

SCHOOL OF OPERATIONS RESEARCH  
AND INDUSTRIAL ENGINEERING  
COLLEGE OF ENGINEERING  
CORNELL UNIVERSITY  
ITHACA, NEW YORK 14853-3801

TECHNICAL REPORT NO. 981

October 1991

A POTENTIAL REDUCTION ALGORITHM  
WITH USER-SPECIFIED PHASE I - PHASE II  
BALANCE, FOR SOLVING A LINEAR PROGRAM  
FROM AN INFEASIBLE WARM START

by

Robert M. Freund\*

\*Visiting Cornell from Sloan School of Management, MIT, Cambridge, MA 02139. Research supported in part by NSF, AFOSR, and ONR through NSF grant DMS-8920550.

Abstract:

This paper develops a potential reduction algorithm for solving a linear-programming problem directly from a “warm start” initial point that is neither feasible nor optimal. The algorithm is of an “interior point” variety that seeks to reduce a single potential function which simultaneously coerces feasibility improvement (Phase I) and objective value improvement (Phase II). The key feature of the algorithm is the ability to specify beforehand the desired balance between infeasibility and nonoptimality in the following sense. Given a prespecified balancing parameter  $\beta > 0$ , the algorithm maintains the following Phase I - Phase II “ $\beta$ -balancing constraint” throughout:

$$(c^T x - z^*) < \beta \xi^T x,$$

where  $c^T x$  is the objective function,  $z^*$  is the (unknown) optimal objective value of the linear program, and  $\xi^T x$  measures the infeasibility of the current iterate  $x$ . This balancing constraint can be used to either emphasize rapid attainment of feasibility (set  $\beta$  large) at the possible expense of good objective function values or to emphasize rapid attainment of good objective values (set  $\beta$  small) at the possible expense of a lower infeasibility gap. The algorithm exhibits the following advantageous features: (i) the iterate solutions monotonically decrease the infeasibility measure, (ii) the iterate solutions satisfy the  $\beta$ -balancing constraint, (iii) the iterate solutions achieve constant improvement in both Phase I and Phase II in  $O(n)$  iterations, (iv) there is always a possibility of finite termination of the Phase I problem, and (v) the algorithm is amenable to acceleration via line-search of the potential function.

Key words: Linear program, potential function, interior-point algorithm, polynomial-time complexity.

Running Header: balanced “warm start” algorithm for LP

## 1. Introduction

This paper is concerned with the problem of solving a linear programming problem directly from an infeasible “warm start” solution that is hopefully close to both feasibility and to optimality. Quite often in the practice of using a linear programming model, a practitioner needs to solve many slightly-altered versions of the same base case model. It makes sense in this scenario that the optimal solution (or optimal basis) of a previous version of the linear programming model ought to serve as an excellent “warm start” starting point for the current version of the model, if the two versions of the model are similar. Experience with the simplex method over the years has borne this out to be true in practice; the optimal basis for a previous version of the model usually serves as an excellent starting basis for the next version of the model, even when this basis is infeasible. Intuitively, a good “warm start” infeasible solution (that is not very infeasible and whose objective value is not far from optimality) should give an algorithm valuable information and should be a good starting point for an algorithm that will solve the linear programming model to feasibility and optimality. In spite of the success of “warm start” solutions in solving linear programming problems efficiently with the simplex method, there is no underlying complexity analysis that guarantees faster running times for such starting solutions. This is due to the inevitable combinatorial aspects of the simplex method itself.

In the case interior point algorithms for linear programming, the research on algorithms for solving a linear programming directly from an infeasible “warm start” are part of the research on combined Phase I - Phase II methods for linear programming. The underlying strategy in a combined Phase I - Phase II algorithm is to simultaneously work on the Phase I problem (to attain feasibility) and the Phase II problem (to attain optimality). The starting point for such an algorithm then need not be feasible, and a “warm start” starting point again should serve as an excellent starting point for a combined Phase I - Phase II algorithm. Perhaps the first interior-point combined Phase I - Phase II algorithm is de Gellinck and Vial [9]. Anstreicher [1] also contributed to the early literature in this area, see also Todd [15] and Todd and Wang [16]. These approaches all used the strategy of potential reduction and projective transformations, as originally developed by Karmarkar [12]. Other approaches to the problem using trajectories of optimal solutions to parametric families of shifted barrier problems

were studied by Gill et. al. [10], [7], and Polyak [13]. Later, after direct potential reduction methods were developed by Gonzaga [11], Ye [20], and [6], these methods were extended to the combined Phase I - Phase II problem, see [8], Anstreicher [2], and Todd [17].

While all of these algorithms simultaneously solve the Phase I and Phase II problems, they are all interior point algorithms and so they are only guaranteed to converge to a solution. The algorithm is terminated in theory after the appropriate gap (feasibility gap for Phase I, duality gap for Phase II) is less than  $2^{-L}$ , where  $L$  is the bitsize representation of the problem data, and is terminated in practice when this gap is less than some prescribed small number, e.g.,  $10^{-6}$ .

The formulation of the Phase I - Phase II problem that has been developed by Anstreicher [1,2] is to solve the linear program:

$$\begin{aligned} & \underset{\mathbf{x}}{\text{minimize}} \quad \mathbf{c}^T \mathbf{x} \\ & \text{s.t.} \quad \mathbf{A} \mathbf{x} = \mathbf{b} \\ & \quad \quad \quad \boldsymbol{\xi}^T \mathbf{x} = 0 \\ & \quad \quad \quad \mathbf{x} \geq 0, \end{aligned}$$

where we are given an infeasible “warm start” vector  $\mathbf{x}^0$  that is feasible for the following Phase I problem:

$$\begin{aligned} & \underset{\mathbf{x}}{\text{minimize}} \quad \boldsymbol{\xi}^T \mathbf{x} \\ & \text{s.t.} \quad \mathbf{A} \mathbf{x} = \mathbf{b} \\ & \quad \quad \quad \boldsymbol{\xi}^T \mathbf{x} \geq 0 \\ & \quad \quad \quad \mathbf{x} \geq 0, \end{aligned}$$

and has the Phase I objective value  $\boldsymbol{\xi}^T \mathbf{x}^0 > 0$ . If  $\mathbf{z}^*$  is the optimal value of LP, then  $\mathbf{c}^T \mathbf{x} - \mathbf{z}^*$  measures the optimal value gap, and  $\boldsymbol{\xi}^T \mathbf{x}$  measures the feasibility gap.

The Balance of Priorities between Phase I and Phase II. One might think that in solving any linear programming problem, that both Phase I and Phase II are equally important: for surely, without feasibility, the problem is not solved, and without optimality, the problem is not optimized. However, in practice, there are many instances where this simple logic breaks down, and that different problems naturally lend themselves to very different ways of prioritizing the balance between improving the Phase I objective, i.e., reducing the feasibility gap, and improving the Phase II objective, i.e., improving the optimal value gap. Consider the following list of instances:

(i) In some practical modelling problems, the constraints of the problem are specified easily, but the objective function is not so easy to specify. This may be because of accounting criteria and the problem of ascertaining the true “variable cost” of the activities. Or it may be because it is not clear just what the actual objective is in the practical problem. In these problems, attaining feasibility is important, but attaining exact optimality is not so important, because of a lack of confidence that the linear programming objective function is a good representation of the true objective of the underlying practical problem.

(ii) There are instances of practical situations where the user is primarily interested in obtaining a feasible solution, and the objective function is not very important. In these instances, much more priority should be given to Phase I than to Phase II.

(iii) In other practical modelling problems, a feasible solution that is not optimal may be of no use at all. This type of situation arises frequently when using linear programming to solve partial equilibrium economic models, see, e.g., Wagner [19]. In these models, feasibility may be easy to attain, but the partial equilibrium solution is obtained by looking at both the primal and the prices that arise as the solution to the dual problem. A nonoptimal primal feasible solution conveys virtually no information about the underlying economic model. In this application of linear programming, Phase II should achieve a much higher priority than it would in instances (i) or (ii) above.

(iv) In using linear programming in branch and bound routines for solving mixed-integer programming problems, a sequence of linear programs is generated and solved as the branch and bound

routine runs its course. When solving a particular one of these linear programs, we may only be interested in looking at the bounds generated by the algorithm. In this case, attaining a feasible solution may be completely unnecessary, and it may suffice to generate a bound that is sufficiently positive to signal that the branch of the underlying tree should be pruned. In this case, attaining feasibility may be unimportant, and should receive much less priority than it would in instances (i) and (ii) above.

These instances suggest that an algorithm for the combined Phase I - Phase II problem should have as a parameter some measure of the relative importance or “balance” between the goals of reducing the feasibility gap (Phase I) and reducing the optimal value gap (Phase II) in solving a given linear programming problem. In this paper, we propose a measure of this balance concept, a parameter for setting this measure for a particular problem, and a polynomial-time algorithm for solving the linear programming problem from an infeasible “warm start” that incorporates this measure and parameter into the algorithm. The notion that we develop herein is denoted as “ $\beta$ -balancing” and is developed as follows.

Let  $\beta$  be a positive scalar constant that is specified by the user, called the “balancing parameter.” Given the prespecified balancing parameter  $\beta > 0$ , the algorithm maintains the following Phase I -Phase II “ $\beta$ -balancing constraint” throughout:

$$(c^T x - z^*) < \beta \xi^T x, \quad (1.1)$$

where  $c^T x$  is the objective function,  $z^*$  is the (unknown) optimal objective value of the linear program, and  $\xi^T x$  measures the infeasibility of the current iterate  $x$ . The left side of (1.1) is the optimal value gap, and the right side is  $\beta$  times the feasibility gap. Thus (1.1) states that the optimal value gap must be less than or equal to  $\beta$  times the feasibility gap.

If  $\beta$  is set to be very large, then (1.1) does not coerce a very tight optimal value gap. And so even when the feasibility gap is small, the optimal value gap can still be quite large (although when the

feasibility is zero, clearly from (1.1) the optimal value gap must also be zero.) Thus the larger the value of  $\beta$ , the more Phase II is de-emphasized, i.e., the more Phase I is emphasized.

If  $\beta$  is set to be very small, then (1.1) coerces a very tight (or negative) optimal value gap as the feasibility gap is narrowed to zero. And so even for a relatively large infeasibility gap, the optimal value gap must be small (or even negative) in order to satisfy (1.1). Thus the smaller the value of  $\beta$ , the more Phase II is emphasized in the algorithm.

The algorithm developed in this paper also has the following other desirable features: (i) the iterate solutions monotonically decrease the infeasibility gap, (ii) the iterate solutions satisfy the  $\beta$ -balancing constraint (1.1), (iii) the iterate solutions achieve constant improvement in both Phase I and Phase II in  $O(n)$  iterations, (iv) there is a possibility of finite termination of the Phase I problem (whether or not the objective values are superoptimal), and (v) the algorithm is amenable to acceleration via line-search of the potential function.

