Optimal Face Identification Methods and Bounded Variable Linear Programs

Pamela J. Williams, Amr S. El-Bakry, and Richard A. Tapia

CRPC-TR98755-S
August 1998

Center for Research on Parallel Computation
Rice University
6100 South Main Street
CRPC - MS 41
Houston, TX 77005

Submitted August 1998; Also available as Rice CAAM TR 98-16

OPTIMAL FACE IDENTIFICATION METHODS AND BOUNDED VARIABLE LINEAR PROGRAMS

PAMELA J. WILLIAMS AMR S. EL-BAKRY RICHARD A. TAPIA

Abstract. We extend optimal face identification methods to bounded variable linear programming problems. Distance to the lower and upper bounds are incorporated into a projection model and Mehrotra–Ye's solution technique to prevent the computed solution from violating the bound constraints. Empirical and theoretical evidence are provided that support use of the new models to compute an exact solution. We also introduce a nonlinear weighted projection method to solve the optimal face identification problem.

Key words. active set, finite termination, infeasible primal-dual interior-point algorithms, linear programming, optimal face, optimal partition, polynomiality

AMS subject classifications. 90C05, 49M35, 65K05

1. Introduction.

Finite termination procedures are methods inserted into iterative algorithms to compute an exact solution in a finite number of steps. In particular, finite termination procedures can be added to interior-point methods to advance from an approximate solution to an exact solution.

Optimal face identification methods identify the face upon which the objective function attains its optimal value. The optimal face is uniquely defined by the active set, the set of variables which are zero at the solution. Once the active set has been identified, the exact solution of the linear program can be obtained by computing an interior feasible point on the face. For a survey of optimal face identification methods; see Williams, El-Bakry, and Tapia [52].

Adding optimal face identification methods to the interior-point framework can lead to computational savings and highly accurate solutions. Moreover, a point on the optimal face can be used to generate an optimal basic solution in strongly polynomial time; see for example Megiddo [30], Andersen [2], and Andersen and Ye [3]. Knowledge of the optimal face in sensitivity analysis in the context of interior-point methods was assumed by Adler and Monteiro [1]; Monteiro and Mehrotra [38]; Jansen, Roos, and Terlaky [19, 20]; Jansen, Roos, Terlaky; and Vial [21]; and Greenberg [16, 17].

In this work, we extend finite termination procedures based on the optimal face identification methods of Ye [57, 58, 59] and Mehrotra and Ye [37] to linear programs with upper bound constraints.

1.1. Background.

We consider linear programs of the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & l \le x \le u \end{array} \tag{1}$$

where $c, x \in \mathbf{R}^n$, $b \in \mathbf{R}^m$, $A \in \mathbf{R}^{m\times n}$ $(m \le n)$ and A has full rank m. The vector $l \in \mathbf{R}^n$ represents the vector of lower bounds and $u \in \mathbf{R}^n$ represents the vector of upper bounds for the vector x. Without loss of generality, we assume all the variables have lower bounds of zero and finite upper bounds. Problem (1) rewritten in standard form is

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x + s = u \\ & x, s \ge 0 \end{array} \tag{2}$$

where $s \in \mathbf{R}^n$ is the primal slack vector.

The corresponding dual problem is

$$\begin{array}{lll} & \text{maximize} & b^T y - u^T w \\ (3) & \text{subject to} & A^T y + z - w = c \\ & & z, w \geq 0 \end{array}$$

where $y \in \mathbf{R}^m$ are the Lagrange multipliers corresponding to the equality constraints, $z \in \mathbf{R}^n$ are the Lagrange multipliers corresponding to the inequality constraints, $w \in \mathbf{R}^n$ are the Lagrange multipliers corresponding to the upper bound constraints.

The optimality conditions for (1) are

$$F(x, y, z, s, w) = \begin{pmatrix} Ax - b \\ x + s - u \\ A^T y + z - w - c \\ XZe \\ SWe \end{pmatrix} = 0, \quad (x, z, s, w) \geq 0, \tag{4}$$

where $X = \text{diag}(x)$, $Z = \text{diag}(z)$, $S = \text{diag}(s)$, $W = \text{diag}(w)$, and e is the n-vector of all ones.

The Jacobian of (4) is

$$F'(x, y, z, s, w) = \begin{pmatrix} A & 0 & 0 & 0 & 0 \\ I & 0 & 0 & I & 0 \\ 0 & A^T & I & -I & 0 \\ Z & 0 & X & 0 & 0 \\ 0 & 0 & 0 & W & S \end{pmatrix}. \tag{5}$$

The vectors x and s are feasible for the primal if $Ax = b, x + s = u$, and (x, s) is nonnegative. We say that w and z are feasible for the dual if there exists y such that (y, z, w) is feasible for (3). A point (x, y, z, s, w) is said to be strictly feasible if it satisfies $Ax = b$, $x + s = u$, $A^T y + z - w = c$, and $(x, s, z, w) > 0$.

We denote the solution set of (4) as

$$\mathcal{S} = \{(x, y, z, s, w) : F(x, y, z, s, w) = 0, \ \ (x, z, s, w) \geq 0\}.$$

If a solution satisfies $x + z > 0$ and $s + w > 0$, in addition to $XZe = 0$ and $SWe = 0$, then this solution is said to satisfy the strict complementarity condition or strict complementarity. Given feasible iterates, we see that $||F(x, y, z, s, w)||_1 = x^T z + s^T w$. It can be shown that the expression $x^T z + s^T w$ is equal to the duality gap, which vanishes at any solution.

1.2. Algorithmic Framework. Kojima, Mizuno, and Yoshise [25] introduced the primal-dual interior-point method in linear programming. It is well-known that their method can be viewed as perturbed and damped Newton's method on the first order optimality conditions. In this section, we describe an infeasible primal-dual Newton interior-point method.

ALGORITHM 1 (Infeasible Primal-Dual Interior-Point Algorithm).
Given $v^0 = (x^0, y,^0, z^0, s^0, w^0)$ *with* $(x^0, z^0, s^0, w^0) > 0$, *for* $k = 0, 1, \ldots,$ *do*
1. *Choose* $\sigma^k \in (0, 1)$ *and set* $\mu^k = ((x^k)^T z^k + (s^k)^T w^k)/2n$.
2. *Solve for the step* Δv^k

$$F'(v^k)\Delta v^k = -F(v^k) + \mu\tilde{e} \ \textit{where}\ \mu \geq 0$$

3. *Choose* $\tau^k \in (0,1)$ *and set* $\alpha^k = min(1, \tau^k \hat{\alpha}^k)$, *where*

$$\hat{\alpha}^k = \frac{-1}{\min((X^k)^{-1}\Delta x^k, (Z^k)^{-1}\Delta z^k, (S^k)^{-1}\Delta s^k, (W^k)^{-1}\Delta w^k)}.$$

4. *Let* $v^{k+1} := v^k + \alpha^k \Delta v^k$.

5. *Test for convergence.*

In Step 2, $\tilde{e} = (0,\ldots,0,1,\ldots,1,1,\ldots,1)^T$ with $2n+m$ zero components. The optimality conditions (4) are perturbed so that the Newton direction obtained from the perturbed KKT conditions does not point towards the boundary. If $\sigma^k = 0$ (i.e., no perturbation), global convergence may be precluded. See Proposition 3.1 of Gonzalez-Lima [26] for a proof. The iteration sequence is damped to maintain the nonnegativity requirement.

1.3. Notation. For notational convenience in the statements and proofs of the theory, we introduce the following notation:

$$\tilde{x} = (x, s) \in \mathbf{R}^{2n}$$
$$\tilde{z} = (z, w) \in \mathbf{R}^{2n}.$$

In all other instances we will refer to the point (x, y, z, s, w).

