delta^{-1}(x,k)e \\ AU^{-1}\Delta(x,k)r = 0 \end{cases}$$

or

$$(32)\quad \begin{cases} r = \Delta(x,k)A^T v - (\Delta(x,k)p - Ue) \\ AU^{-1}\Delta(x,k)r = 0. \end{cases}$$

Then system (32) can be rewritten as follows,

$$(33)\quad \begin{cases} U^{-\frac{1}{2}} r = U^{-\frac{1}{2}}\Delta(x,k)A^T v - U^{-\frac{1}{2}}(\Delta(x,k)p - Ue) \\ AU^{-\frac{1}{2}}\Delta(x,k)U^{-\frac{1}{2}} r = 0. \end{cases}$$

By setting $U^{-\frac{1}{2}} r = h$, from (33) we obtain

$$(34)\quad \begin{cases} h = U^{-\frac{1}{2}}\Delta(x,k)A^T v - U^{-\frac{1}{2}}(\Delta(x,k)p - Ue) \\ AU^{-\frac{1}{2}}\Delta(x,k)h = 0 \end{cases}$$

or

$$(35)\quad \begin{pmatrix} I^{n,n} & -\Delta(x,k)U^{-\frac{1}{2}}A^T \\ AU^{-\frac{1}{2}}\Delta(x,k) & 0 \end{pmatrix} \begin{pmatrix} h \\ v \end{pmatrix} = \begin{pmatrix} U^{-\frac{1}{2}}(Ue - \Delta(x,k)p) \\ 0 \end{pmatrix}.$$

The least square problem (35) can be rewritten as

$$(36)\quad v(x, u, k) = v(\bullet) = argmin\{\|(A\Delta(x,k)U^{-\frac{1}{2}})^T v - U^{-\frac{1}{2}}(\Delta(x,k)p - Ue)\|_2^2 / v \in R^m\}.$$

The least square problem (36) is equivalent to the next normal system of equations

$$(A\Delta^2(x,k)U^{-1}A^T)v = A\Delta(x,k)(U^{-1}\Delta(x, k)p - e).$$

To describe the Primal Newton MBF Method we define the indicator function $v(w) : \Omega_k \times \mathbf{R}^m \to R^1_+$ by formula

$$\nu(w) \equiv \nu(x, \mathrm{v}) = \max\{\max_{1 \le i \le n}\{-x_i\}, \max_{1 \le i \le n}\{(\mathrm{v}A - p)_i\}, \sum_{i=1}^{n} |(\mathrm{v}A - p)_i \| x_i|\}.$$

It is easy to see that

$$\nu(w) = 0 \Leftrightarrow w = w^* = (x^*, \mathrm{v}^*)$$

and for any bounded subset $W \subset \Omega_k \times \mathbf{R}^m$ there exists $M > 0$ such that

$$\nu(w) = \nu(w) - \nu(w^*) \le M\|w - w^*\|, \ \forall w \in W.$$

So for a fixed $0 < \gamma < 0.5$ there exists such $k_0 > 0$ that for any $k \ge k_0$ and for a sequence $\{w^s = (x^s, \mathrm{v}^s)\}_{s=0}^{\infty}$, that is defined by (24), the next inequality

$$\nu(w^s) \le \gamma^s.$$

holds true.

Now we are going to describe the Primal Newton MBF Method.

To maintain the Primal MBF Method properties without solving the unconstrained optimization problem (24) at every step, one has to use a finite procedure, which enables to find an approximation for $\hat{x}$ with sufficient accuracy to remain the estimate (23). Below we describe such a type of method, that is based on the Primal MBF Method and on the global converging step size version of the Newton Method.

Let $\varepsilon > 0$ be small enough, $\{k_s\}_{s=0}^{\infty} : k_{s+1} > k_s, \lim_{s \to \infty} k_s = \infty$, $k = k(0) = k_0 > 0$, $d(0) = 1$, $0 < \gamma \le 1/2$ is fixed and $\Omega_k^+ = \{x = (x_1, \dots, x_n) : x_i \ge k^{-1}, i = 1, \dots, n\}$.

We start with $x := x^0 \in \Omega_k^+, u^0 = e = (1, \dots, 1) \in R^m$ and suppose $\bar{x}, x^s, u^s, \mathrm{v}^s$, $k(s), d(s)$ has already been determined. Then to find the approximation $(x^{s+1}, u^{s+1}, \mathrm{v}^{s+1})$ one has to fulfill the following operations:

0. start with $x := \bar{x}$

1. set $u := u^s$, $k := k(s)$, $d := d(s)$

2. find

$$\zeta = \zeta(x, u, k) = (k\,U)^{-1}\,\Delta^2(x, k)(v(x, u, k)A - p + U\,\Delta^{-1}(x, k)e)$$

and set $t := 1$

3. check $x + t\zeta \in \Omega_k$ and $F(x + t\zeta, u, k) - F(x, u, k) \le 1/3t(F_x'\,(x, u, k), \zeta)$

4. if $x + t\zeta \in \Omega_k$, the last inequality is fulfilled and $t = 1$ set $x := x + \zeta$ and go to 5; if $x + t\zeta \in \Omega_k$, the last inequality is fulfilled and $t < 1$ set $x := x + t\zeta$ and go to 2; if $x + t\zeta \,\overline{\in}\, \Omega_k$, and/or the inequality is not fulfilled set $t := t/2$ and go to 3.

5. if $\| \zeta \| \le \varepsilon$ go to 6; otherwise go to 2

6. set $\tilde{x} := x, \tilde{u} = \Delta^{-1}(\tilde{x}, k)u, \tilde{v} = (p - \tilde{u})A^T(A\,A^T)^{-1}, \tilde{w} = (\tilde{x}, \tilde{v})$; if $v(\tilde{w}) \le \gamma^{d+1}$ set $x^{s+1} = \tilde{x}, u^{s+1} = \tilde{u}; v^{s+1} = \tilde{v}, \bar{x} := x^{s+1}$, $d(s+1) = d(s) + 1, k(s+1) = k(s),\ s := s + 1, \varepsilon := \varepsilon\gamma$ and go to 1.

7. if $v(\tilde{w}) > \gamma^{d+1}$, set $\tilde{x}^{s+1} = \text{argmin}\left\{f_0(x^i) \mid i = 1, \dots, s+1\right\}, t_{s+1} = \max\left\{t \mid \bar{x} + t(\tilde{x}^{s+1} - \bar{x}) \in \Omega_k^+\right\}, \bar{x} := t_{s+1}\,\tilde{x}^{s+1} + (1 - t_{s+1})\bar{x}, u^{s+1} = u^0$, $\varepsilon := \varepsilon k^{-1}, k(s+1) = k_{s+1}, d(s+1) = 1, s := s + 1$ and go to 0.

<u>Assertion 4</u>. If the dual pair linear programming problems (1) - (2) are nondegenerate, then for a small enough $\varepsilon > 0$ and $0 < \gamma \le 0.5$ there exists such s_0 that for $s \ge s_0$:

1) the penalty parameter is permanent, i.e. $k(s) = k_{s_0} = k$ and the step size $t = 1$;

2) every Primal Newton MBF Method step, i.e. every Lagrange multiplier update ("large" step) requires $O\left(lg_2 lg_2 \varepsilon^{-1}\right)$ Newton steps;

3) the sequence $\left\{w^s = (x^s, v^s)\right\}_{s=0}^{\infty}$ converges to $w^* = (x^*, v^*)$ and the estimate

$$\max\left\{ \| x^s - x^* \|,\ \| u^s - u^* \| \| v^s - v^* \| \right\} \le \gamma^s, \quad s \ge s_0 \tag{37}$$

holds true.

Assertion 5 follows from theorems 1 and 2 and the Newton Method properties (see [29]). We will call the approximation (x^{s_0}, u^{s_0}), i.e. the moment when the Primal Newton MBF Method switches to the MBF trajectory, a "hot" start. Beginning at this moment, one can update u, i.e. improve the current approximation twice ($\gamma \le 1/2$) in every $O\left(lg_2 lg_2 \varepsilon^{-1}\right)$ Newton steps in the worst case.

The number s_0 depends on the measure of nondegeneracy $\sigma > 0$ and $\|B^{-1}\|$ as well as on the size of the problem (see section 9) and can be decreased by increasing k_0. Moreover, there exists such $k_0 > 0$, that $s_0 = 1$ for any $k \ge k_0$. The "hot start" phenomenon will be considered in detail from both theoretical as well as practical view points in forthcoming papers.

Remark 5.

The Primal Newton MBF Method can be furnished with a basis search procedure after updating u. Let $x_{i_1} \le x_{i_2} \le \cdots \le x_{i_n}$ and $u_{j_1} \ge u_{j_2} \ge \cdots \ge u_{j_n}$. If $\{i_1, i_2, \dots, i_{n-m}\} \cup \{j_{n-m+1}, \dots, j_n\} = \{1, \dots, n\}$ form the matrix $B = (a_i, i = i_{n-m+1}, \dots, i_n)$, and let $\det B \ne 0$ then find $x_B = B^{-1} q$ and $\mathrm{v} = p_B B^{-1}$. If $x_B > 0$ and $u_N = p_N - \mathrm{v}N > 0$ then $x = (0, x_B) = x^*$, $u = (u_N, 0) = u^*$, $\mathrm{v} = \mathrm{v}^*$.

7. MBF for the Dual Problem.

In this section we introduce and investigate the MBF for dual problem (2).