The paper is organized as follows. In Section 2, the notation used in the paper is presented, the formulation of the “warm start” problem is presented, and the  $\beta$ -balancing constraint is developed and discussed further. Also, we show how to convert any linear programming problem with an infeasible “warm start” and an initial objective function lower bound into the standard form that also satisfies the  $\beta$ -balancing constraint (1.1). Section 3 contains the development of the potential reduction problem that will be used to solve the linear programming problem, and contains convergence properties of the potential reduction problem. Section 4 describes the algorithm that is used to solve (in a limiting sense) the potential reduction problem developed in Section 3. Section 5 discusses modifications and enhancements to the algorithm of Section 4, that are designed to speed convergence and give more useful information. In particular, Section 5 discusses ways to accelerate the algorithm via line-searches, improved dual updates via Fraley’s restriction of the dual problem [5], finite termination of the Phase I problem, and obtaining explicit convergence constants related to the potential function.

## 2. Notation, Problem Formulation, and Conversions

### Notation

Throughout the paper,  $\mathbf{e}$  denotes the vector of ones,  $\mathbf{e} = (1, 1, \dots, 1)^T$ , where the dimension is  $n$ . For any vector  $\bar{\mathbf{x}}$ , etc.,  $\bar{\mathbf{X}}$  denotes the diagonal matrix whose diagonal components correspond to  $\bar{\mathbf{x}}$ . If  $\mathbf{v} \in \mathbb{R}^n$ ,  $\|\mathbf{v}\|$  denotes the Euclidean norm, i.e.,  $\|\mathbf{v}\| = (\sum_{j=1}^n v_j^2)^{\frac{1}{2}}$ .

### Problem Formulation and the $\beta$ -Balancing Constraint

The combined Phase I - Phase II linear programming problem is usually expressed in the format:

$$\text{LP:} \quad z^* = \underset{\mathbf{x}}{\text{minimize}} \quad \mathbf{c}^T \mathbf{x} \quad (2.1a)$$

$$\text{s.t.} \quad \mathbf{A} \mathbf{x} = \mathbf{b} \quad (2.1b)$$

$$\boldsymbol{\xi}^T \mathbf{x} = 0 \quad (2.1c)$$

$$\mathbf{x} \geq 0, \quad (2.1d)$$

where we assume that there is a given infeasible “warm start” vector  $\mathbf{x}^0$  that satisfies  $\mathbf{A} \mathbf{x}^0 = \mathbf{b}$  (2.1b),  $\mathbf{x}^0 \geq 0$  (2.1d), but for which  $\boldsymbol{\xi}^T \mathbf{x}^0 > 0$ , see [1], [2], and [16]. Thus  $\mathbf{x}^0$  is “almost feasible” for LP, and the extent to which  $\mathbf{x}^0$  is infeasible is precisely the quantity  $\boldsymbol{\xi}^T \mathbf{x}^0$ . (At the end of this section, we will show how to convert any linear programming problem with an initial infeasible warm start into an instance of LP above.) Considering LP as the Phase-II problem, the Phase-I problem for LP then is the problem:

$$\text{P1:} \quad \underset{\mathbf{x}}{\text{minimize}} \quad \boldsymbol{\xi}^T \mathbf{x} \quad (2.2a)$$

$$\text{s.t.} \quad \mathbf{A} \mathbf{x} = \mathbf{b} \quad (2.2b)$$

$$\boldsymbol{\xi}^T \mathbf{x} \geq 0 \quad (2.2c)$$

$$\mathbf{x} \geq 0, \quad (2.2d)$$

and now note that  $x^0$  is feasible (but not optimal) for P1. We also assume that we are given an initial lower bound  $B^0$  on the optimal value  $z^*$  of LP, i.e.  $B^0 \leq z^*$ . Such a bound may be readily available, or can be produced by the algorithm in Todd [17].

In the design of an algorithm for solving LP and P1, that will produce iterate values  $x^1, x^2, x^3, \dots$ , we would like  $\xi^T x^k \rightarrow 0$  as  $k \rightarrow \infty$ , i.e., the iterates converge to a feasible solution to LP (and solve P1 to optimality). We also would like  $c^T x^k \rightarrow z^*$  as  $k \rightarrow \infty$ , i.e., the iterates' objective values converge to the optimal objective value. Let  $\beta > 0$  be a given (user-specified) “balancing parameter” that will be used to enforce the following Phase I - Phase II balancing condition at each iteration:

$$c^T x^k - z^* < \beta \xi^T x^k. \quad (2.3)$$

The left side of (2.3) is the optimal objective value gap at iteration  $k$ , and the right side is  $\beta$  times the feasibility gap. Thus (2.3) states that the optimal objective value gap must be less than  $\beta$  times the feasibility gap. An alternate way to write (2.3) is

$$\frac{c^T x^k - z^*}{\xi^T x^k} < \beta. \quad (2.4)$$

In this form, we see that the ratio of the optimal objective value gap to the infeasibility gap cannot exceed  $\beta$ .

If  $\beta$  is given and (2.3) is enforced throughout the algorithm, then  $\beta$  acts as a pre-specified balancing factor that will bound the optimal objective value gap in terms of the feasibility gap. For example, if feasibility is much more important than optimality, then  $\beta$  can be chosen to be a large number ( $\beta = 1,000$ , for example) whereby from (2.3) we see that the feasibility gap does not coerce a small optimal value gap. If on the other hand, staying near the optimal objective value is more important, then  $\beta$  can be chosen to be a small number ( $\beta = 0.001$ , for example). Then from (2.3), the feasibility gap does coerce a small optimal objective value gap. From (2.3), we see that at iteration

$k$ , the deviation from the optimal objective value  $(c^T x^k - z^*)$  is bounded in terms of the extent of infeasibility  $(\xi^T x^k)$  by the constant  $\beta$ , i.e.,

$$c^T x^k - z^* < \beta \xi^T x^k.$$

However,  $z^*$  is not known in advance; only a lower bound  $B^0$  on  $z^*$  is known in advance. The algorithm developed in this paper will produce an increasing sequence of bounds  $B^1, B^2, \dots$ , on  $z^*$ , where  $B^k$  is the bound produced at iteration  $k$ . Since we do not know  $z^*$  in advance, the algorithm will enforce the following balancing condition:

$$c^T x^k - B^k < \beta \xi^T x^k, \tag{2.5}$$

where (2.5) is identical to (2.3) except that  $z^*$  is replaced by the bound  $B^k$ . Note that since  $B^k \leq z^*$ , then (2.5) implies (2.3), i.e.

$$c^T x^k - z^* < \beta \xi^T x^k \text{ whenever } c^T x^k - B^k < \beta \xi^T x^k.$$

We can rearrange (2.5) into the more standard format:

$$(-\beta \xi + c)^T x^k < B^k. \tag{2.6}$$

We refer to (2.6) as the “ $\beta$ -balancing constraint” at iteration  $k$ . In order to satisfy (2.6) at the start of the algorithm, we will need the initial assumption that  $(-\beta \xi + c)^T x^0 < B^0$ . (At the end of this section, we will show how to convert any linear programming problem with an infeasible warm start  $x^0$  into an instance of LP for which  $(-\beta \xi + c)^T x^0 < B^0$  is satisfied). We now summarize the data and other assumptions we will use for the rest of this study:

A(i) The given data for LP is the array  $(A, \xi, b, c, x^0, B^0, \beta)$

A(ii)  $Ax^0 = b, \xi^T x^0 > 0, x^0 > 0, B^0 \leq z^*$

A(iii)  $\beta > 0$  and  $(-\beta\xi + c)^T x^0 < B^0$

A(iv) The set of optimal solutions of LP is a bounded set.

A(v)  $n \geq 3$ .

Assumptions A(i)-A(iii) have been reviewed above. Assumption A(iv) is a standard (though nontrivial) assumption needed for convergence of all interior-point algorithms. (See Vial [18] and Anstreicher [4] for ways to mitigate this assumption.) Assumption A(v) is trivial, since for  $n \leq 2$  the problem LP lends itself to instant analysis. We now show how to convert a linear program satisfying A(iv) and A(v) into the standard form LP of (2.1) and that satisfies all assumptions A(i)-A(v).

Converting a Linear Program into an Instance of LP Satisfying A(i)-A(v)

Suppose we want to solve the linear program:

$$\hat{LP}: \quad z^* = \min_{\hat{x}} \hat{c}^T \hat{x} \quad (2.7a)$$

$$\text{s.t. } \hat{A} \hat{x} = b \quad (2.7b)$$

$$\hat{x} \geq 0, \quad (2.7c)$$

where  $\hat{A}$  is  $m \times n$  and it is assumed that  $n \geq 3$ , and  $\hat{x}^0$  is a given “warm start” that is hopefully near-feasible and near-optimal. Also suppose that  $B^0$  is a known given lower bound on  $z^*$ . Then the given data for the problem  $\hat{LP}$  is the array  $(\hat{A}, b, \hat{c}, \hat{x}^0, B^0)$ . In a typical situation,  $\hat{x}^0$  may be the optimal solution to a previous version of  $\hat{LP}$  that is hopefully a good near-feasible and near-optimal for the current linear program  $\hat{LP}$ . Alternatively,  $\hat{x}^0$  may be a basic solution to  $\hat{LP}$  for a basis that is suspected of being close to the optimal basis. Knowledge of  $B^0$  can be given in a number of ways. If  $L$  is the size of the array  $(A, b, c)$  (i.e.,  $L$  is the number of bits needed to encode the data  $(A, b, c)$  in binary form), then one value of  $B^0$  that can be used is  $-2^L$ , but this is not practical. A more

practical approach would be to set  $B^0$  to be some large negative number such as  $-10^{12}$ . However, if the user has a good knowledge of the program  $\hat{LP}$ , he/she may be able to set  $B^0$  fairly accurately. (For example, suppose  $\hat{LP}$  is a refinery problem. Then it is reasonable that a lower bound on the cost of operating the refinery is readily apparent from knowledge of the data that have been used to generate the program  $\hat{LP}$ ). It should also be pointed out that an algorithm for generating a reasonable bound  $B^0$  has been developed in Todd [17].

We first assume that  $\hat{x}^0$  satisfies the equations 2.7b, i.e.,  $\hat{A}\hat{x}^0 = b$ . This will certainly be the case if  $\hat{x}^0$  is a basic solution for a (hopefully near-optimal) basis of  $\hat{A}$ . If  $\hat{x}^0$  does not satisfy (2.7b), then  $\hat{x}^0$  can be projected onto the linear manifold  $\{\hat{x} | \hat{A}\hat{x} = b\}$  by choosing any suitable projection, e.g.,

$$\hat{x}^0 \leftarrow D[I - D\hat{A}^T(\hat{A}D^2\hat{A}^T)^{-1}\hat{A}D]\hat{x}^0 + D^2\hat{A}^T(\hat{A}D^2\hat{A}^T)^{-1}b,$$

where  $D$  is any positive-definite matrix (e.g.,  $D = I$  or  $D$  is a positive-diagonal matrix). It is assumed that  $\hat{x}^0 \not\geq 0$ , for otherwise  $\hat{x}^0$  would be an interior feasible solution to (2.7) and there would be no need for a Phase I procedure to be part of the solution to (2.7).