If $\mathcal{S} \neq \emptyset$, then the relative interior of $\mathcal{S}$, $ri(\mathcal{S})$, is nonempty. In this case, the solution set $\mathcal{S}$ has the following structure (see El-Bakry, Tapia, and Zhang [9] for a proof): (i) all points in the relative interior satisfy strict complementarity; (ii) the zero-nonzero pattern of points in the relative interior is invariant. for any $(\tilde{x}^*, y^*, \tilde{z}^*)$ in the relative interior of the solution set of (1), we define the index sets $\tilde{\mathcal{B}}$ and $\tilde{\mathcal{N}}$ as

$$\tilde{\mathcal{B}} = \{j : \tilde{x}_j^* > 0,\ 1 \leq j \leq 2n\} \text{ and } \tilde{\mathcal{N}} = \{j : \tilde{x}_j^* = 0,\ 1 \leq j \leq 2n\}.$$

For more details, see Güler and Ye [18] and McLinden [28]. Moreover,

$$\tilde{\mathcal{B}} \bigcup \tilde{\mathcal{N}} = \{1, \ldots, 2n\} \quad \text{and} \quad \tilde{\mathcal{B}} \bigcap \tilde{\mathcal{N}} = \emptyset.$$

Thus the sets $\tilde{\mathcal{B}}$ and $\tilde{\mathcal{N}}$ define the optimal partition of the set $\{1, 2, \ldots, 2n\}$.

We define the index sets $\mathcal{B}$ and $\mathcal{N}$ as

$$\mathcal{B} = \{j : x_j^* > 0,\ 1 \leq j \leq n\} \text{ and } \mathcal{N} = \{j : x_j^* = 0,\ 1 \leq j \leq n\}.$$

In the following sections, the columns of A corresponding to the indices of $\mathcal{B}$ comprise the matrix $A_{\mathcal{B}}$. The matrix $A_{\mathcal{N}}$ is formed in an analogous manner. We represent the components of the vector x whose indices are in $\mathcal{B}$ by $x_{\mathcal{B}}$. Unless otherwise specified, $||\cdot||$ is the Euclidean norm. The cardinality of set $\mathcal{B}$ is denoted by $|\mathcal{B}|$. We use the notation

$$\min u = \min_{1 \leq i \leq n} u_i \text{ for } u \in \mathbf{R}^n.$$

The inequality $x \geq 0$ denotes component-wise nonnegativity.

The paper is organized as follows. We provide an historical overview of the optimal face identification problem in Section 2. Section 3 contains technical results. Section 4 deals with mathematical models and solution techniques to solve the optimal face identification problem. Sections 4.1 and 4.2 are devoted to methods specifically designed for linear programs with upper bound constraints. We offer computational results in Section 5. Section 6 contains concluding remarks.

2. Historical Overview. In 1989, stopping tests to compute optimal solutions in interior-point methods for linear programming were proposed by Gay [13]. While these tests did not constitute a finite termination procedure because the primal and dual optimality checks were iterative methods, they were clearly predecessors of current optimal face identification methods. Ye [57] popularized the study of finite termination in interior-point methods for linear programming. He was motivated by the fact that the simplex method for linear programming has the finite termination property and also by research activity in efficient algorithmic termination techniques. Ye [57] established a theoretical base for Gay's tests when added to primal-dual interior-point algorithms which generate iteration sequences that converge to strict complementarity solutions. Mehrotra and Ye [37] developed a solution technique based on Gaussian elimination to compute an interior feasible point on the optimal primal and dual faces. Previously Tardos [49] used Gaussian elimination to calculate a feasible point on the optimal face of an integer program. More recently, Ye [59] proposed an optimal face identification model, which incorporated bound information, for the standard linear program with nonnegativity constraints. Ye [59] proved for k sufficiently large the solutions of his proposed optimal face identification methods can be obtained in finite time when included in a feasible primal-dual interior-point method.

The theoretical aspects of finite termination procedures in infeasible interior-point methods for linear programming have been studied by Potra [42], Williams [51], and Anstreicher, Ji, Potra, and Ye in [4], where a probabilistic analysis was given. In particular, Williams [51] extended the analysis of Ye's weighted projection models to infeasible interior-point algorithms.

Optimal face identifications methods have not been restricted to linear programming. Researchers have considered optimal face identification in linear complementarity problems and network flow problems. Monteiro and Wright [39], as well as Ji and Potra [22] investigated finite termination procedures in infeasible interior-point algorithms for degenerate monotone linear complementarity problems (LCPs). Resende and Veiga [43] identified the optimal dual face and generated a primal basic solution to derive robust stopping criteria for minimum cost network flow problems. Subsequent research into identifying the optimal dual face for network flow problems appeared in Portugal, Resende, Veiga, and Judice [41] and Resende, Tsuchiya, and Veiga [44].

The optimal primal face of (1) can be written as

$$\Theta_p = \{\tilde{x} : Ax = b,\ x + s = u,\ \tilde{x} \geq 0,\ \tilde{x}_j = 0\ j \in \tilde{\mathcal{N}}\}.$$

and the optimal dual face as

$$\Theta_d = \{(y, \tilde{z}) : A^T y + z - w = c,\ \tilde{z} \geq 0,\ \tilde{z}_j = 0\ j \in \tilde{\mathcal{B}}\}.$$

If the set of strictly feasible points is nonempty, then $\tilde{\Theta}_p$ and $\tilde{\Theta}_d$ are nonempty and bounded. Hence, $\tilde{\xi}_p^*$ and $\tilde{\xi}_d^*$ are bounded, where

$$\begin{aligned}
\tilde{\xi}_p^* &= \min_{j \in \tilde{\mathcal{B}}} \{\max \tilde{x}_j,\ \text{s.t.}\ \tilde{x} \in \Theta_p\};\\
\tilde{\xi}_d^* &= \min_{j \in \tilde{\mathcal{N}}} \{\max \tilde{z}_j,\ \text{s.t.}\ (y, \tilde{z}) \in \Theta_d\};\\
\tilde{\xi}^* &= \min(\tilde{\xi}_p^*, \tilde{\xi}_d^*).
\end{aligned}$$

An ideal property of an optimal face identification methods is that it is computationally inexpensive. The cost of computing a point on the optimal faces should be the same as an interior-point iteration. This property plays a pivotal role in our optimal dual face identification models.

3. Technical Results. The following lemma provides a theoretical basis for finite termination procedures for linear programs of the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \quad x \geq 0, \end{array} \tag{6}$$

where $c, x \in \mathbf{R}^n$, $b \in \mathbf{R}^m$, $A \in \mathbf{R}^{m \times n}$ $(m \leq n)$ and A has full rank m.

LEMMA 3.1. *(Güler-Ye [18]) Let $\{(x^k, y^k, z^k)\}$ be an iteration sequence generated by an interior-point algorithm. Furthermore, let x^k and z^k satisfy*

$$\frac{min(X^k Z^k e)}{(x^k)^T z^k / n} \geq \gamma \tag{7}$$

where $\gamma > 0$ and is independent of k. Then every limit point of $\{(x^k, z^k)\}$ satisfies the strict complementarity condition.

Güler-Ye [18] showed that iterates of feasible path-following algorithms satisfy Inequality (7), which is one of many central path proximity measures used in linear programming.

All points in the relative interior of the solution set satisfy the strict complementarity condition. Therefore, Lemma 3.1 is sufficient to guarantee that all limit points of the iteration sequence are in the relative interior of the solution set; see Güler and Ye [18]. It is well-known that in the relative interior the nonzero-zero pattern of points is invariant; see El-Bakry, Tapia, and Zhang [9]. Consequently, the optimal primal and dual faces are uniquely defined.

It is easy to see that an $l_{-\infty}$ central path neighborhood that includes infeasible points for problem (1) can be defined as

$$\begin{array}{c} \mathcal{N}_{-\infty}(\gamma, \beta) = \{(\tilde{x}, y, \tilde{z}) | \quad \|(r_p, r_u, r_d)\| \leq [\|(r_p^0, r_u^0, r_d^0)\| / \mu^0] \beta \mu, \\ (\tilde{x}, \tilde{z}) > 0, \quad \min(\tilde{X}\tilde{Z}e) \geq \gamma \mu, \quad \mu = (\tilde{x}^T \tilde{z})/2n\} \end{array} \tag{8}$$

where $\gamma \in (0,1)$, $\beta \geq 1$, $r_p = b - Ax$, $r_u = u - x - s$, and $r_d = c - A^T y - z + w$. The first inequality of (8) known as the Feasibility Priority Principle, requires that infeasibility decreases at least as fast as complementarity; see Zhang [60].