Let $V = \{\mathrm{v} : p - A^T\mathrm{v} \ge 0\} = \{\mathrm{v} : r_i(\mathrm{v}) = (p - A^T\mathrm{v})_i \ge 0, i = 1, \dots, n\} = \{\mathrm{v} : (p_i - (a_i, x)) \ge 0, i = 1, \dots, n\}$ be the feasible set of dual problem (2). Therefore, for any $k > 0$ we have $\mathrm{V} = \{\mathrm{v} : k^{-1} \ln(k\, r_i(\mathrm{v}) + 1) \ge 0, i = 1, \dots, n\}$.

In order to introduce the MBF for the dual problem, we consider the Classical Lagrangian for the next problem

$$\text{(38)} \qquad v^* = \arg\ \max\{(q, v) / k^{-1} \ln(k\, r_i(v) + 1) \geq 0,\ i = 1, \dots, n\}$$

that is equivalent to (2).

Let $y = (y_1, \dots, y_n)$ is a vector of Lagrange Multipliers for problem (38), $V_k = \{v : r_i(v) \geq -k^{-1}, i = 1, \dots, n\}$. Then

$$L(v, y, k) = (q, v) + k^{-1} \sum_{i=1}^{n} y_i \ln(k\, r_i(v) + 1) : V_k \times \mathbb{R}_+^n \times \mathbb{R}_+^1 \to \mathbb{R}^1,$$

is a Classical Lagrangian for problem (38). The MBF for dual problem (38) we define by formula

$$P(v, y, k) = \begin{cases} L(v, y, k), & v \in \operatorname{int} V_k \\ -\infty, & v \notin \operatorname{int} V_k \end{cases}.$$

Recall that $A^T = \begin{pmatrix} N^T \\ B^T \end{pmatrix}$, so $p_N - N^T v^* > 0^{n-m}, p_B - B^T v^* = 0^m$. Condition (3) can be rewritten as

$$\text{(39)} \qquad \begin{aligned} &y^*_i = 0,\ p_i - (N^T v^*)_i > 0, i = 1, \dots, n - m; \\ &y^*_i > 0, p_i - (B^T v^*)_i = 0, i = n - m + 1, \dots, n. \end{aligned}$$

In other words the dual pair LP is nondegenerate if

$$\sigma = \min \{\min \{p_i - (N^T v^*)_i \mid i = 1, \dots, n - m\}, \min\{y^*_i \mid i = n - m + 1, \dots, n\}\} > 0$$

and rank $B =$ rank $A = m$.

For $y_N = (y_1, \dots, y_{n-m})$ and $y_B = (y_{n-m+1}, \dots, y_n)$ we have $y^*_N = x^*_N = 0^{n-m}$; $y^*_B = x^*_B > 0^m$. Taking into account (39) we obtain

1°. $$P(\text{v}^*, y^*, k) = (q, \text{v}^*) = (p, y^*) = (p, x^*)$$

for any $k > 0$.

Further, we have

$$P'_{\text{v}}(\text{v}, y, k) = q - \sum_{i=1}^{n} \frac{y_i a_i}{k\, r_i(x) + 1},$$

where $a_i \in R^m$ – i column of the matrix A. Therefore again from (39) we obtain

2° $$P'_{\text{v}}(\text{v}, y^*, k)\,|_{\text{v}=\text{v}^*} = P'_{\text{v}}(\text{v}^*, x^*, k) = q - \sum_{i=n-m+1}^{n} \frac{y^*_i a_i}{k\, r_i(\text{v}^*) + 1}$$

$$= q - \sum_{i=n-m+1}^{n} y^*_i a_i = q - By^*_B = q - Bx^*_B = 0^m.$$

for any $k > 0$.

The function $P(\text{v}, y, k)$ is concave in $\text{v} \in V_k$ for any $y \in R^n_+$ and $k > 0$ so it follows from $P'_{\text{v}}(\text{v}, y^*, k)\,|_{\text{v}=\text{v}^*} = 0^m$ that

$$\text{v}^* \in \operatorname{Arg\,max}\{P(\text{v}, y^*, k) \mid \text{v} \in R^m\}.$$

Let $\Delta(\text{v}, k) = [\operatorname{diag}(k\, r_i(\text{v}) + 1)]_{i=1}^{n}$, $\Delta_N(\text{v}, k) = [\operatorname{diag}(k\, r_i(\text{v}) + 1)]_{i=1}^{n-m}$, $\Delta_B(\text{v}, k) = [\operatorname{diag}(k\, r_i(\text{v}) + 1)]_{i=n-m+1}^{n}$, $Y = [\operatorname{diag} y_i]_{i=1}^{n}$, $Y_N = [\operatorname{diag} y_i]_{i=1}^{n-m}$, $Y_B = [\operatorname{diag} y_i]_{i=n-m+1}^{n}$, then

$$P''_{\text{vv}}(\text{v}, y, k) = -k\, A\Delta^{-2}(\text{v}, k) Y A^T.$$

Taking into account (39) and $Y^*_B = [\operatorname{diag} y^*_i]_{i=n-m+1}^{n} = [\operatorname{diag} x^*_i]_{i=n-m+1}^{n} = X^*_B$, $Y^*_N = 0^{n-m, n-m}$, $\Delta_B(\text{v}^*, k) = I^m$

we obtain

3°. $$P''_{vv}(v, y^*, k)_{v=v^*} = P''_{vv}(v^*, x^*, k) = -k\, B\, Y^*_B\, B^T = -k\, B\, X^*_B\, B^T.$$

Let $\underline{y}^* = \min\{y^*_i \mid i = n-m+1, \dots, n\}, \bar{y}^* = \max\{y^*_i \mid i = n-m+1, \dots, n\}$ and

$$\mu = \text{mineigval}\, B\, B^T,\ M = \text{maxeigval}\, B\, B^T.$$

Then the next inequalities

4° $$k\, \bar{y}^*\, M(v, v) \geq (-P''_{vv}(v^*, y^*, k)v, v) \geq k\underline{y}^*\, \mu(v, v),\ \forall v \in \mathbb{R}^m$$

hold true for any $k > 0$.

So the condition number of the Hessian $-P''_{vv}(v^*, y^*, k)$, i.e.

$$\text{cond}\,(k\, B\, X^*_B\, B^T) = \frac{\text{mineigenval}\, B\, X^*_B\, B^T}{\text{maxeigval}\, B\, X^*_B\, B^T}$$

is independent on $k > 0$ and can be estimated by

$$\pi = (\underline{y}^*\, \mu)(\bar{y}^* M)^{-1}.$$

Moreover, the cond $P''_{vv}(v, y, k)$ is stable in the neighborhood of $(v^*, y^*) = (v^*, x^*)$ for any fixed $k > 0$.

The function $P(v, y^*, k)$ is strongly concave in $v \in V_k$ for any $k > 0$, so taking into account 2° we obtain

5° $$v^* = \text{argmax}\{P(v, y^*, k) \mid v \in \mathbb{R}^m\}$$

for any $k > 0$.

So the next theorem which describes the dual MBF $P(\mathrm{v}, y, k)$ properties at the primal and dual solution takes place.

Theorem 3. If the dual pair linear programming problem is nondegenerate, then the primal $x^* = y^*$ and the dual v^* solutions are unique and the properties 1°-5° take place for any $k > 0$.

Now we are going to prove a theorem similar to Theorem 1 for the dual MBF $P(\mathrm{v}, y, k)$. Let $\varepsilon > 0$ and $\delta > 0$ be small enough and the number $k_0 > 0$ is large enough. We consider m sets $D_i(y^*, k_0, \delta, \varepsilon) = D_i(\bullet) = \{(y_i, k) : y_i \geq \varepsilon, |y_i - y^*_i| \leq \delta k, k \geq k_0\}$, $i = n - m + 1, \dots, n$ and $(n - m)$ sets $D_i(y^*, k_0, \delta, \varepsilon) = D_i(\bullet) = \{(y_i, k) : 0 \leq y_i \leq \delta k, k \geq k_0\}$. Then $D(y^*, k_0, \delta, \varepsilon) = D(\bullet) = D_1(\bullet) \otimes \dots \otimes D_i(\bullet) \otimes \dots \otimes D_n(\bullet)$.

Also for any fixed $k \geq 0$ we define sets $Y^i_k = \{y_i : \max\{\varepsilon, y^*_i - \delta k\} \leq y_i \leq y^*_i + \delta k\}$, $i = n - m + 1, \dots, n$ and $Y^i_k = \{y_i : 0 \leq y_i \leq \delta k\}$, $i = 1, \dots, n - m$,. Now let $Y_k = Y^1_k \otimes \dots \otimes Y^i_k \otimes \dots \otimes Y^n_k$, then $D(\bullet) = \{(y, k) : y \in Y_k, k \geq k_0\}$.

Theorem 4. If the dual pair linear programming problem is nondegenerate, then there exists $k_0 > 0$ and small enough $\varepsilon > 0$ and $\delta > 0$ such that for any $(y, k) \in D(y^*, k_0, \delta, \varepsilon)$ the following statements hold true.