Now let  $h \geq 0$  be any vector that satisfies  $\hat{x}^0 + h \geq 0$ . Then our problem  $\hat{LP}$  is equivalent to:

$$\hat{LP}^2: \quad \min_{\hat{x}, \hat{w}} c^T \hat{x} \quad (2.8a)$$

$$\text{s.t. } \hat{A}\hat{x} = b \quad (2.8b)$$

$$\hat{x} + \hat{w}h \geq 0 \quad (2.8c)$$

$$\hat{w} = 0, \quad (2.8d)$$

where we note that  $(\hat{x}, \hat{w}) = (\hat{x}^0, 1)$  is feasible for (2.8) except for the last constraint (2.8d), which measures the infeasibilities of  $\hat{x}^0$ . If  $\hat{LP}^2$  is the Phase-II problem, then the Phase-I problem can be written as:

$$\hat{\text{LP}}^1: \quad \underset{\hat{\mathbf{x}}, \hat{\mathbf{w}}}{\text{minimize}} \quad \hat{\mathbf{w}} \quad (2.9\text{a})$$

$$\text{s.t.} \quad \hat{\mathbf{A}}\hat{\mathbf{x}} = \mathbf{b} \quad (2.9\text{b})$$

$$\hat{\mathbf{x}} + \hat{\mathbf{w}}\mathbf{h} \geq 0 \quad (2.9\text{c})$$

$$\hat{\mathbf{w}} \geq 0. \quad (2.9\text{d})$$

Notice that  $(\hat{\mathbf{x}}, \hat{\mathbf{w}}) = (\hat{\mathbf{x}}^0, 1)$  is feasible for  $\hat{\text{LP}}^1$ , and in fact,

$$(\hat{\mathbf{x}}, \hat{\mathbf{w}}) = (\hat{\mathbf{x}}^0, \epsilon)$$

is feasible for  $\hat{\text{LP}}^1$  for any  $\epsilon \geq 1$ , due to the fact that  $\mathbf{h} \geq 0$ .

Let  $\beta$  be the pre-specified balancing parameter discussed previously. Then if  $\hat{\mathbf{w}}^0$  is given by:

$$\hat{\mathbf{w}}^0 = \max\left\{1, 1 + \frac{(\hat{\mathbf{c}}^\top \hat{\mathbf{x}}^0 - \mathbf{B}^0)}{\beta}\right\}, \quad (2.10)$$

then  $(\hat{\mathbf{x}}^0, \hat{\mathbf{w}}^0)$  will satisfy

$$\hat{\mathbf{A}}\hat{\mathbf{x}}^0 = \mathbf{b} \quad (2.11\text{a})$$

$$\hat{\mathbf{x}}^0 + \hat{\mathbf{w}}^0\mathbf{h} > 0 \quad (2.11\text{b})$$

$$(-\beta\hat{\mathbf{w}}^0 + \hat{\mathbf{c}}^\top \hat{\mathbf{x}}^0) < \mathbf{B}^0 \quad (2.11\text{c})$$

Thus the pair  $(\hat{\mathbf{x}}^0, \hat{\mathbf{w}}^0)$  is feasible for the Phase-I problem  $\hat{\text{LP}}^1$  (2.9) and also satisfies (2.11c), which is the analog of Assumption A(iii) for this problem. Also,  $(\hat{\mathbf{x}}^0, \hat{\mathbf{w}}^0)$  satisfies all constraints of  $\hat{\text{LP}}^2$  (2.8) except (2.8d).

In order to convert  $\hat{\text{LP}}^2$  (2.8) to an instance of LP (2.1), we proceed as follows. First, let  $\mathbf{x}$  denote the slack vector

$$\mathbf{x} = \hat{\mathbf{x}} + \hat{\mathbf{w}}\mathbf{h}. \quad (2.12)$$

Then  $\mathbf{x}^0 = \hat{\mathbf{x}}^0 + \hat{\mathbf{w}}^0\mathbf{h} > 0$  denotes the starting slack variables for the starting solution  $(\hat{\mathbf{x}}^0, \hat{\mathbf{w}}^0)$  of  $\hat{\text{LP}}^1$  (2.9).

The next step is to eliminate the variable  $\hat{\mathbf{w}}$  from the systems (2.8) and (2.9). To do this, assume with no loss of generality that the vectors  $\hat{\mathbf{A}}\mathbf{h}$  and  $\mathbf{b}$  are not linearly independent. (If this is not the case, a perturbation of  $\mathbf{h} > 0$  will enforce their linear independence). Then let  $\boldsymbol{\lambda} \in \mathbb{R}^m$  be any vector for which

$$\boldsymbol{\lambda}^T \mathbf{b} = 0$$

$$\boldsymbol{\lambda}^T (\hat{\mathbf{A}}\mathbf{h}) = 1,$$

(such a  $\boldsymbol{\lambda}$  is simple to compute), and let

$$\boldsymbol{\xi} = \hat{\mathbf{A}}^T \boldsymbol{\lambda}.$$

Then note that all  $\mathbf{x}, \hat{\mathbf{x}}, \hat{\mathbf{w}}$  that satisfy (2.8b) and (2.12) satisfy

$$\boldsymbol{\xi}^T \mathbf{x} = \boldsymbol{\lambda}^T \hat{\mathbf{A}}\mathbf{x} = \boldsymbol{\lambda}^T \hat{\mathbf{A}}(\hat{\mathbf{x}} + \hat{\mathbf{w}}\mathbf{h}) = \boldsymbol{\lambda}^T \mathbf{b} + \boldsymbol{\lambda}^T \hat{\mathbf{A}}\mathbf{h}\hat{\mathbf{w}} = \hat{\mathbf{w}}. \quad (2.13)$$

Thus we can substitute  $\hat{\mathbf{x}}$  and  $\hat{\mathbf{w}}$  by  $\mathbf{x}$  (from 2.12) and  $\boldsymbol{\xi}^T \mathbf{x}$  (from 2.13) in (2.8) and (2.9). If we define

$$\mathbf{c} = \hat{\mathbf{c}} - \hat{\mathbf{c}}^T \mathbf{h} \boldsymbol{\xi} \quad \text{and} \quad \mathbf{A} = \hat{\mathbf{A}} - \hat{\mathbf{A}}\mathbf{h}\boldsymbol{\xi}^T,$$

then  $\hat{\text{LP}}^2$  transforms to:

$$\begin{aligned} \text{LP:} \quad & \underset{\mathbf{x}}{\text{minimize}} \quad \mathbf{c}^T \mathbf{x} \\ & \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \quad \mathbf{x} \geq 0 \\ & \quad \boldsymbol{\xi}^T \mathbf{x} = 0, \end{aligned}$$

and  $\hat{\text{LP}}^1$  transforms to:

$$\begin{aligned}
\text{P1:} \quad & \underset{\mathbf{x}}{\text{minimize}} \quad \xi^T \mathbf{x} \\
& \text{s.t.} \quad \mathbf{A} \mathbf{x} = \mathbf{b} \\
& \quad \mathbf{x} \geq 0 \\
& \quad \xi^T \mathbf{x} \geq 0,
\end{aligned}$$

Also,  $\mathbf{x}^0 = \hat{\mathbf{x}}^0 + \hat{\mathbf{w}}^0 \mathbf{h}$  satisfies  $\mathbf{A} \mathbf{x}^0 = \mathbf{b}$ ,  $\mathbf{x}^0 > 0$ ,  $\xi^T \mathbf{x}^0 = \hat{\mathbf{w}}^0 > 0$ , and (2.11c) transforms to

$$(-\beta \xi + \mathbf{c})^T \mathbf{x}^0 < \mathbf{B}^0. \quad (2.14)$$

Note that  $\xi^T \mathbf{x}^0 = \hat{\mathbf{w}}^0 = \max\left\{1, 1 + \frac{(\hat{\mathbf{c}}^T \hat{\mathbf{x}}^0 - \hat{\mathbf{B}}^0)}{\beta}\right\}$ , from (2.10). Also, if  $\hat{\mathbf{L}}\mathbf{P}^2$  satisfies A(iv), then it is easy to verify that LP does as well.

### 3. The Potential Reduction Problem for Solving LP and Convergence Properties

In this section we consider solving the “standard form” problem

$$\begin{aligned}
\text{LP:} \quad & z^* = \min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \\
& \mathbf{A} \mathbf{x} = \mathbf{b} \\
& \xi^T \mathbf{x} = 0, \\
& \mathbf{x} \geq 0,
\end{aligned} \quad (3.1)$$

whose dual is

$$\begin{aligned}
\text{LD:} \quad & \max_{\pi, \theta, \mathbf{s}} \mathbf{b}^T \pi \\
& \mathbf{A}^T \pi + \xi \theta + \mathbf{s} = \mathbf{c} \\
& \mathbf{s} \geq 0.
\end{aligned} \quad (3.2)$$

It is assumed that the data array  $(A, \xi, b, c, x^0, B^0, \beta)$  satisfies assumptions A(i)-A(vi) of the previous section.

The Phase-I problem for LP then is to solve

$$\begin{aligned}
 \text{P1:} \quad & \underset{x}{\text{minimize}} \quad \xi^T x \\
 & \text{s.t.} \quad Ax = b \\
 & \quad \quad \xi^T x \geq 0 \\
 & \quad \quad x \geq 0.
 \end{aligned} \tag{3.3}$$

We will not work with this problem (P1), but will instead augment P1 with the additional balancing constraint involving the lower bound  $B^0$  on the optimal value  $z^*$  of LP and the balancing parameter  $\beta$  discussed in Section 2. Suppose  $B^0$  is the given lower bound on  $z^*$  and that  $\beta$  is the balancing parameter for which the starting point  $x^0$  satisfies

$$(-\beta\xi + c)^T x^0 < B^0, \tag{3.4}$$

see Assumption A(iii). (The method for satisfying (3.4) was discussed in Section 2.) Now consider the parametric family of augmented Phase-I problems

$$\begin{aligned}
 \text{P}_B: \quad & z_B = \underset{x, t}{\text{minimize}} \quad \xi^T x \\
 & \text{s.t.} \quad Ax = b \\
 & \quad \quad (-\beta\xi + c)^T x + t = B \\
 & \quad \quad x \geq 0, t \geq 0
 \end{aligned} \tag{3.5}$$

whose dual is given by:

$$\begin{aligned}
D_B: \quad z_B &= \underset{\pi, \mu, s}{\text{maximize}} \quad b^T \pi - B\mu \\
&A^T \pi - (-\beta\xi + c)\mu + s = \xi \\
&s \geq 0, \mu \geq 0.
\end{aligned} \tag{3.6}$$

The following are elementary properties  $P_B$ :

Proposition 3.1. For all  $B \in [B^0, z^*]$

- (i)  $P_B$  is feasible,
- (ii)  $z_B > 0$  if  $B < z^*$ ,  
 $z_B = 0$  if  $B = z^*$ .
- (iii) The set of optimal solutions to  $P_B$  is nonempty and bounded.
- (iv) For all  $x$  feasible for  $P_B$ ,

$$c^T x \leq z^* + \beta\xi^T x.$$

Proof: (i) Let  $t^0 = B^0 - (-\beta\xi + c)^T x^0$ . From Assumption A(ii) and (3.4) it follows that  $(x^0, t^0)$  is feasible for  $P_{B^0}$  and so  $(x^0, t)$  is also feasible for  $P_B$  for any  $B \geq B^0$ , where  $t = B - (-\beta\xi + c)^T x^0$ .