Theorem 3.2 shows that the optimal partition can be identified once the duality gap is sufficiently small.

THEOREM 3.2. *Let $(\tilde{x}^*, y^*, \tilde{z}^*) \in \mathcal{S}$, $\tilde{\mathcal{B}}^k = \{\tilde{x}_j^k > \tilde{z}_j^k\}$, and $\{(\tilde{x}^k, y^k, \tilde{z}^k)\}$ be the iteration sequence generated by Algorithm 1. Assume further that*

$$(\tilde{x}^k)^T \tilde{z}^k < \gamma (\tilde{\xi}^*)^2 / 2\tilde{\tau}^2 n,$$

for some constant,

$$\tilde{\tau} = \beta \left[\frac{2 - \alpha^0}{\alpha^0} + \frac{(\tilde{x}^0)^T \tilde{z}^* + (\tilde{z}^0)^T \tilde{x}^*}{(\tilde{x}^0)^T \tilde{z}^0} \right] > 0, \text{ then } \tilde{\mathcal{B}}^k = \tilde{\mathcal{B}}.$$

Proof. If we let $\beta = 1$, the proof follows directly from Lemma 4.1 and Proposition 5.2 of Potra [42] with $\tilde{x}$, $\tilde{z}$ replacing x and z. □

Subsequent theory requires bounds on the primal and dual residuals. Rather than repeatedly deriving the bounds, we establish them at this juncture. Note that the bound depends on the current duality gap, the initial duality gap, and the minimum positive value in the solution set, which is not known a priori.

LEMMA 3.3. *(Williams [51]) Let $\{(\tilde{x}^k, y^k, \tilde{z}^k)\}$ be the iteration sequence generated by Algorithm 1. Further, assume that $\tilde{\mathcal{B}}^k = \tilde{\mathcal{B}}$. Then,*

$$\|b - A_{\mathcal{B}} x^k_{\mathcal{B}}\| \leq (2n\mu^0 \tilde{\xi}^*)^{-1} \left(\tilde{\xi}^* \beta \|r^0\| + 2n\sqrt{2n}\mu^0 \tilde{\tau} \|A_{\mathcal{N}}\|\right) \left((\tilde{x}^k)^T \tilde{z}^k\right)$$
$$\|c_{\mathcal{B}} - A^T_{\mathcal{B}} y^k\| \leq (2n\mu^0 \tilde{\xi}^*)^{-1} (\tilde{\xi}^* \beta \|r^0\| + 4n\sqrt{2n}\mu^0 \tilde{\tau}) \left((\tilde{x}^k)^T \tilde{z}^k\right).$$

Proof. The proof uses the Feasibility Priority Principle, the bounds on Θ_p and Θ_d, and Theorem 3.2. □

4. Mathematical Models and Solution Techniques. Generating an exact solution for the bounded variable linear program is complicated by the fact that a solution of any finite termination procedure must not only satisfy $x_{\mathcal{B}} \geq 0$ but also $x_{\mathcal{B}} \leq u_{\mathcal{B}}$, where $u_{\mathcal{B}}$ is the subvector of u corresponding to $x_{\mathcal{B}}$.

The bounded variable linear program (1) contains two sets of inequalities involving the vector x ($x \geq 0$ and $x \leq u$). The component-wise distance of x^k to its lower bound is

$$(x^k_j - 0) \text{ or simply } x^k_j.$$

Similarly, $(u_j - x^k_j)$ is the component-wise distance of x^k to its upper bound.

4.1. Modified Weighted Projection. Williams [51] introduced a diagonal weighting matrix D, where

$$d_{jj} = \min(x^k_j, u_j - x^k_j) \quad \text{for } j \in \mathcal{B}, \tag{9}$$

which incorporates both lower and upper bound information into a finite termination procedure. Weighting the objective function by D penalizes the movement of the variables in the direction of their nearest bound. Therefore, if x^k_j for $j \in \mathcal{B}$ is close to its upper bound, the weight in (9) prevents the *jth* component of the solution vector $x_{\mathcal{B}}$ from violating its upper bound as well as its lower bound which is the desired result.

We propose the following weighted projection model

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\|D^{-1}(x_{\mathcal{B}} - x^k_{\mathcal{B}})\|^2 \\ \text{subject to} & A_{\mathcal{B}} x_{\mathcal{B}} = b \end{array} \tag{10}$$

where $d_{jj} = \min(x^k_j, u_j - x^k_j)$ for $j \in \mathcal{B}$ for the optimal primal face identification problem. When $x^k_j = u_j$ for $j \in \mathcal{B}$, corresponding columns from $A_{\mathcal{B}}$ can be removed and the right-hand side updated. If no upper bounds exist, problem (10) reduces to Ye's weighted projection model; see Ye [59].

Now, we demonstrate that, under certain conditions, projection problem (10) generates a positive solution. Without loss of generality, we assume $A_{\mathcal{B}}$ has full row rank.

THEOREM 4.1. *Let $\{(\tilde{x}^k, y^k, \tilde{z}^k)\}$ be generated by Algorithm 1. Assume $\tilde{\mathcal{B}}^k = \tilde{\mathcal{B}}$ and D is a nonsingular matrix such that*

$$\min(d_{jj}) > \left[(2n\mu^0 \tilde{\xi}^*)^{-1} (\tilde{\xi}^* \beta \|r^0\| + 2n\sqrt{2n}\mu^0 \tilde{\tau} \|A_{\mathcal{N}}\|)\right] \|A^T_{\mathcal{B}} (A_{\mathcal{B}} A^T_{\mathcal{B}})^{-1}\| \left(\tilde{x}^{kT} \tilde{z}^k\right).$$

Then the solution $x_{\mathcal{B}}$ obtained from solving

$$\begin{array}{ll} \textit{minimize} & \frac{1}{2}\|D^{-1}(x_{\mathcal{B}} - x_{\mathcal{B}}^k)\|^2 \\ \textit{subject to} & A_{\mathcal{B}} x_{\mathcal{B}} = b \end{array} \tag{11}$$

satisfies $0 < x_{\mathcal{B}} \le u_{\mathcal{B}}$, *and* $0 \le s_{\mathcal{B}} < u_{\mathcal{B}}$.

Proof. Let $dx_{\mathcal{B}} = D^{-1}\Delta x_{\mathcal{B}}$, where $\Delta x_{\mathcal{B}} = x_{\mathcal{B}} - x_{\mathcal{B}}^k$. Therefore,

$$\begin{aligned} \|dx_{\mathcal{B}}\| &= \|DA_{\mathcal{B}}^T(A_{\mathcal{B}}D^2A_{\mathcal{B}}^T)^{-1}(b - A_{\mathcal{B}}x_{\mathcal{B}}^k)\| \\ &= \|DA_{\mathcal{B}}^T(A_{\mathcal{B}}D^2A_{\mathcal{B}}^T)^{-1}A_{\mathcal{B}}DD^{-1}[A_{\mathcal{B}}^T(A_{\mathcal{B}}A_{\mathcal{B}}^T)^{-1}(b - A_{\mathcal{B}}x_{\mathcal{B}}^k)]\| \\ &\le \|D^{-1}\|\|A_{\mathcal{B}}^T(A_{\mathcal{B}}A_{\mathcal{B}}^T)^{-1}\|\|(b - A_{\mathcal{B}}x_{\mathcal{B}}^k)\| \end{aligned} \tag{12}$$

From Lemma 3.2, we have

$$\|dx_{\mathcal{B}}\| \le \|D^{-1}\|\|A_{\mathcal{B}}^T(A_{\mathcal{B}}A_{\mathcal{B}}^T)^{-1}\| \left[(2n\mu^0\xi^*)^{-1}(\xi^*\beta\|r^0\| + \mu^0\tau 2n\sqrt{2}n\|A_{\mathcal{N}}\|)\right](\tilde{x}^k)^T\tilde{z}^k.$$

Consequently, $\|dx_{\mathcal{B}}\|_\infty \le \|dx_{\mathcal{B}}\|_2 < 1$. The first inequality holds from the fact that for any vector v, $\|v\|_\infty \le \|v\|_2$, the second inequality from assumption (2).