1) There exists a unique vector $\hat{\mathrm{v}} = \hat{\mathrm{v}}(y, k) = \operatorname{argmin} \max\{P(\mathrm{v}, y, k) \mid \mathrm{v} \in \mathbb{R}^m\}$ such that

$$P'_{\mathrm{v}}(\hat{\mathrm{v}}, y, k) = q - \sum_{i=1}^{n} \frac{y_i a_i}{k\, r_i(\hat{\mathrm{v}}) + 1} =$$

$$q - \sum_{i=1}^{n} \hat{y}_i a_i = q - A\hat{y} = A\Delta^{-1}(\hat{\mathrm{v}}, k)y = q - A\hat{y} = 0^m,$$

2) the estimate

$$\max\{\|\hat{\mathrm{v}} - \mathrm{v}^*\|, \|\hat{y} - y^*\|\} \leq c k^{-1} \|y - y^*\| \tag{40}$$

holds true, and $c > 0$ is independent of $k \geq k_0$;

3) for the couple $\hat{v} = \hat{v}(y, k)$ and $\hat{y} = \hat{y}(y,k) = \Delta^{-1}(\hat{v}, k)y$ we have $v(y^*, k) = v^*$ and $\hat{y}(y^*, k) = y^*$, i.e. $y^* \in \mathbf{R}^n_+$ is a fixed point of the mapping $y \rightarrow \hat{y}(y,k)$ for any $k > 0$;

4) for any fixed $k \geq k_0$ there exists $0 < \alpha_k < 1$ such that

$$\operatorname{cond}(-P''_{vv}(\hat{v}, \hat{y}, k)) = \operatorname{cond} A\, \Delta^{-2}(\hat{v}, k) Y A^T \geq \alpha_k \pi$$

for any $y \in Y_k$.

Proof 1) We will prove Theorem 4 following the scheme that we used to prove Theorem 1. Let $k > 0$ is large enough and $\varepsilon > 0$ and $\delta > 0$ are small enough. We consider three vectors: $t = (t_i = (y_i - y^*_i)k^{-1}, i = 1, \dots, n), \hat{y}_B = (\hat{y}_i, i = n - m + 1, \dots, n)$ and $h(v, t, k) = k \sum_{i=1}^{n-m} t_i(k\, r_i(v) + 1)^{-1} a_i$. Also we consider a set $S(0, \delta) = \{t : |t_i| \leq \delta, i = 1, \dots, n\}$ and a diagonal matrix $Y_B = [\operatorname{diag} y_i]^n_{i = n - m + 1}$. For any $k > 0$ and $x \in S(x^*, \varepsilon) = \{x : \|x - x^*\| \leq \varepsilon\}$, $t \in S(0, \delta)$ the vector function $h(v, t, k)$ is smooth and $h(v^*, 0, k) = 0^m$ $h'_v(v^*, 0, k) = 0^{m,m}$, $h'_{\hat{y}_B}(v, 0, k) = 0^{m,m}$.

On the $S(x^*, \varepsilon) \times S(\hat{y}_B, \varepsilon) \times S(0, \delta) \times (0, +\infty)$, we consider the map $\Phi(v, \hat{y}_B, t, k)$: $\mathbf{R}^{2m+n+1} \rightarrow \mathbf{R}^{2m}$ defined by formula :

$$\Phi(v, \hat{y}_B, t, k) = (q - \sum_{l=n-m+1}^{n} \hat{y}_l a_l - h(v, t, k);$$

$$k^{-1}(kt_i + y^*_i)(k\, r_i(v) + 1)^{-1} - k^{-1} \hat{y}_i,\ i = n - m + 1, \dots, n)$$

Taking into account (39) and $h(v^*, 0, k) = 0^m$ we obtain $\Phi(x^*, y^*_B, 0, k) = 0^{2m}$ for $\forall\, k > 0$. In view of $h'_v(v^*, 0, k) = h'_{\hat{y}_B}(v^*, 0, k) = 0^{m,m}$ we obtain

$$\Phi_k = \Phi'_{v\hat{y}_B}(v^*, y^*_B, 0, k) = \begin{pmatrix} 0^{m,m} & -B \\ -Y^*_B B^T & -k^{-1} I^m \end{pmatrix}.$$

Because of $Y^*_B > 0$ and the fact that rank $B = m$, matrix Φ_k is invertible and

$$\Phi_k^{-1} = \begin{pmatrix} k^{-1}(Y^*_B B^T)^{-1} B^{-1} & -(Y^*_B B^T)^{-1} \\ -B^{-1} & 0^{m,m} \end{pmatrix}. \tag{41}$$

Let $k_1 > k_0$ be large enough, $K = \{0 \in \mathbf{R}^n\} \times [k_0, k_1]$. Since $\Phi(x^*, y^*_B, 0, k) = 0^{2m}$ and the matrix Φ_k is nonsingular for all $k \in [k_0, k_1]$ it follows from the second implicit function theorem (see [3] p. 12) that there exist $\varepsilon > 0$ and $\delta > 0$ and smooth enough vector-functions $v(t,k) = (v_1(t,k), \dots, v_m(t,k))$ and $\hat{y}_B(t,k) = (\hat{y}_{n-m+1}(t,k), \dots, \hat{y}_n(t,k))$ defined uniquely in the neighborhood of $S(K, \delta) = \{(t,k) : |t_i| \le \delta, i = 1, \dots, n,$
$k \in [k_0, k_1]\}$ that $v(0, k) = v^*, \hat{y}_B(0, k) = y^*_B$ and

$$\Phi(v(t,k), \hat{y}_B(t,k), t, k) = \Phi(v(\bullet), \hat{y}_B(\bullet), \bullet) \equiv 0^{2m}$$

for $\forall (t, k) \in S(K, \delta)$.

2) To prove estimation (40) we are rewriting the last identity. We obtain

$$q - \sum_{i=n-m+1}^{n} \hat{y}_i(t, k) a_i - h(v(t, k), t, k) \equiv 0^m \tag{42}$$

$$k^{-1} \hat{y}_i(t, k) - (t_i + k^{-1} y^*_i)(k\, r_i(v(t, k)) + 1)^{-1} \equiv 0, i = n - m + 1, \dots, n. \tag{43}$$

We also define

$$\hat{y}_i(t, k) = k\, t_i (k\, r_i(v(t, k)) + 1)^{-1}, i = 1, \dots, n - m. \tag{44}$$

Note that $\hat{y}_i(0, k) = y^*_i = x^*_i = 0, i = 1, \dots, n - m.$ Therefore for the vector $\hat{y}_N(t, k)) = \hat{y}_N(\bullet) = (\hat{y}_1(t, k), \dots, \hat{y}_{n-m}(t, k))$ we obtain $\|\hat{y}_N(\bullet) - y^*_N\| = \|\hat{y}_N(\bullet)\|$. If

$\delta > 0$ is small enough, then for any $(t, k) \in S(K, \delta)$, we have $\|v(\bullet) - v^*\| \le \varepsilon$ and $r_i(v(\bullet)) \ge \frac{\sigma}{2}$, $i = 1, \dots, n-m$, therefore $\hat{y}_i(t, k) = \frac{y_i - y^*_i}{k} \bullet (r_i(v(\bullet)) + k^{-1})^{-1}$, so we have $\hat{y}_i(\bullet) = \frac{2}{\sigma} \; \frac{y_i - y^*_i}{k}$, $i = 1, \dots, n-m$, and $\|\hat{y}_N(t, k)\| = \|\hat{y}_N(t, k) - y^*_N\| \le \frac{2}{\sigma} \bullet k^{-1} \|y_N - y^*_N\|$.

Now we are going to show that estimate (40) holds for $v(t, k) = v(\bullet)$ and $\hat{y}_B(t, k) = \hat{y}_B(\bullet) = (\hat{y}_{n-m+1}(\bullet), \dots, \hat{y}_n(\bullet))$. To this end we differentiate the identities (42) and (43) with respect to t.

From (42) we obtain

$$0^{m,m} \bullet v'_t(t, k) - B\hat{y}'_{B,t}(t, k) = h'_t(v(t, k), t, k). \tag{45}$$

Let $\Delta_B(t, k) = \Delta_B(\bullet) = [\mathrm{diag}(k\, r_i(v(t, k)) + 1)]_{i=n-m+1}^{n})$, $\Delta_N(t, k) = \Delta_N(\bullet) = [\mathrm{diag}(k\, r_i(v(t, k)) + 1)]_{i=1}^{n-m}$.

From (43) we obtain

$$\begin{aligned} &[\mathrm{diag}(k\, t_i + y^*_i)]_{i=n-m+1}^{n} \Delta_B^{-2}(\bullet) B^T v'_t(\bullet) - k^{-1} \hat{y}_{B,t}(\bullet) \\ &= [0^{m,n-m}; \Delta_B^{-1}(\bullet)]. \end{aligned} \tag{46}$$

Multiplying both sides of the system to $\Delta_B^2(\bullet)$ we obtain

$$\begin{aligned} &[\mathrm{diag}(k\, t_i + y^*_i)]_{i=n-m+1}^{n} B^T v'_t(\bullet) - k^{-1} \Delta_B^2(\bullet) \hat{y}_{B,t}(\bullet) \equiv \\ &[0^{m,n-m}; \Delta_B(\bullet)]. \end{aligned} \tag{47}$$

Let

$$\Phi'(\bullet) = \begin{bmatrix} 0^{m,m} & -B \\ -[\mathrm{diag}(k\, t_i + y^*_i)] B^T & -k^{-1} I^m \end{bmatrix}$$

then combining (45) and (47), we obtain

$$(48)\qquad \begin{bmatrix} v'_t(\bullet) \\ \hat{y}'_{B,t}(\bullet) \end{bmatrix} = (\Phi'(\bullet))^{-1} \begin{bmatrix} h'_t(v(\bullet),\bullet) \\ [0^{m,n-m};\Delta_B(\bullet)] \end{bmatrix} = (\Phi'(\bullet))^{-1} \times R(\bullet).$$