(ii) Suppose  $B < z^*$ . Then if  $z_B \leq 0$ , there exists  $x$  for which  $Ax = b$ ,  $x \geq 0$ ,  $\xi^T x = 0$ ,  $(-\beta\xi + c)^T x \leq B < z^*$ , and so  $c^T x < z^*$ , violating the definition of  $z^*$ . Thus  $z_B > 0$ . If  $B = z^*$ , then a similar argument establishes that  $z_B = 0$ .

(iii) Suppose the set of optimal solutions to  $P_B$  is not bounded. Then there is a direction  $d \neq 0$  that satisfies  $\xi^T d = 0$ ,  $Ad = 0$ ,  $d \geq 0$ ,  $(-\beta\xi + c)^T d \leq 0$ , and so  $c^T d \leq 0$ . Therefore  $d$  is a nontrivial ray of the optimal solution set of LP, violating the assumption that LP has a bounded set of optimal solutions.

(iv)  $(-\beta\xi + c)^T x \leq B$  implies  $c^T x \leq \beta\xi^T x + B \leq \beta\xi^T x + z^*$ .  $\square$

Now consider the following potential reduction problem related to  $P_B$  and  $D_B$ :

$$\text{PR:} \quad \underset{x, t, B}{\text{minimize}} \quad F(x, t) = q \ln(\xi^T x) - \sum_{j=1}^n \ln x_j - \ln t \quad (3.7)$$

$$\text{s.t.} \quad Ax = b \quad (3.8a)$$

$$(-\beta \xi + c)^T x + t = B \quad (3.8b)$$

$$x > 0, t > 0 \quad (3.8c)$$

$$B \leq z^*, \quad (3.9)$$

where  $q$  is a parameter satisfying  $q \geq n+1$ , and (3.8) reflects feasibility for  $P_B$ , for  $B \leq z^*$ , which is given in (3.9). The following lemma relates potential function values to the objective function  $\xi^T x$  of problem  $P_B$ .

Lemma 3.1. Suppose  $(x^0, t^0, B^0)$  is the starting point of an algorithm for solving PR and suppose that  $z^* < +\infty$ . Suppose  $(\bar{x}, \bar{t}, \bar{B})$  is a feasible point generated by the algorithm, and that  $\xi^T \bar{x} \leq \xi^T x^0$ . Let

$$\Delta = F(x^0, t^0) - F(\bar{x}, \bar{t}). \quad (3.10)$$

Then

$$\xi^T \bar{x} \leq (\xi^T x^0) C_1 e^{-\Delta/q}$$

where  $C_1$  is computed from the optimal value of the following linear program:

$$\text{PC:} \quad (C_1)^{\left(\frac{q}{n+1}\right)} = (n+1)^{-1} \max_{x, t} e^T (X^0)^{-1} x + (t^0)^{-1} t \quad (3.11a)$$

$$\text{s.t.} \quad Ax = b \quad (3.11b)$$

$$(-\beta \xi + c)^T x + t \leq z^* \quad (3.11c)$$

$$\xi^T x \leq \xi^T x^0 \quad (3.11d)$$

$$-\xi^T x \leq 0 \quad (3.11e)$$

Proof: The pair  $(x^0, t^0)$  is feasible for PC. If PC is unbounded, then there exists  $(d, v)$  satisfying  $d \geq 0$ ,  $v \geq 0$ ,  $Ad = 0$ ,  $(-\beta\xi + c)^T d + v \leq 0$ ,  $\xi^T d \leq 0$ ,  $-\xi^T d \leq 0$ , and  $e^T (X^0)^{-1} d + (t^0)^{-1} v > 0$ . Thus  $d \geq 0$ ,  $Ad = 0$ ,  $\xi^T d = 0$ ,  $c^T d + v \leq 0$ ,  $v \geq 0$ , and so  $c^T d \leq 0$  and  $d \neq 0$ , contradicting the assumption that the set of optimal solutions to LP is a bounded set. Thus  $C_1$  is well-defined, and  $0 < C_1 < +\infty$ .

Since  $(\bar{x}, \bar{t})$  is also feasible for PC,

$$e^T (X^0)^{-1} \bar{x} + (t^0)^{-1} \bar{t} \leq (n+1)(C_1)^{\left(\frac{q}{n+1}\right)}$$

Thus by the arithmetic-geometric mean inequality, it then follows that

$$\sum_{j=1}^n \ell n \bar{x}_j + \ell n \bar{t} - \sum_{j=1}^n \ell n x_j^0 - \ell n t^0 \leq q \ell n C_1. \quad (3.12)$$

Next, from (3.10) and (3.7), we have

$$\begin{aligned} q \ell n (\xi^T \bar{x}) &= q \ell n (\xi^T x^0) + \sum_1 \ell n \bar{x}_j + \ell n \bar{t} - \sum_1 \ell n x_j^0 - \ell n t^0 - \Delta \\ &\leq q \ell n (\xi^T x^0) + q \ell n C_1 - \Delta, \end{aligned}$$

where the inequality follows from (3.12). Exponentiating and rearranging yields the result.  $\square$

Note that if the size of  $\beta$ ,  $x^0$ , and  $t^0$  are  $O(L)$ , then  $C_1 \leq 2^L$ , provided that  $z^* < +\infty$ .

We will demonstrate an algorithm in Section 4 that will reduce  $F(x, t)$  by a fixed constant  $\delta \geq \frac{1}{6}$  at each iteration, if  $q \geq n + 1 + \sqrt{n+1}$ , with the additional property that the values of  $\xi^T x$  monotonically decrease at each iteration. This will be the basis for the following convergence theorem:

Theorem 3.1 (Convergence). Suppose  $(x^k, t^k, B^k)$ ,  $k = 0, \dots$ , is a sequence of feasible solutions to problem PR with the property that  $F(x^{k+1}, t^{k+1}) \leq F(x^k, t^k) - \frac{1}{6}$ , and  $\xi^T x^{k+1} \leq \xi^T x^k$ ,  $k = 0, 1, \dots$ . Suppose that  $z^* < +\infty$ . Then with  $C_1$  as given in (3.11),

$$(i) \quad 0 < \xi^T x^k \leq (\xi^T x^0) C_1 e^{-k/6q}.$$

Let  $(\pi^*, \theta^*, s^*)$  be any optimal solution to LD. Then

$$(ii) \quad -|\theta^*|(\xi^T x^0) C_1 e^{-k/6q} \leq c^T x^k - z^* \leq \beta(\xi^T x^0) C_1 e^{-k/6q}$$

$$(iii) \quad -|\theta^*|(\xi^T x^0) C_1 e^{-k/6q} \leq c^T x^k - B^k \leq \beta(\xi^T x^0) C_1 e^{-k/6q}$$

$$(iv) \quad 0 \leq z^* - B^k \leq (\beta + |\theta^*|) C_1 e^{-k/6q}.$$

Theorem 3.1(i) states that fixed improvement in the phase-I objective value  $\xi^T x^k$  is obtained in  $O(q)$  iterations. The convergence results in Theorem 3.1(ii) relate the convergence of the phase-II objective value  $c^T x^k$  to the optimal value  $z^*$ . Similar convergence results for the lower bounds  $B^k$  are given in (iii) and (iv) of the theorem.

Proof of Theorem 3.1: Letting  $(\bar{x}, \bar{t}) = (x^k, t^k)$ , (i) follows from Lemma 3.1, where from (3.10)

$\Delta \geq k/6$ . From the convexity properties of linear programming duality, we obtain from Proposition A.3 of the Appendix that

$$c^T x \geq z^* + \theta^* \xi^T x \text{ for any } x \text{ satisfying } Ax = b, x \geq 0, \quad (3.15)$$

$$\text{and so} \quad c^T x^k - B^k \geq c^T x^k - z^* \geq \theta^* \xi^T x^k \geq -|\theta^*|(\xi^T x^0) C_1 e^{-k/6q}. \quad (3.16)$$

Furthermore, from (3.8b), we obtain

$$c^T x^k - z^* \leq c^T x^k - B^k \leq \beta \xi^T x^k \leq \beta(\xi^T x^0) C_1 e^{-k/6q}, \quad (3.17)$$

and (3.16) and (3.17) combine to prove (ii) and (iii). (iv) is a consequence of (ii) and (iii).  $\square$

#### 4. The Algorithm for Solving the Potential Reduction Problem PR

In this section, we present an algorithm that obtains a decrease of  $\delta \geq \frac{1}{6}$  in the potential function  $F(x, t)$  of problem PR (3.7-3.9) at each iteration, and that is monotone decreasing in the values of  $\xi^T x$ , given  $q \geq n + 1 + \sqrt{n+1}$ .

Suppose the current iterate values for PR is the array  $(\bar{x}, \bar{t}, \bar{B})$ , which is feasible for PR. As in the standard potential reduction algorithm (see Ye [20], Gonzaga [11], [6], and Anstreicher [3]), we seek to compute a primal direction that will decrease the potential function. Since the primal variables are  $(x, t) = (\bar{x}, \bar{t})$ , we seek a direction  $(\tilde{d}, \tilde{r})$  and a suitable step-length  $\alpha$  for which  $F(\bar{x} - \alpha\tilde{d}, \bar{t} - \alpha\tilde{r})$ , achieves a constant decrease over  $F(\bar{x}, \bar{t})$ . Analogous to [20], [6], and [3], we let  $(\tilde{d}, \tilde{r})$  be the solution to the following optimization problem:

$$Q: \quad \underset{d, r}{\text{maximize}} \left( \frac{q}{\xi^T \bar{x}} \xi^T - e^T \bar{X}^{-1} \right) d - \bar{t}^{-1} r - \frac{1}{2} d^T \bar{X}^{-2} d - \frac{1}{2} \bar{t}^{-2} r^2 \quad (4.1a)$$

$$\text{s.t. } A d = 0 \quad (\pi) \quad (4.1b)$$

$$(-\beta \xi + c)^T d + r = 0 \quad (-\theta) \quad (4.1c)$$

$$\xi^T d \geq 0 \quad (\delta), \quad (4.1d)$$

where the quantities  $(\pi)$ ,  $(-\theta)$ ,  $(\delta)$  indicated are the dual multipliers on the constraints.