Using assumption (2) and the fact that $\|dx_{\mathcal{B}}\|_\infty \le \|dx_{\mathcal{B}}\|_2 < 1$, we have the following description of the solution.

Case 1: If $x_j^k < u_j - x_j^k$ for all $j \in \mathcal{B}$, then

$$-1 < \frac{\Delta x_j}{x_j^k} < 1.$$

Hence $x_j^k - x_j^k < x_j^k + \Delta x_j < x_j^k + x_j^k$ and consequently, $0 < x_{\mathcal{B}} < 2x_{\mathcal{B}} < u_{\mathcal{B}}$.

Case 2: If $u_j - x_j^k < x_j^k$ for all $j \in \mathcal{B}$, then

$$-1 < \frac{\Delta x_j}{u_j - x_j^k} < 1.$$

Hence $x_j^k - (u_j - x_j^k) < x_j^k + \Delta x_j < x_j^k + (u_j - x_j^k)$ and consequently, $0 < x_{\mathcal{B}} < u_{\mathcal{B}}$.

From the constraint $x + s = u$ and Cases 1 and 2, we have $0 \ge s_{\mathcal{B}} < u_{\mathcal{B}}$. This completes the proof. ☐

Now, let us consider the optimal dual face identification problem. To find a feasible point on the optimal dual face for bounded variable problems, we solve

$$\min_y \quad \frac{1}{2}\|D^{-1}(A_{\mathcal{B}}^Ty - c_{\mathcal{B}})\|^2 \tag{13}$$

for Δy. Recall that $\Delta y = y - y^k$. Here, the matrix D is defined as in (9) to save computational expense. With this formulation only one matrix factorization is needed to solve the both the optimal primal and dual face identification problems.

To conclude that our finite termination procedure is successful, we must now show that the dual variables $z_{\mathcal{N}}$ and $w_{\mathcal{N}}$ are nonnegative.

LEMMA 4.2. *Assume* $c_{\mathcal{B}} \in range(A_{\mathcal{B}}^T)$, $\tilde{\mathcal{B}}^k = \tilde{\mathcal{B}}$, *and*

$$\begin{aligned} \min(\tilde{z}_{\mathcal{N}}^k) > (2n\mu^0\xi^*)^{-1}[&\beta\xi^*\left(1 + \|A_{\mathcal{N}}^T\|\|(A_{\mathcal{B}}A_{\mathcal{B}}^T)^{-1}A_{\mathcal{B}}\|\right)\|r^0\| \\ &+ 4n\sqrt{2n}\tau\mu^0\|A_{\mathcal{N}}^T\|\|(A_{\mathcal{B}}A_{\mathcal{B}}^T)^{-1}A_{\mathcal{B}}\|]\left((x^k)^Tz^k + (s^k)^Tw^k\right). \end{aligned}$$

Then $z_{\mathcal{N}}$ *and* $w_{\mathcal{N}}$ *are nonnegative.*

Proof. Let $\Delta y = y - y^k$. Then

$$\begin{aligned} \Delta y &= (A_{\mathcal{B}} D^2 A_{\mathcal{B}}^T)^{-1} A_{\mathcal{B}} D^2 (c_{\mathcal{B}} - A_{\mathcal{B}}^T y) \\ &= (A_{\mathcal{B}} A_{\mathcal{B}}^T)^{-1} A_{\mathcal{B}} (c_{\mathcal{B}} - A_{\mathcal{B}}^T y) \text{ since } c_{\mathcal{B}} \in range(A_{\mathcal{B}}^T). \end{aligned} \tag{14}$$

From problem (3), we have $A^T y + z - w = c$ which gives us

$$c_{\mathcal{N}} - A_{\mathcal{N}}^T y = z_{\mathcal{N}} - w_{\mathcal{N}}.$$

If $c_{\mathcal{N}} - A_{\mathcal{N}}^T y < 0$, then $z_{\mathcal{N}} = 0$ and $w_{\mathcal{N}} > 0$. Therefore, we have to check for optimality(i.e., that $w_{\mathcal{N}}$ is positive). From the dual constraint, it is easy to see that

$$\begin{aligned} \|dw_{\mathcal{N}}\| &= \|A_{\mathcal{N}}^T (y - y^k) - r_{d_{\mathcal{N}}}^k\| \\ &\le \|A_{\mathcal{N}}^T\| \|(A_{\mathcal{B}} A_{\mathcal{B}}^T)^{-1} A_{\mathcal{B}}^T\| \|c_{\mathcal{B}} - A_{\mathcal{B}}^T y^k\| + \|r_d^k\|. \end{aligned} \tag{15}$$

From Lemma 3.1 and the bound on $\tilde{z}_{\mathcal{B}}^k$, we have

$$\begin{aligned} \|dw_{\mathcal{N}}\| \le (2n\mu^0\xi^*)^{-1} [\beta\xi^* \left(\|A_{\mathcal{N}}^T\| \|(A_{\mathcal{B}} A_{\mathcal{B}}^T)^{-1} A_{\mathcal{B}}^T\| + 1 \right) \|r^0\| \\ + 4n\sqrt{2n}\tau\mu^0 \|A_{\mathcal{N}}^T\| \|(A_{\mathcal{B}} A_{\mathcal{B}}^T)^{-1} A_{\mathcal{B}}^T\|] \left((\tilde{x}^k)^T \tilde{z}^k \right). \end{aligned}$$

From the third assumption of the lemma, we have $\|dw_{\mathcal{N}}\|_2 < 1$. Thus $\|dw_{\mathcal{N}}\|_\infty \le \|dw_{\mathcal{N}}\|_2 < 1$. Therefore

$$-1 < \frac{\Delta w_j}{w_j^k} < 1 \Rightarrow \quad w_j^k - w_j^k < w_j^k + \Delta w_j < w_j^k + w_j^k \quad \text{ for all } j \in \mathcal{N}.$$

If $c_{\mathcal{N}} - A_{\mathcal{N}}^T y > 0$, then $z_{\mathcal{N}} > 0$ and $w_{\mathcal{N}} = 0$. Hence, we only have to check that $z_{\mathcal{N}}$ is positive. The proof that $z_{\mathcal{N}} > 0$ follows the same format as the preceding proof that $w_{\mathcal{N}} > 0$. □

4.2. Scaled Mehrotra-Ye. We also study the effectiveness of using Gaussian elimination to solve the linear feasibility problems associated with the optimal primal and dual faces. We propose a strategy that explicitly incorporates the current interior-point iterate into the model. The authors proposed column scaling the constraint matrix by the current interior-point iterate to potentially bias the Gaussian elimination so that the columns corresponding to the smallest components of d are chosen as pivots last. Skeel [47] provided a theoretical basis for column scaling as an effective tool to achieve numerical stability.

Skeel [47] proved that Gaussian elimination with row pivoting is numerically stable when the matrix is column scaled by $D = \text{diag}\,|\hat{v}|$, where $\hat{v}$ is the computed solution of $Av = g$. Unfortunately, the value of $\hat{v}$ is not known when the factorization begins. The theory assumes that columns of A corresponding to negligible components of $\hat{v}$ are selected as pivots last.

Skeel's theorem motivated the authors to compute a feasible point on the optimal primal and dual faces, respectively, by solving the following linear systems

$$(A_{\mathcal{B}} D^k) dx_{\mathcal{B}} = b - A_{\mathcal{B}} x_{\mathcal{B}}^k \quad \text{and} \quad (D^k A_{\mathcal{B}}^T) dy = D(c_{\mathcal{B}} - A_{\mathcal{B}}^T y^k). \tag{16}$$

Once again we use D as defined in (9). The authors row scale the dual linear system by D so that LU factorization of $A_{\mathcal{B}} D$ can be used to solve the dual linear system. Any other scale would necessitate two matrix factorizations in the finite termination

procedure which would make the procedure twice as expensive as an interior-point iteration.