Considering the last system for $t = 0^n$ we obtain

$$\begin{bmatrix} v'_t(0,k) \\ \hat{y}_{B,t}(0,k) \end{bmatrix} = \left[\Phi_k^{-1}\right] \begin{bmatrix} h'_t(v(0,k),0,k) \\ [0^{m,n-m}, I^m] \end{bmatrix}.$$

Taking into account $h'_t(v(0,k),0,k) = [N\,k[\mathrm{diag}(k\,r_i(v^*)+1)^{-1}]_{i=1}^{n-m};0^{m,m}] =$ $[N(k\,\Delta_N^{-1}(0,k));0^{m,m}] = [N[\mathrm{diag}(p_N - N^T v^*)_i + k^{-1}]_{i=1}^{n-m};0^{m,m}]$ and (41) we obtain

$$\begin{bmatrix} v'_t(0,k) \\ \hat{y}'_{B,t}(0,k) \end{bmatrix} = \begin{bmatrix} k^{-1}(Y^*_B B^T)^{-1} B^{-1} & -(Y^*_B B^T)^{-1} \\ -B^{-1} & 0^{m,m} \end{bmatrix} \times$$

$$\begin{bmatrix} N(k\,\Delta_N^{-1}(0,k)) & 0^{m,m} \\ 0^{m,n-m} & I^m \end{bmatrix} =$$

$$\begin{bmatrix} k^{-1}(Y^*_B B^T)^{-1} B^{-1} N(k\,\Delta_N^{-1}(0,k)) & -(Y^*_B B^T)^{-1} \\ -B^{-1} N(k\,\Delta_N^{-1}(0,k)) & 0^{m,m} \end{bmatrix} = \Psi(A,p,q,k).$$

Now we are going to show that there exists $k_0 > 0$ and an independent on $k \ge k_0$ estimate for the norm $\|\Psi(A,p,q,k)\|$.

Let $A_{11} = k^{-1}(Y^*_B B^T)^{-1} B^{-1} N(k\,\Delta_N^{-1}(0,k))$, $A_{12} = -(Y^*_B B^T)^{-1}$,

$$A_{21} = -B^{-1} N(k\,\Delta_N^{-1}(0,k)),\quad A_{22} = 0^{m,m}.$$

Then

$$\|\Psi(A,p,q,k)\| \le \max\{\|A_{11}\| + \|A_{12}\|, \|A_{21}\|\}.$$

Further, we have $k\Delta_N^{-1}(0,k) = [\mathrm{diag}((p_N - N^T v^*)_i + k^{-1})]_{i=1}^{n-m}$ and $\|k\Delta_N^{-1}(0,k)\| \le$ $\|U^{*\,-1}_N\|$. Therefore

$$\|A_{11}\| \le k^{-1}\|Y^{*-1}_B\|\,\|B^{-T}\|\,\|B^{-1}\|\,\|N\|\,\|U^{*-1}_N\|, \|A_{12}\| \le \|Y^{*-1}_B\|\,\|B^{-1}\| \text{ and}$$

$$\|A_{21}\| \le \|B^{-1}\|\,\|N\|\,\|U^{*-1}_N\|.$$

Taking into account (39) we obtain

$$\|A_{11}\| \le k^{-1}\sigma^{-2}\|B^{-T}\|\,\|B^{-1}\|\,\|N\|, \|A_{12}\| \le \sigma^{-1}\|B^{-T}\| \text{ and}$$
$$\|A_{21}\| \le \sigma^{-1}\|B^{-1}\|\,\|N\|.$$

Hence

$$\|\Psi(A, p, q, k)\| \le \sigma^{-1}\max\{k^{-1}\sigma^{-1}\|B^{-T}\|\,\|B^{-1}\|\,\|N\| + \|B^{-T}\|, \|B^{-1}\|\,\|N\|\}$$
$$\le \sigma^{-1}\max\{\|B^{-T}\|(k^{-1}\sigma^{-1}\|B^{-1}\|\,\|N\| + 1), \|B^{-1}\|\,\|N\|\}.$$

Therefore

$$\|\Psi(A, p, q, k)\| \le 2\sigma^{-1}\max\{\|B^{-T}\|, 0.5\|B^{-1}\|\,\|N\|\} = c_1$$

for any $k \ge k_0 = \sigma^{-1}\|B^{-1}\|\,\|N\|$. Hence

$$\max\{\|v'_t(0, k)\|, \|\hat{y}'_{B,t}(0, k)\|\} \le c_1, \quad \forall k \ge k_0.$$

Taking into account (48) we obtain for small enough $\delta > 0$ and all $(t, k) \in S(K, \delta)$ that

$$\|\Phi'^{-1}(v(\tau t, k), \hat{y}_B(\tau t, k); \tau t, k)\mathbf{R}(v(\tau t, k), \tau t, k)\| \le 2\, c_1$$

for $\forall \tau \in [0, 1]$. So

$$\begin{bmatrix} v(t, k) - v(0, k) \\ \hat{y}_B(t, k) - \hat{y}_B(0, k) \end{bmatrix} = \begin{bmatrix} v(t, k) - v^* \\ \hat{y}_B(t, k) - y^*_B \end{bmatrix} =$$
$$\int_0^1 \Phi'^{-1}(v(\tau t, k), \hat{y}_B(\tau t, k); \tau t, k)\mathbf{R}(v(\tau t, k); \tau t, k)[t]d\tau$$

and

$$\max\{\|\mathrm{v}(t, k) - \mathrm{v}^*\|, \|\hat{y}_B(t, k) - y^*_B\|\} \le 2\, c_1 \|t\| = \frac{2\, c_1}{k} \|y - y^*\|.$$

Let $\hat{\mathrm{v}}(y, k) = \mathrm{v}(\frac{y - y^*}{k}, k)$ and $\hat{y}(y, k) = (\hat{y}_N(\frac{y - y^*}{k}, k), \hat{y}_B(\frac{y - y^*}{k}, k))$ then for $c = 2 \max\{c_1, \sigma^{-1}\}$, the next estimate

$$\max\{\|\hat{\mathrm{v}}(y, k) - \mathrm{v}^*\|, \|\hat{y}(y, k) - y^*\|\} \le \frac{c}{k} \|y - y^*\|$$

holds true for $\forall (y, k) \in D(y^*, k_0, \delta, \varepsilon)$, i.e. we have proven estimate (40).

3) Statement (3) follows from estimate (40).

4) Statement 4) follows from the nonsingularity conditions of (39) and estimate (40). We have proven Theorem 4.

<u>Remark 6.</u>

All results of Theorem 4 remain true if instead of $P(\mathrm{v}, y, k)$ one will consider next the multiparameter MBF for the dual problem

$$P(\mathrm{v}, y, K) = \begin{cases} (q, \mathrm{v}) + \sum k_i^{-1} y_i \, ln(k_i r_i(\mathrm{v}) + 1), & \mathrm{v} \in \operatorname{int} V_k \\ -\infty \quad , & \mathrm{v} \notin \operatorname{int} V_k \end{cases}$$

where $V_k = \{\mathrm{v} : r_i(\mathrm{v}) \ge -k_i^{-1}, i = 1, \dots, n\}$, $\min_{1 \le i \le n} k_i \ge k_0 > 0$ and $k_0 > 0$ is large enough.

The same is true for the next version of MBF that corresponds to Carroll's function for the dual problem

$$Q(\mathrm{v}, y, K) = \begin{cases} (q, \mathrm{v}) - \sum_{i=1}^{n} k_i^{-1} y_i((k_i r_i(\mathrm{v}) + 1)^{-1} - 1), & \mathrm{v} \in \mathrm{int} V_k \\ \infty & , \mathrm{v} \notin \mathrm{int} V_k \end{cases}$$

Now we will consider the one parameter Shifted Barrier Function for the dual problem. The one parameter Shifted Barrier Function for the dual LP is obtained by setting $y = e = (1, \dots, 1) \in \mathbf{R}^n$ in the Modified Barrier Function for the dual problem. We have

$$N(\mathrm{v}, k) \equiv P(\mathrm{v}, e, k) = \begin{cases} (q, \mathrm{v}) + k^{-1} \sum \ln(k\, r_i(\mathrm{v}) + 1), \mathrm{v} \in \mathrm{int}\ V_k \\ -\infty & , \mathrm{v} \notin \mathrm{int}\ V_k \end{cases}.$$

The next assertion is a consequence of Theorem 4.

Assertion 5. Let V be a polytope, then

1) for any $k > 0$ there exists a vector

$$\mathrm{v}(k) = \mathrm{argmax}\{N(\mathrm{v}, k) \mid \mathrm{v} \in \mathbf{R}^m\}$$

such that

$$N'_{\mathrm{v}}(\mathrm{v}(k), k) = q - A^{-1} \Delta(\mathrm{v}(k), k) e = 0^m$$

2) if the complementary conditions in (39) are fulfilled and rank A = rank $B = m$, then $\mathrm{v}(k)$ is unique and there exists $k_0 > 0$ such that for the couple of vectors $\mathrm{v}(k)$ and $y(k) = \Delta^{-1}(\mathrm{v}(k), k)e$, the estimate

(49) $$\max\{\|\mathrm{v}(k) - \mathrm{v}^*\|, \|y(k) - y^*\|\} \le c\, k^{-1}$$

holds and the constant $c > 0$ is independent of $k \ge k_0$.

Remark 7.

All statements of Assertion 5 hold true for multiparameter Shifted Barrier Functions $P(\mathrm{v}, e, K)$ and $Q(\mathrm{v}, e, K)$, if $\min_{1 \le i \le n} k_i \ge k_0$.