This problem has a strictly concave quadratic objective, and since  $(d, r) = (0, 0)$  is a feasible solution, it will attain its optimum uniquely. Program Q can be interpreted as the standard rescaled projection of the rescaled gradient of the potential function onto the null space of the equations (3.8a-b), with the simple monotonicity constraint (see Anstreicher [3]) added as well in (4.1d). The unique solution  $(\tilde{d}, \tilde{r})$  to Q is obtained by solving the following Karush-Kuhn-Tucker conditions for  $(\tilde{d}, \tilde{r}, \tilde{\pi}, \tilde{\theta}, \tilde{\delta})$ :

$$A \tilde{d} = 0 \quad (4.2a)$$

$$(-\beta \xi + c)^T \tilde{d} + \tilde{r} = 0 \quad (4.2b)$$

$$\xi^T \tilde{d} \geq 0 \quad (4.2c)$$

$$\frac{q}{\xi^T \bar{x}} \xi - \bar{X}^{-1} e - \bar{X}^{-2} \tilde{d} = A^T \tilde{\pi} - (-\beta \xi + c) \tilde{\theta} - \tilde{\delta} \xi \quad (4.2d)$$

$$-\bar{t}^{-1} - \bar{t}^{-2} \tilde{r} = -\tilde{\theta} \quad (4.2e)$$

$$\tilde{\delta} \geq 0, \quad (\xi^T \tilde{d}) \tilde{\delta} = 0. \quad (4.2f)$$

It will be convenient to set

$$\tilde{s} = \bar{X}^{-1}(e + \bar{X}^{-1}\tilde{d}) \quad (4.2g)$$

and to rewrite (4.2e) as

$$\tilde{\theta} = \bar{t}^{-1}(1 + \bar{t}^{-1}\bar{r}). \quad (4.2h)$$

Next we define

$$\bar{\gamma} = \sqrt{\tilde{d}^T \bar{X}^{-2} \tilde{d} + (\bar{r}/\bar{t})^2} \quad (4.3)$$

and note from (4.2g) and (4.2h) that

$$\bar{\gamma} = \sqrt{(e - \bar{X}\tilde{s})^T (e - \bar{X}\tilde{s}) + (1 - \bar{t}\tilde{\theta})^2} \quad (4.4)$$

and from (4.2a-f) that

$$\frac{q}{\xi^T \bar{x}} \xi^T \tilde{d} - e^T \bar{X}^{-1} \tilde{d} - \bar{t}^{-1} \bar{r} = \bar{\gamma}^2. \quad (4.5)$$

Just as in [6], for example, we have:

**Theorem 4.1 (Primal Improvement).** For  $0 \leq \alpha < 1$ ,

- (i)  $(\bar{x} - (\alpha/\bar{\gamma})\tilde{d}, \bar{t} - (\alpha/\bar{\gamma})\bar{r}, \bar{B})$  is feasible for PR,
- (ii)  $F(\bar{x} - (\alpha/\bar{\gamma})\tilde{d}, \bar{t} - (\alpha/\bar{\gamma})\bar{r}) \leq F(\bar{x}, \bar{t}) - \alpha\bar{\gamma} + \frac{\alpha^2}{2(1-\alpha)}$ , and
- (iii) if  $\bar{\gamma} \geq \frac{4}{5}$  and  $\alpha = \frac{2}{5}$ ,

$$F(\bar{x} - (\alpha/\bar{\gamma})\tilde{d}, \bar{t} - (\alpha/\bar{\gamma})\bar{r}, \bar{s}, \bar{u}) \leq F(\bar{x}, \bar{t}) - \frac{1}{6}.$$

**Proof:** (i) Since the only variables that change are  $x, t$ , (3.9) is still satisfied. And since  $(\tilde{d}, \bar{r})$  lies in the null space of (3.8a-b), it only remains to show that  $\bar{x} - (\alpha/\bar{\gamma})\tilde{d} > 0$ ,  $\bar{t} - (\alpha/\bar{\gamma})\bar{r} > 0$ . This will follow from (4.3), which implies that

$$|(\bar{X}^{-1}\tilde{d})_j| < \bar{\gamma}, \quad j = 1, \dots, n, \quad \text{and} \quad |\bar{r}/\bar{t}| < \bar{\gamma}.$$

Therefore

$$\bar{x} - (\alpha/\bar{\gamma})\tilde{d} = \bar{X}(e - (\alpha/\bar{\gamma})\bar{X}^{-1}\tilde{d}) > 0 \quad \text{for } \alpha \in [0, 1)$$

and

$$\bar{t} - (\alpha/\bar{\gamma})\bar{r} = \bar{t}(1 - (\alpha/\bar{\gamma})(\bar{r}/\bar{t})) > 0 \quad \text{for } \alpha \in [0, 1).$$

Furthermore, from Proposition A.2 of the Appendix,

$$\begin{aligned}
& \sum_{j=1}^n \ell_n(1 - (\alpha/\bar{\gamma})(\bar{X}^{-1}\tilde{d})_j) + \ell_n(1 - (\alpha/\bar{\gamma})(\tilde{r}/\bar{t})) \\
& \geq -(\alpha/\bar{\gamma})e^T \bar{X}^{-1}\tilde{d} - \sum_{j=1}^n \frac{(\alpha/\bar{\gamma})^2(\bar{X}^{-1}\tilde{d})_j^2}{2(1-\alpha)} - (\alpha/\bar{\gamma})(\tilde{r}/\bar{t}) - \frac{(\alpha/\bar{\gamma})^2(\tilde{r}/\bar{t})^2}{2(1-\alpha)} \\
& = -(\alpha/\bar{\gamma})[e^T \bar{X}^{-1}\tilde{d} + \tilde{r}/\bar{t}] - \frac{\alpha^2}{2(\bar{\gamma}^2)(1-\alpha)} (\tilde{d}^T \bar{X}^{-2}\tilde{d}) + (\tilde{r}/\bar{t})^2 \\
& = -(\alpha/\bar{\gamma})[e^T \bar{X}^{-1}\tilde{d} + \tilde{r}/\bar{t}] - \frac{\alpha^2}{2(1-\alpha)} \quad (\text{from 4.3}) \quad (4.6)
\end{aligned}$$

$$(ii) \quad F(\bar{x} - (\alpha/\bar{\gamma})\tilde{d}, \bar{t} - (\alpha/\bar{\gamma})\tilde{r}) - F(\bar{x}, \bar{t})$$

$$\begin{aligned}
& = q \ell_n\left(1 - (\alpha/\bar{\gamma}) \frac{\xi^T \tilde{d}}{\xi^T \bar{x}}\right) - \sum_{j=1}^n \ell_n\left(1 - (\alpha/\bar{\gamma})(\bar{X}^{-1}\tilde{d})_j\right) - \ell_n(1 - (\alpha/\bar{\gamma})(\tilde{r}/\bar{t})) \\
& \leq -\frac{q}{\xi^T \bar{x}} (\alpha/\bar{\gamma})\xi^T \tilde{d} + (\alpha/\bar{\gamma})[e^T \bar{X}^{-1}\tilde{d} + \tilde{r}/\bar{t}] + \frac{\alpha^2}{2(1-\alpha)}
\end{aligned}$$

(which follows from (4.6) and Proposition A.1 of the Appendix)

$$\begin{aligned}
& = (-\alpha/\bar{\gamma}) \left( \frac{q}{\xi^T \bar{x}} \xi^T \tilde{d} - e^T \bar{X}^{-1}\tilde{d} - \tilde{r}\bar{t}^{-1} \right) + \frac{\alpha^2}{2(1-\alpha)} \\
& = (-\alpha/\bar{\gamma}) \bar{\gamma}^2 + \frac{\alpha^2}{2(1-\alpha)} \quad (\text{from 4.5}) \\
& = -\alpha\bar{\gamma} + \frac{\alpha^2}{2(1-\alpha)}.
\end{aligned}$$

This proves (ii). Then (iii) follows by direct substitution.  $\square$

Theorem 4.1 guarantees a decrease in  $F(x,t)$  if the value of  $\gamma = \bar{\gamma}$  is sufficiently large, e.g., if  $\bar{\gamma} \geq 4/5$ . In the case when  $\bar{\gamma}$  is small, we can obtain a reduction in the potential function by replacing the bound  $\bar{B}$  on  $z^*$  by a new bound  $\hat{B}$  generated from new dual variables  $(\hat{\pi}, \hat{\theta}, \hat{s})$  for LD.

Lemma 4.1 (Dual Improvement). Suppose  $\bar{\gamma} < 1$  and  $q \geq n + 1 + \sqrt{n+1}$ . Define

$$(\hat{\pi}, \hat{\theta}, \hat{s}) = \left( \frac{\tilde{\pi}}{\tilde{\theta}}, \beta - \frac{\tilde{\delta}}{\tilde{\theta}} - \frac{q}{\tilde{\theta} \xi^T \bar{x}}, \frac{\tilde{s}}{\tilde{\theta}} \right), \quad (4.7)$$

where  $\tilde{\pi}, \tilde{\theta}, \tilde{s}, \tilde{\delta}$  are given in the solution to (4.2). Then

$(\hat{\pi}, \hat{\theta}, \hat{s})$  is feasible for LD, with dual objective value

$$\hat{B} \triangleq b^T \hat{\pi} > \bar{B} + \frac{\sqrt{n+1} (1-\bar{\gamma})}{(1+\bar{\gamma})} \bar{t}. \quad (4.8)$$

Proof: If  $\bar{\gamma} < 1$ , it follows from (4.4) that  $\tilde{s} > 0$  and  $\tilde{\theta} > 0$ . Therefore  $(\hat{\pi}, \hat{\theta}, \hat{s})$  is well defined

(since we also have  $\xi^T \bar{x} > 0$ ), and  $\hat{s} > 0$ . Then it is easily verified from (4.2d) and (4.2g) that

$$A^T \hat{\pi} + \xi \hat{\theta} + \hat{s} = c, \text{ and so } (\hat{\pi}, \hat{\theta}, \hat{s}) \text{ is feasible for LD, with dual objective value } \hat{B} \triangleq b^T \hat{\pi} =$$

$$\bar{x}^T A^T \hat{\pi}$$

$$= c^T \bar{x} - \bar{x}^T \hat{s} - \xi^T \bar{x} \hat{\theta}$$

$$= c^T \bar{x} - \frac{1}{\tilde{\theta}} e^T (e + \bar{X}^{-1} \tilde{d}) - \xi^T \bar{x} \left( \beta - \frac{\tilde{\delta}}{\tilde{\theta}} - \frac{q}{\tilde{\theta} \xi^T \bar{x}} \right) \quad (\text{from (4.2g) and (4.7)})$$

$$= c^T \bar{x} - \beta \xi^T \bar{x} + \frac{1}{\tilde{\theta}} (-n - e^T \bar{X}^{-1} \tilde{d} + \tilde{\delta} \xi^T \bar{x} + q)$$

$$= \bar{B} - \bar{t} + \frac{1}{\tilde{\theta}} (q - n - e^T \bar{X}^{-1} \tilde{d}) + \frac{\tilde{\delta} \xi^T \bar{x}}{\tilde{\theta}} \quad (\text{from (3.8b)})$$

$$= \bar{B} + \frac{1}{\tilde{\theta}} (q - n - \tilde{\theta} \bar{t} - e^T \bar{X}^{-1} \tilde{d}) + \frac{\tilde{\delta} \xi^T \bar{x}}{\tilde{\theta}}$$

$$= \bar{B} + \frac{1}{\tilde{\theta}} (q - n - (1 + \tilde{r}/\bar{t}) - e^T \bar{X}^{-1} \tilde{d}) + \frac{\tilde{\delta} \xi^T \bar{x}}{\tilde{\theta}} \quad (\text{from (4.2h)})$$

$$= \bar{B} + \frac{1}{\tilde{\theta}} (q - (n+1) - [e^T \bar{X}^{-1} \tilde{d} + \tilde{r}/\bar{t}]) + \frac{\tilde{\delta} \xi^T \bar{x}}{\tilde{\theta}}$$

$$\geq \bar{B} + \frac{1}{\tilde{\theta}} (\sqrt{n+1} - \sqrt{n+1} \bar{\gamma}) + \frac{\tilde{\delta} \xi^T \bar{x}}{\tilde{\theta}} \quad (\text{from (4.3)})$$

$$\geq \bar{B} + \frac{(1-\bar{\gamma})\sqrt{n+1}}{\tilde{\theta}} \geq \bar{B} + \frac{\bar{t}(1-\bar{\gamma})\sqrt{n+1}}{(1+\bar{\gamma})},$$

where the last inequality follows from (4.4) which implies that  $\bar{t} \tilde{\theta} \leq (1+\bar{\gamma})$ .  $\square$