We now prove that the solutions of the linear systems just described will result in strictly positive vectors, $x_{\mathcal{B}}$ and $z_{\mathcal{N}}$. First, we define $A_{\bar{\mathcal{B}}}$ as an arbitrary nonsingular submatrix of $A_{\mathcal{B}}$ with maximal rank. Similarly, $A_{\bar{\mathcal{N}}}$ denotes the corresponding submatrix of $A_{\mathcal{N}}$.

Recall that

$$\tilde{x} = (x, s) \in \mathbf{R}^{2n}$$
$$\tilde{z} = (z, w) \in \mathbf{R}^{2n}.$$

THEOREM 4.3. *Consider the iteration sequence $\{(\tilde{x}^k, y^k, \tilde{z}^k)\}$ generated by Algorithm 1. Assume*

(1) $\tilde{\mathcal{B}}^k = \tilde{\mathcal{B}}$.

(2) $A_{\bar{\mathcal{B}}}$, *and* $\bar{D}$ *are nonsingular matrices, where D is defined in (9).*

(3) $\min(d_{jj}) > \|A_{\bar{\mathcal{B}}}^{-1}\| \left[(2n\mu^0\xi^*)^{-1}(\xi^*\beta\|r^0\| + \mu^0\tau 2n\sqrt{2n}\|A_{\mathcal{N}}\|)\right]\left((\tilde{x}^k)^T\tilde{z}^k\right)$.

(4) $\min(\tilde{z}^k_{\mathcal{N}}) > (2n\mu^0\xi^*)^{-1}[\beta\xi^*\left(1 + \|A^T_{\mathcal{N}}\|\|A_{\bar{\mathcal{B}}}^{-T}\|\right)\|r^0\|$

$$+ 4n\sqrt{2n}\tau\mu^0\|A^T_{\mathcal{N}}\|\|A_{\bar{\mathcal{B}}}^{-T}\|]\left((\tilde{x}^k)^T\tilde{z}^k\right).$$

Then $0 < x_{\mathcal{B}} = x^k_{\mathcal{B}} + Ddx_{\mathcal{B}} \leq u_{\mathcal{B}}$, where $dx_{\mathcal{B}}$ is the solution of the problem

$$(A_{\mathcal{B}}D)dx_{\mathcal{B}} = b - A_{\mathcal{B}}x^k_{\mathcal{B}}, \tag{17}$$

and $0 \leq s_{\mathcal{B}} < u_{\mathcal{B}}$. Moreover, $z_{\mathcal{N}}, w_{\mathcal{N}} \geq 0$ where $z_{\mathcal{N}} - w_{\mathcal{N}} = c_{\mathcal{N}} - A^T_{\mathcal{N}}y$, $y = y^k + \Delta y and \Delta y$ solves

$$(DA^T_{\mathcal{B}})\Delta y = D(c_{\mathcal{B}} - A^T_{\mathcal{B}}y^k). \tag{18}$$

Proof. Assume $A_{\bar{\mathcal{B}}}$ is a nonsingular submatrix of $A_{\mathcal{B}}$. Let $dx_{\bar{\mathcal{B}}} = \bar{D}^{-1}\Delta x_{\bar{\mathcal{B}}}$. Then

$$\begin{aligned}\|\Delta x_{\bar{\mathcal{B}}}\| &= \|\bar{D}dx_{\bar{\mathcal{B}}}\| \\ &= \|\bar{D}(A_{\bar{\mathcal{B}}}\bar{D})^{-1}(\bar{b} - A_{\bar{\mathcal{B}}}x^k_{\bar{\mathcal{B}}})\| \\ &\leq \|A_{\bar{\mathcal{B}}}^{-1}\|(\mu^0 2n\xi^*)^{-1}\left(\xi^*\beta\|r^0\| + \mu^0\tau 2n\sqrt{2}n\|A_{\mathcal{N}}\|\right)(\tilde{x}^k)^T\tilde{z}^k.\end{aligned} \tag{19}$$

The last inequality holds from an application of (??). From assumption (3), the solution $x_{\bar{\mathcal{B}}}$ is positive. Let $dx_j = 0$, for $j \in \mathcal{B} \setminus \bar{\mathcal{B}}$ when column $A_{.j}$ is linearly dependent on columns of $A_{\bar{\mathcal{B}}}$. Hence the entire vector $x_{\mathcal{B}}$ is positive.

We now prove that the dual solution is positive. First, we compute the solution of the scaled dual Mehrotra-Ye formulation.

$$\Delta\bar{y} = A_{\bar{\bar{B}}}^{-T}(c_{\bar{\mathcal{B}}} - A^T_{\bar{\bar{B}}}\bar{y}^k).$$

If $c_{\mathcal{N}} - A^T_{\mathcal{N}}y > 0$, then $z_{\mathcal{N}} > 0$ and $w_{\mathcal{N}} = 0$.

$$\begin{aligned}\|\Delta z_{\mathcal{N}}\| &= \|N^T(y - y^k) - r^k_{d_{\mathcal{N}}}\| \\ &\leq \|N^T(y - y^k)\| + \|r^k_{d_{\mathcal{N}}}\| \\ &\leq \|N^T\|\|\bar{Y}^k(\bar{B}^T\bar{Y}^k)^{-1}(c_{\bar{\mathcal{B}}} - \bar{B}^T\bar{y}^k)\| + \|r^k_{d_{\mathcal{N}}}\| \\ &\leq \|N^T\|\|\bar{B}^{-T}(c_{\mathcal{B}} - \bar{B}^Ty^k)\| + \|r^k_d\| \\ &\leq (2n\mu^0\xi^*)^{-1}\left[\beta\xi^*\left(1 + \|N^T\|\|\bar{B}^{-T}\|\right)r^0\| + 2n\sqrt{2}n\tau\mu^0\|N^T\|\|\bar{B}^{-T}\|\right]\left((x^k)^Tz^k\right).\end{aligned} \tag{20}$$

The last inequality follows from (??) and the Feasibility Priority Principle. Hence by assumption (3), $z_{\mathcal{N}} > 0$.

We similarly prove that $w_{\mathcal{N}} > 0$. ∎

4.3. Nonlinear Model. El-Bakry [10] proposed the following function,

$$\phi(\bar{x}, x) = ||(\bar{X}X)^{-1/2}(\bar{x} - x)||,$$

to measure distances in the cone $\{x \in \mathbf{R}^n : x \geq 0\}$. This function is used by El-Bakry, Farah, and Tapia [11] in the interior-point steepest descent method. It is a generalization of a central path proximity measure in primal-dual interior-point methods which was used by Roos, Terlaky, and Jensen [45], Argaez and Tapia [5], and Nesterov and Todd [40]. Since the distance function $\phi(\bar{x}, x)$ contains boundary information, it is the obvious choice to use in a finite termination procedure to bias the solution away from the boundary.

We now study a nonlinear weighted projection model of the form

$$\begin{aligned} \min \quad & \tfrac{1}{2}||D(x_{\mathcal{B}}, x_{\mathcal{B}}^k)(x_{\mathcal{B}} - x_{\mathcal{B}}^k)||^2 \\ \text{s.t.} \quad & A_{\mathcal{B}} x_{\mathcal{B}} = b, \end{aligned} \tag{21}$$

where $D(x_{\mathcal{B}}, x_{\mathcal{B}}^k) = (X_{\mathcal{B}} X_{\mathcal{B}}^k)^{-1/2}$ to generate a feasible point on the optimal primal face.

Problem (21) incorporates information about the solution of the projection problem into the model as well as the nonnegativity constraints, which is desired. However, because of the nonlinearity of the weighting matrix $D(x_{\mathcal{B}}, x_{\mathcal{B}}^k)$, we can not use a direct method to solve the model. Therefore, the solution of problem (21) cannot be labeled an exact solution.