8. Dual MBF Method

In this section we consider a method for simultaneous solution primal and dual LP, based on MBF for the dual problem. First we consider the permanent parameter version. Let $k \ge k_0$ is fixed, $\mathrm{v}^0 \in \operatorname{int} V_k$ and $y^0 = e = (1, \dots, 1) \in \mathbf{R}^n$ are initial approximations for dual and primal problems.

The permanent parameter version of the dual MBF method consists of finding a sequence $\{\mathrm{v}^s, y^s\}_{s=0}^{\infty}$ by formulas:

(50)
$$\begin{cases} \mathrm{v}^{s+1} = \operatorname{argmax} \{P(\mathrm{v}, y^s, k) \mid \mathrm{v} \in \mathbf{R}^m\} \\ y^{s+1} = \Delta^{-1}(\mathrm{v}^{s+1}, k) y^s . \end{cases}$$

The next assertion is a consequence of Theorem 4 and Assertion 5.

Assertion 6. If the dual pair linear programming problems are nondegenerate, then for any fixed $0 < \gamma < 0.5$ there exists such $k_0 > 0$ that the sequence $\{\mathrm{v}^s, y^s\}_{s=0}^{\infty}$ converge to the dual v^* and primal $y^* = x^*$ solution and the next estimate

(51)
$$\max\{\|\mathrm{v}^s - \mathrm{v}^*\|, \|y^s - y^*\|\} \le \gamma^s$$

holds true for any $k \ge k_0$.

The varying parameter version of the Dual MBF Method one can obtain by changing from step to step parameter k. Let $\mathrm{v}^0 \in V_k, y^0 = e = (1, \dots, 1) \in \mathbf{R}^n$ and we consider the sequence $\{k_s\}_{s=0}^{\infty} : k_s < k_{s+1}, k_s \to \infty$. The varying parameter version of the Dual MBF Method consists of finding a sequence $\{\mathrm{v}^s, y^s\}_{s=0}^{\infty}$ by formulas

$$(52) \qquad \begin{cases} \text{a) } v^{s+1} = \text{argmax}\,\{P(v, y^s, k_s) \mid v \in \mathbf{R}^m\} \\ \text{b) } y^{s+1} = \Delta^{-1}(v^{s+1}, k_s)y^s. \end{cases}$$

The next assertion which is a consequence of Theorem 4 and Assertion 5 takes place.

Assertion 7. If the dual pair linear programming problems are nondegenerate, then for any sequence $\{k_s\}_{s=0}^{\infty}$ there exists a sequence $\{\gamma_s\}_{s=0}^{\infty}$; $0.5 \geq \gamma_0 \geq \gamma_1 \geq \cdots \gamma_s \geq \cdots$ and $\gamma_s \to 0$ that $\{v^s, y^s\}_{s=0}^{\infty}$ converges to the primal-dual solution $(v^*, y^*) = (v^*, x^*)$ and the next estimate

$$\max\{\|v^s - v^*\|, \|y^s - y^*\|\} \leq \gamma_0 \times \gamma_1 \ldots \times \gamma_s$$

holds true.

Now we are going to describe the Dual Newton MBF Method. To this end we introduce an indicator function $\nu(w) : V_k \times R_+^m \to R_+^1$ by formula

$$\nu(w) \equiv \nu(v, y) = \max\{\|P'_v(v, y, k)\|, \max_{1 \leq i \leq n} \{-r_i(v)\}, \sum_{i=1}^{n} |r_i(v)| y_i\}.$$

It is easy to see that

$$\nu(w) = 0 \Leftrightarrow w = w^* = (v^*, y^*) = (v^*, x^*).$$

For any bounded $W \subset V_k \times R_+^m$ there exists such $M > 0$ that

$$\nu(w) = \nu(w) - \nu(w^*) \leq M\|w - w^*\|, \ \forall w \in W.$$

So for $0 < \gamma \leq 0.5$ there exists such $k \geq k_0$ that for the sequence $\{w^s = (v^s, y^s)\}_{s=0}^{\infty}$ defined by (52) the next estimate

$$\nu(w^s) \leq \gamma^s$$

holds true.

The Dual Newton MBF Method consists of using the Newton Method for finding an approximation $\tilde{v}^{s+1}$ for the dual vector v^{s+1} (see (51a)) and update the primal vector $y = (y_1, \dots, y_n)$ by using instead of v^{s+1} the vector $\tilde{v}^{s+1}$ in formula (52b).

Let $\varepsilon > 0$ be small enough, $\{k_s\}_{s=0}^{\infty} : k_{s+1} > k_s, \lim_{s\to\infty} k_s = \infty, k = k(0) = k_0$, $d(0) = 1, 0 < \gamma \le 0.5$ is fixed and $V_k^+ = \{v : r_i(v) \ge k^{-1}, i = 1, \dots, n\}$.

We start with $v = v^0 \in V_k^+, y^0 = e = (1, \dots, 1) \in \mathbf{R}^n$, let $\bar{v}, v^s, y^s, k(s), d(s)$ be already found. To find the next approximation (v^{s+1}, y^{s+1}) one has to fulfill the next operations:

0. start with $v := \bar{v}$
1. set $y := y^s, k := k(s), d := d(s)$;
2. find ζ by solving the normal system of equations

$$A\,\Delta^{-2}(v, k) Y A^T \zeta = q - A\,\Delta^{-1}(v, k) y$$

and set $t := 1$;

3. check $v + t\zeta \in V_k$ and inequality

$$P(v + t\zeta, y, k) - P(v, y, k) \ge 0.5t(P'_v(v, y, k), \zeta); \tag{53}$$

4. if $x + t\zeta \in V_k$, inequality (53) is fulfilled and $t = 1$ set $v := v + \zeta$ and go to 5; if $v + t\zeta \notin V_k$ inequality (53) is fulfilled and $t < 1$ set $v := v + t\zeta$ and go to 2; if $v + t\zeta \notin V_k$ and/or inequality (53) is not fulfilled, set $t := 0.5t$ and go to 3;
5. if $\|\zeta\| \le \varepsilon$ go to 6; otherwise go to 2.
6. set $\tilde{v} := v, \tilde{y} = \Delta^{-1}(\tilde{v}, k)y, \tilde{w} = (\tilde{v}, \tilde{y})$ and if $v(\tilde{w}) \le \gamma^{d+1}$ set $v^{s+1} = \tilde{v}, y^{s+1} = \tilde{y}$, start with $v := v^{s+1}, d(s+1) = d(s) + 1, k(s+1) = k(s), s + 1 := s, \varepsilon := \gamma\varepsilon$ and go to 1.
7. if $v(\tilde{w}) > \gamma^{d+1}$ set $\tilde{v}^{s+1} = \operatorname{argmax}\{(q, v^i) \mid i = 1, \dots, s+1,\}$ $t_{s+1} =$

$\max\{t \mid \bar{v} + t(\tilde{v}^{s+1} - \bar{v}) \in V_k^+\}, \bar{v} := t_{k+1}\tilde{v}^{s+1} + (1 - t_{s+1})\bar{v}, y^{s+1} =$
$y^0 = e = (1, \dots, 1) \in \mathbf{R}^n, \varepsilon := \varepsilon k^{-1}, k(s+1) = k_{s+1}, s+1 := s$ and go to 0.

The next assertion is a consequence of Theorem 4, Assertion 5 and the Newton method properties (see [29]).

Assertion 8. If the dual pair linear programming problems are nondegenerate then for any fixed $0 < \gamma \le 0.5$ there exist a small enough $\varepsilon > 0$ and such number s_0 that for $s \ge s_0$:

1) the parameter $k = k(s) = k_{s_0}$ is permanent, $v + t\zeta \in V_k$ and inequality (53) is fulfilled for $t = 1$.

2) every Dual Newton MBF "large" step, i.e. every update y requires $0(\ln \ln \varepsilon^{-1})$ Newton steps of maximization $P(v, y, k)$ in $v \in V_k$,

3) the sequence $\{w^s = (v^s, y^s)\}_{s=0}^{\infty}$ converges to $w^* = (v^*, y^*) = (v^*, x^*)$ and the estimate

$$\max\{\|v^s - v^*\|, \|y^s - y^*\|\} \le \gamma^s, s \ge s_0$$

holds true.

Moreover, for any $0 < \gamma \le 0.5$ there exists such k_0 that for any $k \ge k_0$ all statements of Assertion 8 take place for $s_0 = 1$. Parameter k_0 depends on parameter $0 < \gamma \le 0.5$, measure of nondegeneracy $\sigma > 0$, mineigval $B^T B$ as well as on the size of the problem. In the next section we shall consider this question in more detail.

9. Condition Numbers of Primal and Dual LP

In this section we introduce the *condition numbers* for the primal and dual LP. These notions are different from the condition for the LP that was introduced in [16]. The condition numbers of the primal and dual of LP will be characterized by the key parameters that accumulate the most important information about the primal and dual LP.

The condition number of the dual pair of LP is responsible for the "hot start", rate of convergence of primal and dual MBF methods and complexity of the primal and dual Newton MBF Methods.

Let us recall that the norm $\|A\| = \|A\|_\infty = \max_{1 \le j \le m} \sum_{i=1}^{n} |a_{ji}|$, is subordinate to the vector norm $\|x\| = \|x\|_\infty = \max_{1 \le i \le n} |x_i|$.