We now can prove:

Theorem 4.2 (Dual Improvement). Suppose  $\bar{\gamma} < 1$  and  $q \geq n + 1 + \sqrt{n+1}$ . Define  $(\hat{\pi}, \hat{\theta}, \hat{s})$  as in (4.7),  $\hat{B}$  as in (4.8) and let

$$\hat{t} = \bar{t} + \hat{B} - \bar{B}. \quad (4.9)$$

Then:

(i)  $(\bar{x}, \hat{t}, \hat{B})$  is feasible for PR, and

$$F(\bar{x}, \hat{t}) \leq F(\bar{x}, \bar{t}) - \ell n \left( 1 + \frac{\sqrt{n+1}(1-\bar{\gamma})}{1+\bar{\gamma}} \right)$$

(ii) If  $\bar{\gamma} < 4/5$ , then

$$F(\bar{x}, \hat{t}) \leq F(\bar{x}, \bar{t}) - \frac{1}{6}.$$

Proof: Because  $(\hat{\pi}, \hat{\theta}, \hat{s})$  is feasible for LD and  $\hat{B} = b^T \hat{\pi}$ , then  $\hat{B} \leq z^*$ , and so  $\hat{B}$  satisfies (3.9).

Also  $\bar{x}$  satisfies (3.8a), and  $\bar{x} > 0$ . Finally, we need to show that  $(-\beta\xi + c)^T \bar{x} + \hat{t} = \hat{B}$ , but this follows easily since  $(-\beta\xi + c)^T \bar{x} + \bar{t} = \bar{B}$  and  $\hat{t} = \bar{t} + \hat{B} - \bar{B}$ , and so (3.8b) is satisfied. Also, since  $\hat{B} > \bar{B}$  (4.8),  $\hat{t} > \bar{t} > 0$ , and so (3.8c) is satisfied. Thus  $(\bar{x}, \hat{t}, \hat{B})$  is feasible in PR.

In order to demonstrate the decrease in the potential function, we note

$$\begin{aligned} F(\bar{x}, \hat{t}) - F(\bar{x}, \bar{t}) &= -\ell n(\hat{t}) + \ell n(\bar{t}) = -\ell n(\hat{t}/\bar{t}) \\ &= -\ell n((\bar{t} + \hat{B} - \bar{B})/\bar{t}) \\ &= -\ell n(1 + (\hat{B} - \bar{B})/\bar{t}) \\ &\leq -\ell n \left( 1 + \frac{\sqrt{n+1}(1-\bar{\gamma})}{(1+\bar{\gamma})} \right) \quad (\text{from (4.8)}) \\ &\leq -\frac{1}{6}, \end{aligned}$$

because  $\bar{\gamma} \leq \frac{4}{5}$  and  $n \geq 3$ .  $\square$

The following algorithm is a summary of the analysis of this section:

Algorithm 1  $(A, b, c, \xi, \beta, x^0, t^0, B^0, q, \gamma)$  ( $q \geq n + 1 + \sqrt{n+1}$ ,  $\gamma \leq 1$ ).

Step 0 (Initialization)  $k = 0$

Step 1 (Compute Primal Direction)  $(\bar{x}, \bar{t}, \bar{B}) = (x^k, t^k, B^k)$

Compute  $(\tilde{d}, \tilde{r}, \tilde{\pi}, \tilde{\theta}, \tilde{\delta}, \tilde{s})$  from (4.2a-h)

Compute  $\bar{\gamma}$  from (4.3)

Step 2 (Determine whether to take Primal Step or to Update Dual Bound)

If  $\bar{\gamma} \geq \gamma$ , go to Step 3.

If  $\bar{\gamma} < \gamma$ , go to Step 4.

Step 3 (Take a Primal Step)

Set  $(\hat{x}, \hat{t}) = (\bar{x} - (\alpha/\bar{\gamma})\tilde{d}, \bar{t} - (\alpha/\bar{\gamma})\tilde{r})$ , where  $\alpha = 2/5$ .

Set  $\hat{B} = \bar{B}$ .

Go to Step 5.

Step 4 (Update Dual Bound)

Compute  $(\hat{\pi}, \hat{\theta}, \hat{s})$  from (4.7).

Compute  $\hat{B}$  from (4.8).

Compute  $\hat{t}$  from (4.9).

Set  $\hat{x} = \bar{x}$ .

Go to Step 5.

Step 5 (Redefine all Variables and Return)

$(x^{k+1}, t^{k+1}, B^{k+1}) = (\hat{x}, \hat{t}, \hat{B})$

$k \leftarrow k+1$ .

Go to Step 1.

With  $q = n + 1 + \sqrt{n+1}$  and  $\gamma = 4/5$ , Theorems 4.1 and 4.2 guarantee a decrease in the  $F(x, t)$  of at least  $\delta = 1/6$  at each iteration of Algorithm 1, yielding the bounds on convergence as

stated in Theorem 3.1.

The next Section discusses ways to accelerate and improve Algorithm 1, and discusses other features as well.

### 5. Modifications and Enhancements to Algorithm 1

Use of a line-search of the potential function. Instead of using a fixed step-length of  $\alpha$  in Step 3,  $\alpha$  can be determined by a line-search of the potential function  $F(x,t)$ . Todd and Burrell [14] have shown that  $F(x,t)$  is quasiconvex, and so the line-search procedure is very simple to execute.

A similar idea can be used to improve the bound  $\hat{B}$  in Step 4 of the algorithm. Suppose  $(\bar{\pi}, \bar{\theta}, \bar{s})$  was a previous solution to the dual LD resulting in the previous bound  $\bar{B} = b^T \bar{\pi}$ . Then at Step 4, the new dual solution is  $(\hat{\pi}, \hat{\theta}, \hat{s})$  with  $\hat{B} = b^T \hat{\pi} > \bar{B}$ , from (4.8). A min-ratio test can be used to compute the largest value  $\alpha^*$  of  $\alpha$  for which the affine combination  $\alpha(\hat{\pi}, \hat{\theta}, \hat{s}) + (1-\alpha)(\bar{\pi}, \bar{\theta}, \bar{s})$  is feasible for LD, and since  $\alpha^* > 1$ ,  $B^* \triangleq b^T(\alpha^* \hat{\pi} + (1-\alpha^*) \bar{\pi}) > \hat{B}$ . This new value  $B^*$  is a valid lower bound on  $z^*$ , and can be used instead of  $\hat{B}$  at Step 4. A further enhancement on the choice of  $B$  is discussed next.

Update the lower bound  $B$  using Fraley's Restriction of the Dual. In Fraley [5], a two-dimensional restriction of the dual problem is developed. This restriction has been used to great advantage in Todd [17], e.g. Here we motivate this problem and show its use in updating the lower bound  $B$  in Algorithm 1. Substituting  $b^T = \bar{x}^T A^T = e^T \bar{X} A^T$  in LD (3.2) and multiplying the constraints by  $\bar{X}$  yields the equivalent form of LD:

$$\text{LD}': \quad \max_{\pi, \theta, s} \quad e^T \bar{X} A^T \pi \quad (5.1a)$$

$$\text{s.t.} \quad \bar{X} A^T \pi + \theta \bar{X} \xi + \bar{X} s = \bar{X} c \quad (5.1b)$$

$$s \geq 0. \quad (5.1c)$$

Note from (5.1b) that 5.1a is equal to  $c^T \bar{x} - \xi^T \bar{x} \theta - \bar{x}^T s$ . Also, note that (5.1b) is equivalent to

$$\theta(\bar{X}\xi)_p + (\bar{X}s)_p = (\bar{X}c)_p,$$

where

$$P = [I - \bar{X}A^T(A\bar{X}^2A^T)^{-1}A\bar{X}] \quad (5.2)$$

and the notation  $v_p$  denotes the quantity  $Pv$ , i.e.  $v_p = Pv$ .

Thus LD is equivalent to

$$LD'': \quad \max_{\theta, s} c^T \bar{x} - \xi^T \bar{x} \theta - \bar{x}^T s \quad (5.3a)$$

$$\text{s.t. } \theta(\bar{X}\xi)_p + (\bar{X}s)_p = (\bar{X}c)_p \quad (5.3b)$$

$$s \geq 0. \quad (5.3c)$$

Now consider the equation system:

$$\theta(\bar{X}\xi)_p + \bar{X}s + \mu(e - e_p) = (\bar{X}c)_p. \quad (5.4)$$

If  $(\mu, \theta, s)$  solves (5.4), then  $(\theta, s)$  solves (5.3b), and so the following program is a restriction of  $LD''$  in the sense that the set of feasible solutions (in  $\theta$  and  $s$ ) is a subset of those of  $LD''$ :

$$FD_{\bar{X}}: \quad z_{\bar{X}} = \max_{\mu, \theta, s} c^T \bar{x} - \xi^T \bar{x} \theta - \bar{x}^T s \quad (5.5a)$$

$$\text{s.t. } \theta(\bar{X}\xi)_p + \bar{X}s + \mu(e - e_p) = (\bar{X}c)_p \quad (5.5b)$$

$$s \geq 0. \quad (5.5c)$$

We denote this linear program as  $FD_{\bar{X}}$  for “Fraley’s restricted dual” and note its dependence on  $\bar{x}$  through (5.5) as well as (5.2). Note also that  $FD_{\bar{X}}$  can be solved as a two-dimensional linear program in  $n$  inequalities in the variables  $\theta$  and  $\mu$ .