Solving the first order conditions of (21) directly, yields

$$x_{\mathcal{B}} = X_{\mathcal{B}}^3 (X_{\mathcal{B}}^k)^{-2} e + X_{\mathcal{B}}^3 (X_{\mathcal{B}}^k)^{-1} A_{\mathcal{B}}^T (A_{\mathcal{B}} X_{\mathcal{B}}^3 (X_{\mathcal{B}}^k)^{-1} A_{\mathcal{B}}^T)^{-1} (b - A_{\mathcal{B}} X_{\mathcal{B}}^3 (X_{\mathcal{B}}^k)^{-2} e). \tag{22}$$

If in (22), we replace $X_{\mathcal{B}}$ on the right-hand side with $X_{\mathcal{B}}^k$, then

$$x_{\mathcal{B}} = x_{\mathcal{B}}^k + (X_{\mathcal{B}}^k)^2 A_{\mathcal{B}}^T (A_{\mathcal{B}} (X_{\mathcal{B}}^k)^2 A_{\mathcal{B}}^T)^{-1} (b - A_{\mathcal{B}} x_{\mathcal{B}}^k), \tag{23}$$

which is the solution of Ye's primal weighted projection problem; see Ye [59].

The corresponding dual model is

$$\min \quad \tfrac{1}{2}||D(x_{\mathcal{B}}, x_{\mathcal{B}}^k)^{-1}(c_{\mathcal{B}} - A_{\mathcal{B}}^T y)||^2.$$

Let $\phi(y) = \frac{1}{2}||D(x_{\mathcal{B}}, x_{\mathcal{B}}^k)^{-1}(c_{\mathcal{B}} - A_{\mathcal{B}}^T y)||^2$. Then $\nabla\phi(y) = -2A_{\mathcal{B}} X_{\mathcal{B}} X_{\mathcal{B}}^k (c_{\mathcal{B}} - A_{\mathcal{B}}^T y)$. The formulation of $\nabla\phi(y)$ would suggest first solving for $x_{\mathcal{B}}$, updating the matrix $D(x_{\mathcal{B}}, x_{\mathcal{B}}^k)$, and then solving the $\nabla\phi(y) = 0$ for Δy. Recall that $y = y^k + \Delta y$. However, numerically, this would require two matrix factorizations to find feasible points on the optimal primal and dual faces. Instead we reuse the matrix factorization from (23).

5. Computational Results. The numerical experiments were conducted on a Sparcworkstation with application hardware Sun4. We used the LIPSOL - Linear programming Interior-Point SOLver- package developed under the MATLAB[1] environment. The software package, written by Zhang [61], implements an infeasible primal-dual predictor-corrector interior point method. The *netlib* suite of linear programming problems comprises the test set.

The initial matrix is scaled in an attempt to achieve row/column equilibration. Preprocessing deletes fixed variables, deletes zero rows and columns from the matrix

[1]MATLAB is a registered trademark of The MathWorks, Inc.

A, solves equations of one variable, and shifts nonzero lower bounds to zero. For problem *greenbea*, preprocessing deletes fixed variables, deletes zero columns from the matrix A, and shifts nonzero lower bounds to zero. No other preprocessing is performed.

5.1. Projections. We implemented the following optimal face identification procedure.

PROCEDURE 1 (Optimal Face Identification - Finite Termination Procedure).

Step 0 *Set attempt = 0.*

Step 1 *Use an interior-point algorithm to generate the iteration sequence* $(x^k, y^k, z^k, s^k, w^k)$.

Step 2 *While*

$$\frac{|c^T x^k - (b^T y^k - u^T w^k)|}{1 + |b^T y^k - u^T w^k|} \leq 10^{-8} \text{ and attempt } \leq 6,$$

set

$$\mathcal{B}^k = \{j : z_j^k \leq 1.e - 14 \quad or \quad |\Delta^p x^k| / x_j^k \leq |\Delta^p z^k| / z_j^k\}.$$

Here variables and the relative change of variables are used to identify the optimal partition.

Step 3 *Solve the optimal primal and dual face identification problems by the modified weighted projection models described in Section 3.1.*

Step 4 *Update,*

$$x = \begin{cases} x_j^k + \Delta x_j & j \in \mathcal{B} \\ 0 & j \in \mathcal{N} \end{cases}$$

and $y = y^k + \Delta y$. *Set* $s = u - x$ *and* $\delta = c - A^T y$, *then*

$$z_j = \begin{cases} 0 & \text{if } \delta_j < 0 \\ \delta_j & \text{else} \end{cases} \quad \text{and } w_j = \begin{cases} -\delta_j & \text{if } \delta_j < 0 \\ 0 & \text{else} \end{cases}$$

Step 5 *Set dual infeasibility (dbi)* $= \max(0, -z_N, -z_B)$. *If the computed solution is complementary and satisfies*

$$\max\left(\frac{||Ax - b||}{1 + ||b||}, \frac{||A^T y + z - w - c||}{1 + ||c||}, \frac{|c^T x - (b^T y - u^T w)|}{1 + |b^T y - u^T w|}\right) \leq 10^{-11}$$

we terminate the algorithm with a solution. If not, we repeat the finite termination procedure at the next interior-point iteration.

In Step 2, the transformed linear systems that define the optimal face are solved in the least squares sense. If the linear system is underdetermined, we factor the matrix $A_{\mathcal{B}} D^2 A_{\mathcal{B}}^T$, where D is as proposed in previous sections. Likewise, if the linear system is overdetermined, we factor the matrix $A_{\mathcal{B}}^T A_{\mathcal{B}}$. The update formula for the dual variables was first used by Resende, et al in [43, 44] and Portugal [41] to generate feasible dual variables.

We now compare three projection models. The first one is the orthogonal projection model where $D = I$, the second is Ye's weighted projection method with $D = X_{\mathcal{B}}^k$, and the third is the modified weighted projection. The first two models were developed by Ye [57, 59].

We tested 35 problems with upper bounds from *netlib*. Column 1 gives the number of failed calls to the finite termination procedure the optimal face identification

	Subproblems		
# of misses	OP	WP	MWP
0	13	15	19
1	13	12	11
2	2	2	4
3	2	3	1
4	2	0	0
5	1	1	0
more than 5	2	2	0
TOTAL misses	46	40	22

TABLE 1
Problems with upper bound constraints

problem was solved to the desired accuracy. We consider a call a failure if the procedure does not generate a positive solution that satisfies the optimality and feasibility tolerances. The second column gives the number of problems solved by the orthogonal projection model for the given number of misses. The third column contains the computational results of Ye's weighted projection model and the fourth column the modified weighted projection.

With the modified weighting matrix, we are able to compute interior points on the optimal face for all problems in the test set. The other two models fail to deliver a solution for two problems, *greenbea* and *nesm*. If we weight the constraint matrix of problem *greenbea* by the modified weighting matrix, we can compute an interior point on the optimal face in one projection attempt. The solution agrees to thirteen digits with the CPLEX[2] reoptimized objective function value reported in Table II of [7]. The most accurate solutions are obtained when we weight by the matrix D.

5.2. Scaled Mehrotra-Ye. We compare the effectiveness of the Mehrotra-Ye method and scaled Mehrotra-Ye in finding an interior point on the optimal face. We use a dense implementation of the Gaussian elimination routine; therefore, our test set was restricted to problems where the matrix $A_{\mathcal{B}}$ had approximately 500 rows and columns. Consequently, our test set consisted of 55 problems from the *netlib* suite. The three largest problems tested were *maros* with 835 rows and 1921 columns, *scsd8* with 397 rows and 2750 columns, and *ship08l* with 688 rows and 4339 columns. The removal of columns corresponding to zero variables combined with the elimination of zero rows reduced the original matrix A to the desired dimensions.

If a negligible pivot was encountered, the column was removed from the matrix. The pivot tolerance was

$$\max(m, |\mathcal{B}|) * ||A_{\mathcal{B}}||_1 * 10^{-16},$$

which is the same default tolerance used in MATLAB to determine the numerical rank of a matrix. We did not pivot to minimize fill-in of the triangular factors, L and U. Zero rows were removed before the Gaussian elimination subroutine started. At the completion of the factorization any remaining zero rows were deleted. Components of the solution vector $(\Delta x, \Delta y)$ corresponding to dependent rows and columns were set to zero.

[2]CPLEX is a trademark of CPLEX Optimization, Inc.