Let $\Pi(A,p,q) \equiv \Pi(A,p,q,\infty)$, $\Pi_1(A,p,q) \equiv \Pi_1(A,p,q,\infty)$ and $\Pi_2(A,p,q) \equiv \Pi_2(A,p,q,\infty)$, then

$$\Pi_1(A,p,q) = \begin{bmatrix} -U^{*-1}_N & 0^{n-m,m} \\ B^{-1} N U^{*-1}_N & 0^{m,m} \end{bmatrix}, \Pi_2(A,p,q) = \begin{bmatrix} 0^{n-m,n-m} & N^T B^{-T} X^{*-1}_B \\ 0^{m,n-m} & -B^{-T} X^{*-1}_B \end{bmatrix}$$

and

$$\|\Pi(A,p,q)\| = \max\{\|\Pi_1(A,p,q)\|, \|\Pi_2(A,p,q)\|\}$$

where

$$\|\Pi_1(A,p,q)\| = \max\{\|U^{*-1}_N\|, \|B^{-1} N U^{*-1}_N\|\}$$

and

$$\|\Pi_2(A,p,q)\| = \max\{\|N^T B^{-T} X^{*-1}_B\|, \|B^{-T} X^{*-1}_B\|\}.$$

We will say that the primal $P(A,p,q)$ is scaled by the vector $t = (t_1, \dots, t_j, \dots, t_m) > 0^n$ if instead of the matrix $A = (N, B)$ and the vector q one takes the matrix $\bar{A} = TB = (\bar{N}, \bar{B}) = (TN, TB)$ and the vector $\bar{q} = Tq$ where $T = [\operatorname{diag} t_j]_{j=1}^{m}$ it is clear that the feasible set $\{x : Ax = q, x \geq 0^n\}$ is an invariant with respect to scaling the primal problem $P(A, p, q)$.

Before we estimate the norms $\|\Pi_1(A, p, q)\|$, $\|\Pi_2(A, p, q)\|$ and $\|\Pi(A, p, q)\|$, let us consider some other invariants of the problem $P(A, p, q)$ with respect to scaling.

We consider the scaled matrixes $\overline{N} = TN$ and $\overline{B} = TB$ and the scaled vector $\overline{q} = Tq$, then $\overline{v} = p_B \overline{B}^{-1}$, $\overline{u}_N = p_N - \overline{v}\,\overline{N} = p_N - p_B \overline{B}^{-1} \overline{N} = p_N - p_B (TB)^{-1} TN = p_N - p_B B^{-1} T^{-1} TN = p_N - p_B B^{-1} N = p_N - v^* N = u^*_N$, hence $\overline{U}_N = [\operatorname{diag} \overline{u}_i]_{i=1}^{n-m} = U^*_N$.

Further $\overline{x}_B = \overline{B}^{-1} \overline{q} = (TB)^{-1} Tq = B^{-1} T^{-1} Tq = B^{-1} q = x^*_B$, therefore $\overline{X}_B = [\operatorname{diag} \overline{x}_i]_{i=n-m+1}^{n} = X^*_B$.

Taking into account $\overline{B}^{-1} \overline{N} = (TB)^{-1}(TN) = B^{-1} N$ and $\overline{U}^{-1} = U^{*-1}_N$ we obtain $\overline{B}^{-1} \overline{N}\, \overline{U}_N^{-1} = B^{-1} N U^{*-1}_N$. Also $(\overline{B}^{-1} \overline{N})^T = \overline{N}^T (\overline{B}^{-1})^T = (TN)^T ((TB)^{-1})^T = N^T T (B^{-1} T^{-1})^T = N^T T T^{-1} B^{-T} = N^T B^{-T}$, hence, $\overline{N}^T \overline{B}^{-T} \overline{X}_B = N^T B^{-T} X^{*-1}_B$.

In other words the matrices U^*, X^*_B, $B^{-1} N U^{*-1}_N$ and $N^T B^{-T} X^{*-1}_B$, the vectors $u^* = (u^*_N, u^*_B)$, $x^* = (x^*_N, x^*_B)$ and hence the measure of nondegeneracy $\sigma > 0$ are invariants with respect to scaling the primal problem $P(A, p, q)$.

Later we will assume that the primal problems have already been scaled by the vector $t = (t_1, \dots, t_j, \dots, t_m)$, $t_j = \|a^j\|^{-1} = (\max_{1 \leq i \leq n} |a_{ji}|)^{-1}$, $j = 1, \dots, m$, i.e. we assume that $\|a^j\| = \max_{1 \leq i \leq n} |a_{ji}| = 1, j = 1, \dots, m$.

Problem (1) with $\|a^j\| = 1, j = 1, \dots, m$ we will call *normalized* primal LP. The value

$$\operatorname{cond} P(A, p, q) = \|\Pi(A, p, q)\|$$

we will call *condition number* of the *primal LP problem* $P(A, p, q)$.

Let us estimate cond $P(A,p,q)$ for the normalized $P(A,p,q)$ We have

$$\|\Pi_1(A,p,q)\| \le \max\{\|U^{*}{}_N^{-1}\|, \|B^{-1}N\|\ \|U^{*}{}_N^{-1}\|\} \le$$
$$\sigma^{-1}\max\{1, \|B^{-1}\|\max_{1\le j\le m}\sum_{i=1}^{n-m} |a_{ji}|\} \le$$
$$\sigma^{-1}\max\{1,(n-m)\|B^{-1}\|\}.$$

Then $\|B^{-1}\| \le \sqrt{m}\,\|B^{-1}\|_2$, also $\|B\|_2 \bullet \|B^{-1}\|_2 = (M\mu^{-1})^{1/2}$ and $\|B\|_2 = M^{1/2}$, therefore $\|B^{-1}\| \le (m\mu^{-1})^{1/2}$. So $\|\Pi_1(A,p,q)\| \le \sigma^{-1}\max\{1,(n-m)(m\mu^{-1})^{1/2}\}$.

It is easy to show that for any matrix $B = \|b_{ij}\|_{\substack{i=1,\dots,m\\ j=1,\dots,m}}$ such that $\max_{i,j} |b_{ij}| \le 1$ we have $\mu \le m$ therefore

$$\|\Pi_1(A,p,q)\| \le \sigma^{-1}(n-m)(m\mu^{-1})^{1/2}.$$

For the norm $\|\Pi_2(A,p,q)\|$ we obtain

$$\|\Pi_2(A,p,q)\| \le \max\{\|N^T B^{-T}\|\ \|X^{*}{}_B^{-1}\|, \|(B^{-T})\,X^{*}{}_B^{-1}\|\}$$
$$\le \sigma^{-1}\|B^{-T}\|\max\{\|N^T\|, 1\}$$
$$\le \sigma^{-1}\|B^{-T}\|\max_{1\le i\le n-m}\left\{\sum_{j=1}^{m} |a_{ij}|, 1\right\} \le \sigma^{-1}\|B^{-T}\|m.$$

Then $\|B^{-T}\| \le \sqrt{m}\,\|B^{-T}\|_2 = \sqrt{m}\,\|B^{-T}\|_2 = (m\mu^{-1})^{1/2}$, therefore

$$\|\Pi_2(A,p,q)\| \le \sigma^{-1}(m\mu^{-1})^{1/2} m.$$

<u>Assertion 9</u> For any nondegenerate normalized primal linear programming problem $P(A,p,q)$ the next estimate

$$\mathrm{cond}P(A,p,q) \le \sigma^{-1}(m\mu^{-1})^{1/2}\max\{n-m, m\}$$

holds true.

Now we will introduce and estimate the condition number of the dual problem $D(A, p, q)$.

Let us consider the matrix

$$\Psi(A, p, q) = \Psi(A, p, q, \infty) = \begin{bmatrix} 0^{m, n-m} & -(Y^*_B B^T)^{-1} \\ -B^{-1} N U^{*-1}_N & 0^{m, m} \end{bmatrix}.$$

The value

$$\text{cond } D(A, p, q) = \|\Psi(A, p, q)\|$$

we will call *condition number of the dual LP problem D(A, p, q)*.

Before we estimate the cond $D(A, p, q)$ let us consider the scaling issue for the dual problem. We will say that the dual problem $D(A, p, q)$ is scaled by a vector $s = (s_1, \dots, s_n) > 0^n$ if instead of the matrix $A = (N, B)$ and vector p we consider the matrix $\overline{A} = AS = (\overline{N}, \overline{B}) = (NS_N, BS_B)$, where $S = [\text{diag } s_i]_{i=1}^n$, $S_N = [\text{diag } s_i]_{i=1}^{n-m}$, $S_B = [\text{diag } s_i]_{i=n-m+1}^n$ and vector $\overline{p} = pS = (\overline{p}_N, \overline{p}_B) = (p_N S_N, p_B S_B)$. First of all note that $\overline{\text{v}} = \overline{p}_B \overline{B}^{-1} = p_B S_B (BS_B)^{-1} = p_B S_B (BS_B)^{-1} = p_B S_B S_B^{-1} B^{-1} = p_B B^{-1} = \text{v}^*$.