Now consider a modification of Algorithm 1 that solves  $FD_{\bar{x}}$  at the start of Step 1, and replaces  $(\bar{B}, \bar{t})$  by  $(z_{\bar{x}}, \bar{t} + z_{\bar{x}} - \bar{B})$  whenever  $FD_{\bar{x}}$  has an optimal solution and  $z_{\bar{x}} \geq \bar{B}$ . (If at some iteration  $FD_{\bar{x}}$  is unbounded, i.e.  $z_{\bar{x}} = +\infty$ , then  $z^* = +\infty$  and so LP has no feasible solution.)

We can show that if Fraley's restricted dual is used to update  $\bar{B}$  at the start of Step 1, then it will always be the case that  $\bar{\gamma} \geq 1$ , and so Step 4 of the algorithm will never be encountered. To see this, suppose that Fraley's restricted dual is used to update  $\bar{B}$  at the start of Step 1 as indicated above. Then  $\bar{B} \geq z_{\bar{x}}$ . Consider the following proposition:

**Proposition 5.1.** Let  $(\hat{\pi}, \hat{\theta}, \hat{s})$  be the dual solution to LD at Step 4 of Algorithm 1. Then  $(\hat{\mu}, \hat{\theta}, \hat{s})$  is feasible for  $FD_{\bar{x}}$ , where  $\hat{\mu} = -1/\hat{\theta}$  and  $\hat{\theta}$  is given in (4.2), with objective value  $\hat{B}$  (see (4.8)), and  $\hat{B} > \bar{B}$ .

**Proof:** Consider the system (4.2).  $A\tilde{d} = 0$ , so  $A\bar{X}(\bar{X}^{-1}\tilde{d}) = 0$ , so  $(\bar{X}^{-1}\tilde{d})_p = \bar{X}^{-1}\tilde{d}$ . From (4.2d),  $\bar{X}^{-1}\tilde{d} = -e - \bar{X}A^T\tilde{\pi} - \left(\tilde{\theta}\beta - \tilde{\delta} - \frac{q}{\xi^T\bar{X}}\right)\bar{X}\xi + \tilde{\theta}\bar{X}c$ , and  $\bar{X}^{-1}\tilde{d} = (\bar{X}^{-1}\tilde{d})_p = -e_p - \left(\tilde{\theta}\beta - \tilde{\delta} - \frac{q}{\xi^T\bar{X}}\right)(\bar{X}\xi)_p + \tilde{\theta}(\bar{X}c)_p$ . Thus from (4.2g) and (4.7),

$$\begin{aligned}\bar{X}\hat{s} &= \frac{1}{\hat{\theta}}\bar{X}\tilde{s} = \frac{1}{\hat{\theta}}(e + \bar{X}^{-1}\tilde{d}) = \frac{1}{\hat{\theta}}(e - e_p) - \left(\beta - \frac{\tilde{\delta}}{\hat{\theta}} - \frac{q}{\xi^T\bar{X}\hat{\theta}}\right)(\bar{X}\xi)_p + (\bar{X}c)_p \\ &= \hat{\mu}(e - e_p) - \hat{\theta}(\bar{X}\xi)_p + (\bar{X}c)_p,\end{aligned}$$

and so  $(\hat{\mu}, \hat{\theta}, \hat{s})$  is feasible for  $FD_{\bar{x}}$ , with objective value  $c^T\bar{x} - \hat{\theta}(\xi^T\bar{x}) - \bar{x}^T\hat{s} = b^T\hat{\pi} = \hat{B} > \bar{B}$ .  $\square$

Now from the above Proposition, if  $\bar{\gamma} < 1$ , then we would produce a solution to  $FD_{\bar{x}}$  with objective  $\hat{B} > \bar{B} \geq z_{\bar{x}}$ , a contradiction. Thus, if  $\bar{B}$  is updated using Fraley's restricted dual, Step 4 of Algorithm 1 will never be encountered.

**Check for Finite Termination of Phase I.** Suppose the algorithm is at Step 3. Instead of setting

$\alpha = 2/5$  or determining  $\alpha$  by a linesearch of the potential function, one could first test if  $(\bar{x} - \alpha\tilde{d})$

solves the Phase-I problem for some value of  $\alpha$ . Since  $A\bar{x} = b$ ,  $\xi^T \bar{x} > 0$ ,  $\xi^T \tilde{d} \geq 0$ ,  $\bar{x} > 0$ , this amounts to checking if  $\bar{x} - \left(\frac{\xi^T \bar{x}}{\xi^T \tilde{d}}\right) \tilde{d} \geq 0$  in the case when  $\xi^T \tilde{d} > 0$ . If indeed  $x' = \bar{x} - \frac{\xi^T \bar{x}}{\xi^T \tilde{d}} \tilde{d} \geq 0$ , then  $x'$  solves the Phase-I problem, and then LP can be solved by a purely Phase II-method, of which many abound.

A “ $\zeta$ -optimal” Algorithm for LP. Suppose, instead solving LP, we are interested in finding a feasible solution  $\bar{x}$  to LP whose objective value is within a value  $\zeta > 0$  of  $z^*$ , i.e.  $c^T \bar{x} \leq z^* + \zeta$ . (One can easily imagine a variety of situations where this is a reasonable goal, such as when the objective function is not well-specified, etc.). Thus we seek a point  $\bar{x}$  that satisfies

$$A\bar{x} = b$$

$$\xi^T \bar{x} = 0$$

$$\bar{x} \geq 0$$

$$c^T \bar{x} \leq z^* + \zeta.$$

Algorithm 1 can be easily modified to accomplish this goal, as follows: Whenever the bound  $\bar{B}$  is updated to  $\hat{B}$  in Step 4, replace  $\bar{B}$  by  $\hat{B} + \zeta$ , instead of  $\hat{B}$ . Then all dual bounds  $B^k$  will satisfy  $B^k \leq z^* + \zeta$ , and hence all points  $x^k$  will satisfy

$$(-\beta\xi + c)^T x^k \leq B^k \leq z^* + \zeta.$$

Rearranging gives

$$c^T x^k \leq (z^* + \zeta) + \beta\xi^T x^k.$$

Then Theorem 3.1(i) is still valid and so as  $\xi^T x^k \rightarrow 0$ ,

$$\limsup_{k \rightarrow \infty} c^T x^k \leq z^* + \zeta.$$

Whenever  $\bar{B}$  is updated to  $\hat{B}$  in Step 4,  $\bar{B}$  increases by at least  $\zeta > 0$ . Thus Step 4 can only be visited at most  $\left\lceil \frac{z^* - B^0}{\zeta} \right\rceil$  times. Furthermore, the fact that the bound  $\bar{B}$  is increased by at least  $\zeta$  should accelerate convergence of the algorithm.

An Explicit Convergence Constant for Algorithm 1. Theorem 3.1 together with Lemma 3.1 state that all iterates  $(x^k, t^k, B^k)$  of Algorithm 1 must satisfy

$$(\xi^\top x^k) \leq (\xi^\top x^0) C_1 e^{-k/6q},$$

where  $C_1$  is given in (3.11). Although it is easy to see that  $C_1 \leq 2^L$  when  $z^*$  is finite and  $\beta, x^0$ , and  $t^0$  have size  $O(L)$ , it is impractical to compute  $C_1$ . Knowing  $C_1$  is nevertheless important from the point of view of a prior guarantee that  $\xi^\top x^k$  will be no larger than a certain value after a certain number of iterations. Below we show that if an upper bound  $\bar{U}$  on  $z^*$  is known in advance, then  $C_1$  can be replaced by a known value  $\tilde{C}_1$  (derived below) whenever the algorithm visits Step 4.

Lemma 5.1. (Computing a Substitute value of  $C_1$ ). Suppose Algorithm 1 is in Step 4 at iteration  $k$ . Then let

$$\tilde{D} = \frac{1}{1-\bar{\gamma}} \left[ n + 1 - \bar{\gamma}^2 + \frac{(1+\bar{\gamma})}{t} [\bar{U} - \bar{B}] + \frac{q}{\xi^\top \bar{x}} \xi^\top \bar{d} \right]$$

and

$$\tilde{C}_1 = [(n+1)^{-1} \tilde{D}]^{\left(\frac{n+1}{q}\right)}.$$

Then:

(i) for all subsequent iterations  $i > k$ ,

$$\xi^\top x^i \leq (\xi^\top x^k) \tilde{C}_1 e^{-\left(\frac{i-k}{6q}\right)},$$

and

(ii) if  $\gamma = \frac{4}{5}$ , then  $\bar{\gamma} < \frac{4}{5}$  and  $\tilde{D} \leq 5[q + \frac{2}{t} (\bar{U} - \bar{B})]$ .

Proof: Suppose at iteration  $k$  that the algorithm is in Step 4. Then  $\bar{\gamma} < \gamma \leq 1$ . Consider the linear program:

$$\begin{aligned}
 \text{PB:} \quad & \bar{T} = \max_{\mathbf{x}, t} \mathbf{e}^T \bar{\mathbf{X}}^{-1} \mathbf{x} + \bar{t}^{-1} t \\
 & \text{s.t. } \mathbf{A} \mathbf{x} = \mathbf{b} \\
 & (-\beta \xi + c)^T \mathbf{x} + t \leq \bar{U} \\
 & \xi^T \mathbf{x} \leq \xi^T \bar{\mathbf{x}} \\
 & -\xi^T \mathbf{x} \leq 0.
 \end{aligned}$$

Note that since  $\bar{U} \geq z^* \geq B^i$ ,  $i = k, \dots$ , then all iterates  $(\mathbf{x}^i, t^i)$  will be feasible for PB for  $i \geq k$ .