	Techniques	
# of misses	Mehrotra-Ye	Scaled
0	44	47
1	7	6
2	2	1
3	1	1
4	0	0
5	1	0
TOTAL MISSES	19	11

TABLE 2

The number of misses per technique

Table 4.2 shows the results of our numerical experiments. Column 1 gives the number of calls to the finite termination procedure. Columns 2 through 3 give the number of problems solved by the two variants of the Mehrotra-Ye procedure.

Column scaling saved one interior-point iteration for problems, *boeing2, kb2,* and *seba*. Two interior-point iterations are saved for problems *etamacro, finnis* and *stair*. When we implemented the standard Mehrotra-Ye approach, six tries were needed to find feasible primal and dual points for problem *etamacro*. However, the standard Mehrotra-Ye procedure generates the most accurate solutions. For 91 percent of the problems, the objective function value agrees to thirteen digits with the reoptimized CPLEX objective function value that was reported in [7]. The thirteen digit agreement is 89 percent for scaled Mehrotra-Ye.

When the matrices were scaled, the modified weighted projection method needed one more projection attempt than did scaled Mehrotra-Ye for problems *boeing2* and *finnis*.

5.3. Nonlinear Weighted Projections. Finally we discuss our solution technique for the nonlinear projection model. The major difference in the solution technique for the nonlinear weighted projection method when compared to the one for the modified weighted projection is that we iteratively compute (x, y, z, s, w).

We modify Procedure 1 by replacing Step 3 with the following

PROCEDURE 2 (Nonlinear Model Embedded in a Finite Termination Procedure).

Step 3 *Set $l = 0$ and $(x^{(l)}, y^{(l)}, z^{(l)}, s^{(l)}, w^{(l)}) = (x^k, y^k, z^k, s^k, w^k)$.*

Step 4 *(Inner Loop) While $l \leq 2$ and $maxrerror^{(l)} \geq 1.0e-11$*

Step 4.1 *Solve*

$$A_{\mathcal{B}} X_{\mathcal{B}}^{(0)} \Delta x_{\mathcal{B}} = (b - A_{\mathcal{B}} x_{\mathcal{B}}^{(l)}) \text{ for } \Delta x_{\mathcal{B}}.$$

Step 4.2 *Solve*

$$(A_{\mathcal{B}} (X_{\mathcal{B}}^{(l)})^2 A_{\mathcal{B}}^T) \Delta y = A_{\mathcal{B}} (X_{\mathcal{B}}^{(l)})^2 (c_{\mathcal{B}} - A_{\mathcal{B}}^T y^{(l)}) \text{ for } \Delta y.$$

Step 4.3 *Update,*

$$x^{(l+1)} = \begin{cases} x_j^{(l)} + \Delta x_j & j \in \mathcal{B} \\ 0 & j \in \mathcal{N} \end{cases}$$

$$y^{(l+1)} = y^{(l)} + \Delta y \text{ and } z^{(l+1)} = c - A^T y^{(l+1)}.$$

Problem	Weighted Projection	Nonlinear Projection
boeing1	3/4	3/8
bore3d	3/4	3/8
finnis	5/6	3/6
fit1d	3/4	3/6
ganges	3/4	1/4
grow22	3/4	3/6
pilotja	3/4	3/8
pilotwe	3/4	3/6
standgub	3/4	3/6
TOTAL	29/38	25/58

TABLE 3
Number of matrix factorizations and back substitutions required to reach desired accuracy

Step 4.4 *If upper bounds exist, set* $s^{(l+1)} = u - x^{(l+1)}$, $\delta = c - A^T y^{(l+1)}$,

$$z^{(l+1)} = \begin{cases} 0 & \textit{if } \delta < 0 \\ \delta & \textit{else} \end{cases} \quad \textit{and } w^{(l+1)} = \begin{cases} -\delta & \textit{if } \delta < 0 \\ 0 & \textit{else} \end{cases}$$

Step 4.5 $l \leftarrow l + 1$

Step 4.6 *If* $maxrerror^{(l)} > maxrerror^{(l-1)}$*, then* $attempt \leftarrow attempt + 1$, $k \leftarrow k + 1$*, and goto Step 1.*

Step 4.7 *If* $\max(z_{\mathcal{B}}) > \min(z_{\mathcal{N}})$*, then* $attempt \leftarrow attempt + 1$, $k \leftarrow k + 1$*,and goto Step 1.*

Step 4.8 *Set dual bound infeasibility (dbi)* $= \max(0, -z_{\mathcal{N}}^{(l)}, -z_{\mathcal{B}}^{(l)})$. *If the computed solution is complementary and satisfies* $maxrerror^{(l)} \leq 1.0e-11$ *as well as* $dbi < 1.0e-09$*, we terminate the algorithm with a solution. Otherwise, goto Step 4.*

The primal and dual updates have the flavor of simplified Newton's method since we reuse our initial matrix factorization each time we recompute $\Delta x_{\mathcal{B}}$ and Δy.

Notice that while the number of matrix factorizations decrease with the addition of the nonlinear weighted projection model, the back substitutions increase by more than 50 percent.

6. Concluding Remarks.

REFERENCES

[1] I. Adler and R. Monteiro. A Geometric View of Parametric Linear Programming. *Algorithmica*, 8:161–176, 1992.

[2] E. Andersen. On exploiting problem structure in a basis identification procedure for linear programming. Technical Report, Department of Management, Odense University, Denmark, 1996.

[3] E. Andersen and Y. Ye. Combining interior-point and pivoting algorithms for linear programming. Technical Report, Department of Management Sciences, University of Iowa, 1994.

[4] K. Anstreicher, J. Ji, F. Potra, and Y. Ye. Probabilistic analysis of an infeasible interior-point algorithm for linear programming. Reports on Computational Mathematics, No. 27, Department of Mathematics, University of Iowa, 1993.

[5] M. Argaez and R. Tapia. On the global convergence of a modified augmented Lagrangian linesearch interior-point Newton method for nonlinear programming. Technical Report, TR95-38, Department of Computational and Applied Mathematics, Rice University, 1995.

[6] R. Bixby and M. Saltzman. Recovering an Optimal LP Basis from an Interior-Point Solution. Technical Report, TR91-32, Department of Mathematical Sciences, Rice University, 1991.

[7] R. Bixby. Implementing the Simplex Method: The Initial Basis *ORSA Journal on Computing*, 4:267–284, 1992.

[8] G. Dantzig. *Linear Programming and Extensions*, Princeton University Press, Princeton, NJ, 1963.

[9] A. El-Bakry, R. Tapia, and Y. Zhang. A study of indicators for identifying zero variables in interior-point methods. *SIAM Review*, 36:45–72, 1994.

[10] A. El-Bakry and R. Tapia. A distance function and its applications. Technical Report, TR98-??, Department of Computational and Applied Mathematics, Rice University, 1998.

[11] A. El-Bakry, J. Farah and R. Tapia. Convergence theory and implementation issues of the interior-point steepest descent method. Technical Report, TR98-??, Department of Computational and Applied Mathematics, Rice University, 1998.

[12] R. Freund. An Infeasible-Start Algorithm for Linear Programming whose Complexity depends on the Distance from the Starting Point to the Optimal Solution. Technical Report, Sloan W.P. No.3559-93, Sloan School of Management Massachusetts Institute of Technology, 1993.

[13] D. Gay. Stopping tests that compute optimal solutions for interior-point linear programming algorithms. In S. Gómez, J. Hennart, and R. Tapia, eds., *Advances in Numerical Partial Differential Equations and Optimization: Proceedings of the Fifth Mexico-United States Workshop*, pages 17–42. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1991.

[14] A. Goldman and A. Tucker. Theory of Linear Programming. In H.W. Kuhn and A.W. Tucker, eds., *Linear Inequalties and Related Systems*, pages 53–97. Princeton University Press, Princeton, NJ, 1956.

[15] G. Golub and C. Van Loan. *Matrix Computations*, Second Edition, John Hopkins University Press, Baltimore, MD, 1989.

[16] H. Greenberg. Rim Sensitivity Analysis from an Interior Solution. CCM N0. 86, Center for Computational Mathematics, Mathematics Department, University of Colorado at Denver, 1996.

[17] H. Greenberg. Matrix Sensitivity Analysis from an Interior Solution of a Linear Program. CCM N0. 104, Center for Computational Mathematics, Mathematics Department, University of Colorado at Denver, 1997.

