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

TECHNICAL REPORT NO. 1136

January 1996

**An Infeasible-Interior-Point
Potential-Reduction Algorithm
for Linear Programming¹**

by

Reha H. Tütüncü

¹Research was supported in part by National Science Foundation grant DMS9505155 and ONR grant N00014-96-1-0050.

An Infeasible-Interior-Point Potential-Reduction Algorithm for Linear Programming

Reha H. Tütüncü *

January, 1996

Abstract

This paper studies a new potential-function and an infeasible-interior-point method based on this function for the solution of linear programming problems. This work is motivated by the apparent gap between the algorithms with the best worst-case complexity and their most successful implementations. For example, analyses of the algorithms are usually carried out by imposing several regularity assumptions on the problem, but implementations can solve problems which do not satisfy these assumptions. Furthermore, most state-of-the-art implementations incorporate heuristic tricks, but these modifications are rarely addressed in the theoretical analysis of the algorithms. The method described here and its analysis have the flexibility to integrate any heuristic technique for implementation while maintaining the important polynomial complexity feature.

Key words: linear programming, potential functions, infeasible-interior-point methods, homogeneity, self-dual.

AMS Subject classification: 90C05.

*School of Operations Research and Industrial Engineering Cornell University, Ithaca, New York 14853, USA. E-mail: tutuncu@orie.cornell.edu. Research supported in part by NSF through grant DMS 9505155 and ONR through grant N00014-96-1-0050.

1 Introduction

This paper considers interior-point methods for the solution of linear programming problems given in the following standard form:

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

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ are given, and $x \in \mathbb{R}^n$. The matrix A is assumed to have full row rank without loss of generality. The dual problem associated with (LP) is:

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

where $y \in \mathbb{R}^m$ and $s \in \mathbb{R}^n$.

Although Karmarkar's original work on interior-point methods [4] used a potential-reduction algorithm, in later years, especially following Renegar's work [11], the focus of research on interior-point methods shifted to path-following algorithms. Meanwhile, potential-reduction algorithms received consistent but sporadic attention. See Todd's recent paper for a review of the potential-reduction algorithms for mathematical programming problems [13].

Path-following algorithms restrict the iterates to stay in some neighborhood of the so-called central path, and in this way excellent theoretical convergence results can be obtained. Gonzaga's 1992 paper gives a survey of the developments on path-following methods [2]. But, as observed by Todd, "practical methods inspired by path-following approaches eliminate most of the path-following restrictions, and this leads to an unfortunate gap between theory and practice" [13]. Our purpose here is to develop an algorithm that has theoretical convergence guarantees as well as practical flexibility; therefore we focus on the potential-reduction methods that seem to serve this purpose better.

Tanabe [12] and Todd and Ye [15] independently introduced and analyzed a primal-dual potential function $\Phi_\rho(x, s)$ that has been the main tool of recent potential-reduction algorithms. A succession of articles showed that

this potential function can be reduced by at least a constant amount starting from any feasible interior point, for large enough values of ρ [15, 19, 5]. However, these methods are feasible methods, i.e., they have to start at a strictly feasible initial solution and obtaining such a solution is often hard. Even worse, feasible or strictly feasible solutions may not exist for particular instances of the LP problem. One alternative is to consider algorithms that generate strictly positive iterates that do not necessarily satisfy the equality constraints of the problem, and try to achieve feasibility and optimality simultaneously. These methods are known as infeasible-interior-point methods and seem to be the most popular choice for practical implementations. Although they are usually path-following algorithms, there also exist some infeasible-interior-point potential-reduction algorithms in the literature. One notable example is Mizuno, Kojima, and Todd's $O(nL)$ -iteration primal-dual algorithm [9]. They describe a constrained potential-reduction algorithm based on the Tanabe-Todd-Ye (TTY) potential function, and a pure potential-reduction algorithm based on a variant of it. Furthermore, they suggest the use of some other potential functions in the infeasible case, and note that guaranteed decreases in these potential functions may be very hard to obtain. Here, we describe a new potential function and an infeasible-interior-point algorithm based on this function.