Therefore the solution of the dual problem $D(A, p, q)$ is an invariant with respect to scaling the dual problem, i.e.

$$\text{v}^* = \overline{\text{v}} = \text{argmax}\{(q, \text{v}) \mid \overline{r}(\text{v}) = (\overline{p} - \overline{\text{v}} \overline{A})_i \geq 0, i = 1, \dots, n\}.$$

Matrices $Y^*_B B^T$ and NU^{*-1}_N are also invariant with respect to scaling the dual problem. In fact, we have $\overline{y}_B = (\overline{y}_{n-m+1}, \dots, \overline{y}_n) = \overline{B}^{-1} q = (BS_B)^{-1} q = S_B^{-1} B^{-1} q = S_B^{-1} y^*_B$. So $\overline{Y}_B = [\text{diag} \overline{y}_i]_{i=n-m+1}^n = S_B^{-1} Y^*_B = Y^*_B S_B^{-1}$ and $\overline{Y}_B \overline{B}^T = Y^*_B S_B^{-1} \times S_B B^T = Y^*_B B^T$. Also $r(\text{v}) = (p - \text{v}A) = (p_N - \text{v}N, p_B - \text{v}B) = (u_N, u_B)$, so $\overline{U}_N = [\text{diag}(\overline{p} - \overline{\text{v}}\overline{A})_j]_{j=1}^{n-m} = [\text{diag}(\overline{p} - \text{v}^* \overline{A})_j]_{j=1}^{n-m} = [\text{diag}((p - \text{v}^* A) S_N)_j]_{j=1}^{n-m} = U^*_N S_N$.

Therefore $\overline{N}\,\overline{U}^{-1} = N S_N (U^*_N S_N)^{-1} = N U^{*-1}_N$. Now we are ready to estimate the norm $\|\Psi(A,p,q)\|$.

We will assume that the dual problem $D(A,p,q)$ is normalized, i.e. is scaled by the vector $s = (s_1, \dots s_i, \dots s_n)$, $s_i = \|a_i\|^{-1} = (\max_{1 \le j \le m} |a_{ji}|)^{-1}$, i.e. we assume that $\|a_i\| = 1$, $i = 1, \dots, n$. Then

$$\|\Psi(A,p,q)\| = \max\{\|(Y^*_B B^T)^{-1}\|, \|B^{-1} N U^{*-1}_N\|\} \le \sigma^{-1} \max\{\|B^{-T}\|, \|B^{-1} N\|\}.$$

Taking into account $\|B^{-T}\| \le \sqrt{m}\,\|B^{-T}\|_2 = (m\mu^{-1})^{1/2}$, and $\|N\| = \max_{1 \le j \le m} |\sum_{j=1}^{n-m} a_{ji}| \le (n-m)$, we obtain

$$\|\Psi(A,p,q)\| \le \sigma^{-1}(m\mu^{-1})^{1/2}(n-m).$$

Assertion 10. For any nondegenerage normalized dual linear programming problem the next estimate

$$\operatorname{cond} D(A,p,q) \le \sigma^{-1}(m\mu^{-1})^{1/2}(n-m)$$

holds true.

From the proof of Theorem 1 one can see that the estimate of the norm $\|\Pi(A,p,q,k)\|$ is independent on $k \ge \operatorname{cond} P(A,p,q)$. The constant c in Theorem 1 can be estimated by $O(\operatorname{cond} P(A,p,q))$. Therefore, for a fixed $0 < \gamma \le 0.5$ one can find $k_0 = O(\gamma^{-1} \operatorname{cond} P(A,p,q)$ that for any fixed $k \ge k_0$ Assertion 1 holds true.

The same is true for the dual problem $D(A,p,q)$. The norm $\|\Psi(A,p,q,k)\|$ is independent on $k > \operatorname{cond} D(A,p,q)$, the constant c in Theorem 2 can be estimated by $O(\operatorname{cond} D(A,p,q))$ and for any fixed $0 < \gamma \le 0.5$ one can find $k_0 = O(\gamma^{-1} \operatorname{cond} D(A,p,q))$ that for any fixed $k \ge k_0$ Assertion 6 takes place.

Moreover cond $P(A, p, q)$ is crucial for the "hot start" of the primal Newton MBF (see Assertion 4) while cond $D(A, p, q)$ is crucial for the "hot start" of the Dual Newton MBF (see Assertion 8).

In order to reach the "hot start", one has to perform $0(\sqrt{n}\,\gamma^{-1}\, ln \operatorname{cond} P(A, p, q))$ Newton steps by minimizing Primal $M(x, k)$ or $0(\sqrt{n}\,\gamma^{-1}\, ln \operatorname{cond} D(A, p, q))$ by minimizing Dual $N(\mathrm{v}, k)$ Shifted Barrier Functions, starting at any "warm start" (see [7]) and increasing k by a factor $(1 - \alpha n^{1/2})^{-1}$ after every Newton step, where α is a universal constant (see [12] , [27]).

From the "hot start" on one can decrease $0(\sqrt{n})$ to $0(ln\, n)$ the number of Newton steps required to reduce the gap between primal and dual objective functions by a fixed factor. In the next section we are going to make a few comments on this matter.

10. Concluding Remarks

The difference between Classical Interior Point Methods that are based on CBF and Newton MBF methods result from the difference between primal $\{x(k), u(k)\}$:

$$x(k) = \operatorname{argmin} \{(p, x) - k^{-1} \sum_{j=1}^{n} ln\, x_j \mid x \in \mathbf{R}_+^n \cap Q\}$$

$$\mathrm{v}(k) = k^{-1} [\operatorname{diag} x_j^{-1}(k)]e, \quad k \to \infty$$

or dual $\{\mathrm{v}(k), y(k)\}$:

$$\mathrm{v}(k) = \operatorname{argmax} \{(q, \mathrm{v}) + k^{-1} \sum \ln\, r_j(\mathrm{v}) \mid \mathrm{v} \in \mathbf{R}^m\}$$

$$y(k) = k^{-1} [\operatorname{diag} r_j^{-1}(\mathrm{v}(k))]e, \; k \to \infty.$$

CBF trajectories and primal $\{x(t,k),\hat{u}(t,k)\}$ or dual $\{v(t,k),y(t,k)\}(\|t\|\to 0$ and $k=0(\mathrm{cond}P(A,p,q))$ or $k=0(\mathrm{cond}D(A,p,q))$ are fixed) MBF trajectories.

The Classical Interior Point Methods follow along the CBF trajectory turning from one "warm start" to another "warm start" by performing one Newton step and updating the penalty parameter.

Every Newton step improves the current objective function value by a factor $(1-\alpha n^{-0.5})$ where α is a universal constant (in [27] the corresponding result was proved for $\alpha=41^{-1}$).* So to improve the current approximation twice, one has to perform $0(\sqrt{n})$ Newton steps. The primal or dual Newton MBF methods follow from some point along the MBF trajectory changing Lagrange multipliers instead of the parameter $k>0$.

Having a couple $(\tilde{x}^s,\tilde{u}^s)$ that is *well defined* i.e. $\tilde{x}^s$ is in the Newton area ("warm start") (see [29]) for the problem

$$\hat{x}(\tilde{u}^s,k)=\mathrm{argmin}\{F(x,\tilde{u}^s,k)\mid x\in\Omega_k\cap Q\}$$

and $\tilde{u}^s\in U_k$, $k\geq k_0=0(\gamma^{-1}\mathrm{cond}P(A,p,q))$ one can obtain an approximation $\tilde{x}^{s+1}$ for the vector $x(\tilde{u}^s,k)$ with accuracy $\varepsilon>0$ in $0(ln\, ln\, \varepsilon^{-1})$ Newton steps.

For $\varepsilon>0$ small enough we can maintain estimate (37) for a fixed $0<\gamma\leq 0.5$ and a triple $(\tilde{x}^{s+1},\tilde{u}^{s+1},\tilde{v}^{s+1})$, where $\tilde{u}^{s+1}=\Delta^{-1}(\tilde{x}^{s+1},k)\tilde{u}^s$ and $\tilde{v}^{s+1}=(p-\tilde{u}^{s+1})A^T(AA^T)^{-1}$.

Moreover as it turns out for any nondegenerate dual pair LP and any fixed $0<\gamma\leq 0.5$, there exists such s_0 and $k_0=0(\gamma^{-1}\mathrm{cond}\,P(A,p,q))$ or

*) Recently S. Smale and M. Shub proved that such a result remains true with $\alpha=13^{-1}$.

$k_0 = O(\gamma^{-1}\text{cond } D(A,p,q))$ that for any $s \geq s_0$ and any $k \geq k_0$ the existence of a well defined couple $(\tilde{x}^s, \tilde{u}^s)$ implies that the couple $(\tilde{x}^{s+1}, \tilde{u}^{s+1})$ is also well defined. It means that $\tilde{x}^{s+1}$ is well defined for the problem

$$\hat{x}(\tilde{u}^{s+1}, k) = \text{argmin } \{F(x, \tilde{u}^{s+1}, k) \mid x \in \Omega_k \cap Q\}$$

and $\tilde{u}^{s+1} \in U_k$.

In other words, starting from $(\tilde{x}^{s_0}, \tilde{u}^{s_0})$ ("hot start"), one can improve the current approximation at least twice ($\gamma \leq 0.5$) in every $O(\ln \ln \varepsilon^{-1})$ Newton steps. The "hot start" *phenomenon* was made possible because of the MBF's properties at primal and dual solution as well as in their neighborhoods, where MBF's not only exist and are smooth enough, but have stable condition numbers for their Hessians.

To reach the "hot start" one can use CBF or one-parameter Shifted Barrier Functions starting from any "warm start" and a fixed $k > 0$. Due to the self-concordant properties of CBF or Shifted Barrier Functions (see [20]) one can reach the "hot start" in $O(\sqrt{n}\, \ln(\gamma^{-1}\text{cond } P(A,p,q)))$ or $O(\sqrt{n}\, \ln(\gamma^{-1}\text{cond } D(A,p,q))$ Newton steps by updating at every step the penalty parameter k in a way that we described above.