The dual of PB is:

$$\text{DB:} \quad \bar{T} = \min_{\lambda, \theta, \delta, \mu} \mathbf{b}^T \lambda + \bar{U} \theta + \xi^T \bar{\mathbf{x}} \delta \quad (5.6a)$$

$$\text{s.t. } \mathbf{A}^T \lambda + \theta(c - \beta \xi) + (\delta - \mu) \xi \geq \bar{\mathbf{X}}^{-1} \mathbf{e} \quad (5.6b)$$

$$\theta \geq \bar{t}^{-1} \quad (5.6c)$$

$$\theta, \delta, \mu \geq 0. \quad (5.6d)$$

Suppose we are at Step 4 of the algorithm, and so  $\bar{\gamma} < \gamma \leq 1$ . Then  $\tilde{s} > 0$ , and upon setting

$(\lambda', \theta', \delta', \mu') = \left( \frac{-\tilde{\pi}}{1-\bar{\gamma}}, \frac{\tilde{\theta}}{1-\bar{\gamma}}, \frac{\mathbf{q}}{\xi^T \bar{\mathbf{x}}(1-\bar{\gamma})} + \frac{\tilde{\delta}}{(1-\bar{\gamma})}, 0 \right)$ , it can be verified by rearranging (4.2d) that

$(\lambda', \theta', \delta', \mu')$  is feasible for DB. To see this, note from (4.3) that  $\bar{\mathbf{X}}^{-1} \tilde{\mathbf{d}} \geq -\bar{\gamma} \mathbf{e}$  and  $\bar{t}^{-1} \tilde{r} \geq -\bar{\gamma}$ , and

from (4.2d) that  $\mathbf{A}^T(-\tilde{\pi}) + \tilde{\theta}(-\beta \xi + c) + (\tilde{\delta} + \frac{\mathbf{q}}{\xi^T \bar{\mathbf{x}}}) \xi = \bar{\mathbf{X}}^{-1}(\mathbf{e} + \bar{\mathbf{X}}^{-1} \tilde{\mathbf{d}}) \geq \bar{\mathbf{X}}^{-1} \mathbf{e}(1-\bar{\gamma})$ . Dividing

through by  $(1-\bar{\gamma})$  shows that  $(\lambda', \theta', \delta', \mu')$  solves (5.6b). Also, from (4.2h),

$\theta' = \frac{\tilde{\theta}}{1-\bar{\gamma}} = \frac{1 + \bar{t}^{-1} \tilde{r}}{\bar{t}} \left( \frac{1}{1-\bar{\gamma}} \right) \geq \frac{1-\bar{\gamma}}{1-\bar{\gamma}} \left( \frac{1}{\bar{t}} \right) = \bar{t}^{-1}$ , so (5.6c) is satisfied. Finally, note from (4.2h) and

(4.3) that  $\theta', \delta', \mu' \geq 0$ , so (5.6d) is satisfied. Then  $\bar{T} \leq \mathbf{b}^T \lambda' + \bar{U} \theta' + \delta'^T \xi^T \bar{\mathbf{x}}$

$$\begin{aligned}
&= \frac{1}{1-\bar{\gamma}} [-\bar{\mathbf{x}}^T \mathbf{A}^T \tilde{\boldsymbol{\pi}} + \bar{\mathbf{U}} \tilde{\boldsymbol{\theta}} + \mathbf{q} + \tilde{\delta} \boldsymbol{\xi}^T \bar{\mathbf{x}}] \\
&= \frac{1}{1-\bar{\gamma}} (-(\bar{\mathbf{B}} - \bar{\mathbf{t}}) \tilde{\boldsymbol{\theta}} - \tilde{\delta} \boldsymbol{\xi}^T \bar{\mathbf{x}} - \mathbf{q} + \mathbf{n} + \mathbf{e}^T \bar{\mathbf{X}}^{-1} \tilde{\mathbf{d}} + \bar{\mathbf{U}} \tilde{\boldsymbol{\theta}} + \mathbf{q} + \tilde{\delta} \boldsymbol{\xi}^T \bar{\mathbf{x}}) \quad (\text{from (4.2d)}) \\
&= \frac{1}{1-\bar{\gamma}} (\mathbf{n} + \bar{\mathbf{t}} \tilde{\boldsymbol{\theta}} + (\bar{\mathbf{U}} - \bar{\mathbf{B}}) \tilde{\boldsymbol{\theta}} + \mathbf{e}^T \bar{\mathbf{X}}^{-1} \tilde{\mathbf{d}}) \\
&= \frac{1}{1-\bar{\gamma}} (\mathbf{n} + 1 + \bar{\mathbf{t}}^{-1} \bar{\mathbf{r}} + (\bar{\mathbf{U}} - \bar{\mathbf{B}}) \tilde{\boldsymbol{\theta}} + \mathbf{e}^T \bar{\mathbf{X}}^{-1} \tilde{\mathbf{d}}) \quad (\text{from (4.2h)}) \\
&= \frac{1}{1-\bar{\gamma}} (\mathbf{n} + 1 + (\bar{\mathbf{U}} - \bar{\mathbf{B}}) \tilde{\boldsymbol{\theta}} + \frac{\mathbf{q}}{\boldsymbol{\xi}^T \bar{\mathbf{x}}} \boldsymbol{\xi}^T \tilde{\mathbf{d}} - \bar{\gamma}^2) \quad (\text{from (4.5)}) \\
&\leq \tilde{\mathbf{D}},
\end{aligned}$$

since  $\tilde{\boldsymbol{\theta}} \leq (1+\bar{\gamma})/\bar{\mathbf{t}}$  from (4.2h). Finally, for any  $i > k$ ,  $(\mathbf{x}^i, \mathbf{t}^i)$  is feasible in  $\bar{\mathbf{P}}\bar{\mathbf{B}}$ , so that

$$\mathbf{e}^T \bar{\mathbf{X}}^{-1} \mathbf{x}^i + \bar{\mathbf{t}}^{-1} \mathbf{t}^i \leq \bar{\mathbf{T}} \leq \tilde{\mathbf{D}}.$$

It then follows from the arithmetic-geometric mean inequality that

$$\sum_{j=1}^n \ell_{\mathbf{n}\mathbf{x}_j^i} + \ell_{\mathbf{n}\mathbf{t}^i} - \sum \ell_{\mathbf{n}\bar{\mathbf{x}}_j} - \ell_{\mathbf{n}\bar{\mathbf{t}}} \leq (n+1) \ell_{\mathbf{n}(\tilde{\mathbf{D}}/(n+1))}.$$

The proof of (i) then follows as in Lemma 3.1 and Theorem 3.1(i). To see (ii), note from (4.5) that

$$\begin{aligned}
\frac{\mathbf{q}}{\boldsymbol{\xi}^T \bar{\mathbf{x}}} \boldsymbol{\xi}^T \tilde{\mathbf{d}} &= \bar{\gamma}^2 + \mathbf{e}^T \bar{\mathbf{X}}^{-1} \tilde{\mathbf{d}} + \bar{\mathbf{t}}^{-1} \bar{\mathbf{r}} \\
&\leq \bar{\gamma}^2 + \sqrt{n+1} \bar{\gamma} \leq \bar{\gamma}^2 + \sqrt{n+1}.
\end{aligned}$$

Then with  $\bar{\gamma} < \gamma = \frac{4}{5}$ ,

$$\tilde{\mathbf{D}} \leq 5[\mathbf{q} + \frac{2}{\bar{\mathbf{t}}} (\bar{\mathbf{U}} - \bar{\mathbf{B}})]. \quad \square$$

### Acknowledgement

I would like to acknowledge Michael Todd for many stimulating discussions regarding the warm-start LP problem, which have had an influence on the research herein.

### Appendix

Proposition A.1. If  $x > -1$ ,  $\ln(1+x) \leq x$ .

Proposition A.2. If  $|x| \leq \alpha < 1$ , then  $\ln(1+x) \geq x - \frac{x^2}{2(1-\alpha)}$ .

Proofs of the above two inequalities can be found in [6], among other places.

Proposition A.3. Consider the dual linear programs:

$$\begin{array}{ll}
 \text{LP}_r: & z^*(r) = \min_x c^T x \\
 & \text{s.t. } Ax = b \\
 & \xi^T x = r \\
 & x \geq 0.
 \end{array}
 \qquad
 \begin{array}{ll}
 \text{LD}_r: & \max_{\pi, \theta} b^T \pi + r\theta \\
 & \text{s.t. } A^T \pi + \xi \theta \leq c
 \end{array}$$

Suppose  $(\pi^*, \theta^*)$  solves  $\text{LD}_0$ , and let  $z^* = z^*(0)$ . Then for any  $x$  feasible for  $\text{LP}_r$ ,  $c^T x \geq z^* + \theta^* r$ .

Proof: Because  $(\pi^*, \theta^*)$  is feasible for the dual  $\text{LD}_r$  for any  $r$ ,

$$z^*(r) \geq b^T \pi^* + r\theta^* = z^*(0) + r\theta^* = z^* + r\theta^*.$$

Therefore, if  $x$  is feasible for  $\text{LP}_r$ ,  $c^T x \geq z^*(r) \geq z^* + r\theta^*$ .  $\square$

### References

- [1] K.M. Anstreicher, "A combined phase I - phase II projective algorithm for linear programming," Mathematical Programming 43 (1989) 209-223.
- [2] K.M. Anstreicher, "A combined phase I - phase II scaled potential algorithm for linear programming," CORE Discussion Paper 8939, CORE Catholic University of Louvain, Belgium, 1989.
- [3] K.M. Anstreicher, "On monotonicity in the scaled potential algorithm for linear programming," Linear Algebra and its Applications 152 (1991) 223-232.
- [4] K.M. Anstreicher, "On interior algorithms for linear programming with no regularity assumptions," Manuscript, Dept. of Operations Research, Yale University, New Haven (February 1991).
- [5] C. Fraley, "Linear updates for a single-phase projective method," Operations Research Letters 9 (1990) 169-174.
- [6] R.M. Freund, "Polynomial-time algorithms for linear programming based only on primal scaling and projected gradients of a potential function," Mathematical Programming 51 (1991) 203-222.
- [7] R.M. Freund, "Theoretical efficiency of a shifted barrier function algorithm for linear programming," Linear Algebra and its Applications 152 (1991) 19-41.
- [8] R.M. Freund, "A potential-function reduction algorithm for solving a linear program directly from an infeasible 'warm start'," Working paper 3079-89-MS, Sloan School of Management, MIT, September 1989, to appear in Mathematical Programming.
- [9] G. de Ghellinck and J.-P. Vial, "A polynomial Newton method for linear programming," Algorithmica 1 (1986) 425-453.
- [10] P.E. Gill, W. Murray, M.A. Saunders and M.H. Wright, "Shifted barrier methods for linear programming," Technical Report SOL 88-9, Department of Operations Research, Stanford University, 1988.
- [11] C. C. Gonzaga, "Polynomial affine algorithms for linear programming," Mathematical Programming 49 (1990) 7-21.
- [12] N. Karmarkar, "A new polynomial time algorithm for linear programming," Combinatorica 44 (1984) 373-395.
- [13] R.A. Polyak, "Modified barrier functions," Report RC 14602, IBM T.J. Watson Research Center, Yorktown Heights, NY, 1989.
- [14] M.J. Todd and B. Burrell, "An extension of Karmarkar's algorithm for linear programming using dual variables," Algorithmica 1 (1986) 409-424.
- [15] M.J. Todd, "On Anstreicher's combined phase I - phase II projective algorithm for linear programming," Technical Report No. 776, School of OR&IE, Cornell University, Ithaca, NY, 1988, to appear in Mathematical Programming.

- [16] M.J. Todd and Y. Wang, "On combined phase 1 - phase 2 projective methods for linear programming," Technical Report No. 877, School of OR&IE, Cornell University, Ithaca, NY, 1989, to appear in Algorithmica.
- [17] M.J. Todd, "Combining phase I and phase II in a potential reduction algorithm for linear programming," Technical Report No. 907, School of OR&IE, Cornell University, Ithaca, NY, 1990, to appear in SIAM Journal on Optimization.
- [18] J.P. Vial, "A projective algorithm for linear programming with no regularity condition," University of Geneva, Geneva, Switzerland, 1991.
- [19] M.H. Wagner, "Supply-demand decomposition of the national coal model," Operations Research 29 (1981) 1137-1153.
- [20] Y. Ye, "An  $O(n^3L)$  potential reduction algorithm for linear programming," Mathematical Programming 50 (1991) 239-258.