[18] O. Güler and Y. Ye. Convergence behavior of interior-point algorithms. *Mathematical Programming*, 60:215-228, 1993.

[19] B. Jansen, C. Roos, and T. Terlaky. Sensitivity analysis in linear programming: just be careful! *European Journal of Operations Research*101:15–28, 1997.

[20] B. Jansen, C. Roos, and T. Terlaky. An Interior-Point Approach to Postoptimal and Parametric Analysis in Linear Programming. Report No. 92-21, Faculteit der Technische Wiskunde en Informatica, Technische Universiteit Delft, The Netherlands.

[21] B. Jansen, C. Roos, T. Terlaky, and J.-Ph. Vial. Interior-Point Methodology for Linear Programming: Duality, Sensitivity Analysis and Computation Aspects. In K. Frauendorfer, H. Glavitsch, and R. Bacher, eds., *Optimization in Planning and Operation of Electri Power Systems*, pages 57–123. Physica Berlag, 1993.

[22] J. Ji and F. Potra. Tapia indicators and finite termination of infeasible interior-point methods for degenerate LCP. Technical Report No. 81/1995, Department of Mathematics, University of Iowa, 1995.

[23] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4: 373-395, 1984.

[24] D. Kincaid and W. Cheney. *Numerical Analysis*,Brooks/Cole Publishing Company, Pacific Grove, CA 1991.

[25] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual interior-point method for linear programming. In N. Megiddo, ed., *Progress in Mathematical Programming, Interior Point and Related Methods*, pages 29–47. Springer–Verlag, New York, NY, 1989.

[26] M. Gonzalez-Lima. Effective Computation of the Analytic Center of the Solution Set in Linear Programming Using Primal-Dual Interior-Point Methods. PhD thesis, Department of Computational and Applied Mathematics, Rice University, 1994.

[27] R. Marsten, M. Saltzman, D. Shanno, G. Pierce, and J. Ballintijn. Implementation of a Dual Affine Interior-Point Algorithm for Linear Programming. *ORSA Journal on Computing*, 1:4, 287–297, 1989.

[28] L. McLinden. An analogue of Moreau's proximation theorem with application to nonlinear complementarity problem. *Pacific Journal Math*, 88:101–161, 1988.

[29] N. Megiddo. Pathways to the Optimal Set in Linear Programming. In N. Megiddo, ed., *Progress in Mathematical Programming, Interior Point and Related Methods*, pages 29–

47. Springer–Verlag, New York, NY, 1989.

[30] N. Megiddo. On Finding Primal- and Dual-Optimal Bases. *ORSA Journal on Computing*, 3:1, 63–65, 1991.

[31] S. Mehrotra. Implementations of affine scaling methods: towards faster implementations with complete Cholesky factor in use. Technical Report TR89-15, Dept. of Industrial Engineering and Management Sciences, Northwestern University, 1989.

[32] S. Mehrotra. On Finding a Vertex Solution Using Interior-Point Methods. *Linear Algebra and its Applications*, 152:233–253, 1991.

[33] S. Mehrotra. Finite Termination and Superlinear Convergence in Primal-Dual Methods. Technical Report TR91-13, Dept. of Industrial Engineering and Management Sciences, Northwestern University, 1991.

[34] S. Mehrotra. On the implementation of a primal-dual interior-point method. *SIAM Journal on Optimization*, 2(4):575–601, 1992.

[35] S. Mehrotra. Quadratic Convergence in a Primal-Dual Method. *Mathematics of Operations Research*, 18(3):741–751, 1993.

[36] S. Mizuno. A Predictor-Corrector Infeasible Interior-Point Algorithm for Linear Programming. *Operations Research Letters*,16(2):61–66, 1994.

[37] S. Mehrotra and Y. Ye. Finding an Interior Point in the Optimal Face of Linear Programs. *Mathematical Programming*, 62:497–515, 1993.

[38] R. Monteiro and S. Mehrotra. A General Parametric Analysis Approach and its Implication to Sensitivity Analysis in Interior-Point Methods. *Mathematical Programming*, 72: 65–82, 1996.

[39] R. Monteiro and S. Wright. Local Convergence of Interior-Point Algorithms for Degenerate Monotone LCP. *Computational Optimization and Applications*, 3: 131–155, 1994.

[40] Y. Nesterov and M. Todd.

[41] L. Portugal, M. Resende, G. Veiga, and J. Judice. A Truncated Primal-Infeasible Dual-Feasible Network Interior-Point Method. 1994.

[42] F. Potra. A quadratically convergent predictor-corrector method for solving linear programs from infeasible starting points. *Mathematical Programming*, 67:383–406, 1994.

[43] M. Resende and G. Veiga. An Efficient Implementation of a Network Interior-Point Method. In D. Johnson and C. McGeoch, eds., *Network Flows and Matching: First DIMACS Implementation Challenge, Volume 12 of DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 299–348. American Mathematical Society, 1993.

[44] M. Resende, T. Tsuchiya, and G. Veiga. Identifying the Optimal Face of a Network Linear Program with a Globally Convergent Interior-Point Method. In W. Hager, D. Hearn, and P. Pardalos, eds., *Large-Scale Optimization: State of the Art*, pages 362–387. Kluwer Academic Publishers, 1994.

[45] Roos, Terlaky, and Jansen.

[46] C. M. Samuelsen. The Dikin-Karmarkar Principle for Steepest Descent. Ph.D. thesis, Department of Mathematical Sciences, Rice University, 1992.

[47] R. Skeel. Scaling for Numerical Stability in Gaussian Elimination. *Journal of the Association for Computing Machinery*, 26(3):494–526, 1979.

[48] R. A. Tapia and Y. Zhang. An Optimal Basis Identification Technique for Interior-Point Linear Programming Algorithms. *Linear Algebra and Its Applications*, 152:343–363, 1991.

[49] E. Tardos. A Strongly Polynomial Algorithm to Solve Combinatorial Linear Programs. *Operations Research*, 36(2):250–256, 1986.

[50] S. Vavasis and Y. Ye. Identifying an optimal basis in linear programming. Cornell University, 1995.

[51] P. Williams Effective Finite Termination Procedures in Interior-Point Methods for Linear Programming. Ph.D. thesis. Department of Computational and Applied Mathematics. Rice University.

[52] P. Williams, A. El-Bakry, and R. Tapia. Computing an Exact Solution in Interior-Point Methods for Linear Programming. Technical Report. Department of Computational and Applied Mathematics. Rice University.

[53] S. Wright. Modified Cholesky Factorizations in Interior-Point Algorithms for Linear Programming. Preprint ANL/MCS-P600-0596, Mathematics and Computer Science Division, Argonne National Laboratory.

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

[55] Y. Ye and M. Todd. Containing and Shrinking Ellipsoids in the Path-Following Algorithm. *Mathematical Programming*, 39:305-317, 1987.

[56] Y. Ye Recovering Optimal Basic Variables in Karmarkar's Polynomial Algorithm for Linear Programming. *Mathematics of Operations Research*, 15(3):564-572, 1990.

[57] Y. Ye. On the Finite Convergence of Interior-Point Algorithms for Linear Programming. *Mathematical Programming(Series B)*, 57:325-335,1992.

[58] Y. Ye Toward Probabilistic Analysis of Interior-Point Algorithms for Linear Programming. *Mathematics of Operations Research*, 19(1):38-52, 1994.

[59] Y. Ye. *Interior-Point Algorithm: Theory and Analysis*, John Wiley & Sons, New York, New York, 1997.

[60] Y. Zhang. On the Convergence of a class of Infeasible Interior-Point Methods for the Horizontal Linear Complementarity Problem. *SIAM Journal of Optimization*, 4:208-227, 1994.

[61] Y. Zhang. LIPSOL beta version 2.1. Department of Mathematics and Statistics, University of Maryland Baltimore County, 1995.

[62] Y. Zhang. Solving Large-Scale Linear Programs by Interior-Point Methods Under the MATLAB Environment. Department of Mathematics and Statistics, University of Maryland Baltimore County, 1996.