To estimate the number of Newton steps that one has to perform, beginning at the "hot start" up to an approximation for primal and dual solution with accuracy $\varepsilon = 2^{-L}$ ($L-$ is the input length), we assume that $P(A,p,q)$ or $D(A,p,q)$ are normalized and $\max_{i,j} |a_{ij}| = 1$, $\max_{i,j}\{|p_j|, |q_i|\} \leq 2^l$, $\min_{i,j}\{|a_{ij}|, |p_j|, |q_i|\} \geq 2^{-l} (l \ll m < n)$.
So the input length L can be estimated by $O(n^3)$. Using $\varepsilon = 2^{-2L}$ in Primal or Dual Newton MBF methods, we obtain that, beginning at the "hot start", every Lagrange multipliers update requires $O(\ln L)$ Newton steps and allows us to improve the current approximation at least twice $(0 < \gamma \leq 0.5)$ for a fixed $k > k_0$, where $\ln k_0 = O(\ln \gamma^{-1} \text{ cond } P(A,p,q))$ or $\ln k_0 = O(\ln \gamma^{-1}\text{cond } D(A,p,q,))$.

In other words in $O(\ln n)$ Newton steps one can improve the current approximation at least twice. This means that beginning at the "hot start", one has to perform $O((L - O(\ln \gamma^{-1} \text{cond } P(A, p, q)) \bullet \ln n)$ Newton steps to obtain an approximation for (x^*, v^*) with accuracy $\varepsilon = 2^{-L}$ in case of Primal Newton MBF and $O((L - O(\ln \gamma^{-1} \text{cond } D(A, p, q)) \ln n)$ in the case of Dual Newton MBF. So the total number of Newton steps can be estimated by

$$N \leq O(\sqrt{n} \ln \gamma^{-1} \text{cond } P(A, p, q))) + O((L - \ln \gamma^{-1} \text{cond } P(A, p, q)) \ln n).$$

for the Primal Newton MBF. We have the same type of estimation for the Dual Newton MBF. In the case where the primal or dual condition number is $\gamma 2^L$, we obtain the classical estimation $N \leq O(\sqrt{n} L)$ (see [7], [10], [12], [27], and [30]).

The analysis that was undertaken in this paper shows the principal difference between CBF and MBF approach as well as the possibility to speed up essentially the process of solution dual LP in the final stage by changing the Lagrange multipliers instead of the penalty parameter.

Of course it is possible to use some other procedures to speed up the solution process in the final stage. In particular, the "purification" procedure (see Remark 5) might be effective in the case of nondegenerate dual pair LP.

However, we would like to emphasize that the main results obtained in this paper remain true if only one of the dual pair LP has a unique solution.

Moreover, the implementation of the MBF approach for solving LP that was recently accomplished at the IBM T.J. Watson Research Center, allows us to observe the "hot start" phenomenon practically for all LP (more than 80) that were solved.

The MBF theory also allowed us to develop a primal-dual MBF method that stands to the Primal Newton MBF as Primal-Dual Predictor-Corrector (see [77]) to the Projected Newton Method [9].

The practical aspects of the Newton MBF methods will be considered in a forthcoming paper where the Newton MBF methods will be compared with CBF as well as with primal-dual predictor-corrector approach (see [17]).

It gives me great pleasure to dedicate this paper to Professor J. Ben Rozen on the occasion of his seventieth birthday.

Acknowledgements: It is a great pleasure to express my gratitude to my colleagues Drs. David Jensen and Rina Schneur for many fruitful discussions which led to the improvement of this paper and for their efforts that enabled a much better understanding of the practical aspects of the Newton MBF methods.

I also appreciate very much Eileen Count for her excellent work in preparing the manuscript for publication.

References

1. Bayer, D.A., Lagarias, J.C., "Karmarkar's Linear Programming Algorithm and Newton Method", *Mathematical Programming* 50, N3, (1991) 291-330.
2. Bertsekas, D.P., "Necessary and Sufficient Conditions for a Penalty Method to be Exact", *Mathematical Programming*, Vol. 9, 1975, pp. 87-99.
3. Bertsekas, D., *Constrained Optimization and Lagrange Multiplier Methods*, New York, Academic Press (1982).
4. Carroll, C.W., "The Created Response Surface Technique for Optimizing Nonlinear Restrained Systems," *Operations Research* 9, (1961), pp. 169-184.
5. Fiacco, A.V., and McCormick, G.P., *Nonlinear Programming, Sequential Unconstrained Minimization Techniques*, New York, Wiley, (1968).
6. Fletcher, R., "Recent Developments in Linear and Quadratic Programming", University of Dundee, Numerical Analysis, Report NA/94, (1986).
7. Freund, R., "Theoretical Efficiency of a Shifted Barrier Function Algorithm for Linear Programming", *Operations Research*, pp. 194-89.
8. Frisch, K.R., "The Logarithmic Potential Method of Convex Programming, University Institute of Economics", Oslo, Memorandum of May 13, 1955.
9. Gill, P.E., Murray, M., Saunders, M.A., Tomlin, J., and Wright, M.H., "On projected Newton Barrier Methods for Linear Programming and an Equivalence to Karmarkar's Projective Method", *Mathematical Programming* 36, (1986), pp. 183-209.
10. Goldfarb, D., Xiao, D., "A Primal Projective Interior Point Method for Linear Programming", *Mathematical Programming* 51 (1991), pp. 17-43.
11. Golshtein, E.G., and Tret'ykov, N.V., "Modified Lagrangian Functions, Theory and Optimization Methods", Moscow, Nauka (1989) (In Russian).

12. Gonzaga, C.C., "An Algorithm for solving Linear Programming Problems in $0(n^3L)$ Operations", In N. Megiddo ed. *Progress in Mathematical Programming: Interior Point and Related Methods*, Springer, New York, (1988).

13. Grossman, K., Kaplan, A., *Nonlinear Programming Based on Unconstrained Minimization*, Novosibirski, Navka (1981) (In Russian).

14. Karmarkar, N.A., "A New Polynomial-Time Algorithm for Linear Programming", *Combinatorica* 4, 373-395, (1984).

15. Mangasarian, O.L., "Unconstrained Lagrangians in Nonlinear Programming", *SIAM J. Control*, Vol. 13, New York, (1975) pp. 772-971.

16. Mangasarian, O.L., A Condition Number for Linear Inequalities and Linear Program Methods of Operations Research, Proc. 6th Symposium Operations Research, Augsburg, September 7-9, 1981, Bamberg, G., Opitz, O., eds., Verlagsgruppe Athenaum/Hain/Scriptor/Hanstein, Konigstein, pp. 3-15.

17. Marsten, R., Subramanian, R., Saltzman, M., Lustig, I., and Shanno, D., "Interior Point Methods for Linear Programming: Just Call Newton, Lagrange, and Fiacco and McCormick!", *Interfaces* 20:4, July-August (1990) pp. 105-116.

18. Megiddo, N., Shub, M., "Boundary Behavior of Interior Point Algorithm for Linear Programming", IBM Research Report RC5319, IBM T.J. Watson Research Center, Yorktown Heights, New York (1986).

19. Monteiro, R.D.C., Adler, I., "Interior Path Following Primal-Dual Algorithms, Part 1: Linear Programming", *Mathematical Programming* 44, (1989) pp. 27-42.

20. Nesterov, Yu, E., Nemirovsky, A.S., "Self Concordant Functions and Polynomial Time Methods in Convex Programming", Moscow (1989).

21. Polyak, B.T., Tretjakov, N.V., "The Method of Penalty Estimates for Conditional Extremum Problems", Z. Vychisl. Math. i Mat. Fiz. 13, 1973, pp. 34-46.

22. Polyak, B.T., Tretjakov, N.V., "An Iterative Method for Linear Programming and its Economic Interpretation", *Matecon*, Vol. 10, (1974), pp. 69-81.

23. Polyak, R.A., Controlled Processes in Extremal and Equilibriium Problems. Deposited manuscript, Moscow: VINITI (1986) (in Russian).

24. Polyak, R.A., "Modified Barrier Functions, (Theory and Methods)", *Mathematical Programming* 54, (1992), No. 2, pp. 177-222.

25. Polyak, R.A., "Modified Interior Distance Functions and Center Methods", Research Report RC 17042 (#75556), Math. Sciences Department, IBM T.J. Watson Research Center, Yorktown Heights, New York (1991).

26. Powell, M.J.D., "A Method for Nonlinear Constraints in Minimization Problems", In: *Optimization*, New York, Academic Press, (1969), pp. 283-298.

27. Renegar, J., and Shub, M., "Unified Complexity Analysis for Newton LP Methods", *Mathematical Programming* 53, (1992), N1, pp. 1-16.

28. Rockafellar, R.T., "Augmented Lagrange Multiplier Functions and Duality in Nonconvex Programming", *SIAM J. Control* 12, No. 2, 268-285 (1974).

29. Smale, S., "Newton's Method Estimates from Data at One Point, The Merging of Disciplines in Pure, Applied, and Computational Mathematics" (R.E. Ewing, et al, ed.), Springer-Verlag, New York-Berlin (1986), pp. 185-196.

30. Todd, M.J., Ye, Y., "A Centered Projective Algorithm for Linear Programming", Technical Report No. 763, School of Operations Research and Industrial Engineering, Ithaca, New York, (1987).

31. Todd, M.J., "Recent Developments and New Directions in Linear Programming", Technical Report No. 827, School of Operations Research and Industrial Engineering, Ithaca, New York, Nov. 1988.