Current infeasible-interior-point methods either generate an approximate optimal solution to a nearby problem or establish that any *optimal* solution to the problem must have large norm. This is not satisfactory since the latter conclusion does not necessarily indicate that the problem instance is nearly infeasible or unbounded, but may only imply it is ill-conditioned. Our method achieves a stronger conclusion: we either generate an approximate solution or establish that any *feasible* solution must have large norm. This is accomplished by decreasing a potential function while maintaining a minimal centrality condition. This result is more conclusive since, as shown by Todd and Ye [16], if all feasible solutions for an (LP) instance are large then there is a neighboring problem that is infeasible; alternatively, if all feasible dual solutions are large then there is a neighboring dual infeasible problem.

The potential function we use is inspired by the TTY potential function for the homogeneous self-dual LP formulation of Ye, Todd, and Mizuno [21]. We show how a rich family of potential functions defined on original infeasible variables can be reduced at the same rate as the TTY potential function on the feasible variables of this artificial problem and analyze a particular

member of this family. The significance of this result lies in the fact that it mimics the result in [5] for infeasible-point potential functions, i.e. it provides a recipe to guarantee reductions from any infeasible-point and ensures global polynomial convergence.

This guarantee also gives the method the flexibility to incorporate heuristic tricks to make it more efficient. In particular, once the direction guaranteeing the reduction is generated, line searches along this direction with possibly different step sizes in primal and dual spaces can be employed to reduce the potential function faster. This practical flexibility coupled with the theoretical guarantee narrows the gap between the theory and practice of interior-point-methods.

This paper is organized as follows. Following this introduction, Section 2 reviews the homogeneous self-dual formulation of Ye, Todd, and Mizuno [21]. Section 3 introduces a family of potential functions which can be reduced by at least a constant amount from any infeasible-interior-point. A particular member of this family, a potential function that combines measures of feasibility and optimality, is analyzed in Section 4. Section 5 describes our algorithm and concludes. The notation used here is mostly standard. $\|\cdot\|$ is used exclusively to denote the two-norm of real vectors. We use subscripts to denote components of vectors and superscripts to denote iteration indices. However, we find it convenient to use subscript 0 to indicate the initial solution in the homogeneous self-dual formulation.

2 Homogeneous Self-Dual LP Formulation

A common approach for solving the problems (LP) and (LD) is to combine them into an artificial problem and generate solutions for the original problems using a solution to the artificial problem. One of the earlier examples of this technique is the homogeneous linear feasibility model developed by Goldman and Tucker [1, 17]:

$$\begin{array}{rcll}
 (HLF) & & Ax & - & b\tau & = & 0 \\
 & & -A^T y & & + & c\tau & \geq & 0 \\
 & & b^T y & - & c^T x & & \geq & 0 \\
 & & x & , & \tau & \geq & 0
 \end{array}$$

Here, τ serves as a homogenizing variable for the primal and dual feasibility

constraints. The third constraint is reversed weak duality and ensures the equality of the primal and dual objective functions for feasible solutions. From a solution (y^*, x^*, τ^*) to (HLF) with $\tau^* > 0$ one can obtain optimal solutions for (LP) and (LD) simply by scaling x^* and y^* . Furthermore, if $\tau^* = 0$, but $b^T y^* > 0$ or $c^T x^* < 0$ then it is easily seen that either the primal or the dual problem is infeasible (or both). One interesting feature of the (HLF) model is self-duality. A linear program is called *self-dual* when the dual of the problem is equivalent to the primal. Thus, (HLF) can be viewed as a self-dual LP with the objective function equal to zero.

Recently, Ye, Todd, and Mizuno considered a larger system that allows them to start the search for an optimal solution from an arbitrary interior point [21]. Given any $x_0 > 0$, $s_0 > 0$, $\kappa_0 > 0$, and y_0 , they formulate an artificial homogeneous self-dual linear program :

$$\begin{array}{llllllll}
\min & & & & & & (x_0^T s_0 + \kappa_0)\theta & \\
\text{s.t.} & & Ax & - & b\tau & + & \bar{b}\theta & = 0 \\
(HSDP) & -A^T y & & + & c\tau & - & \bar{c}\theta & \geq 0 \\
& b^T y & - & c^T x & & + & \bar{z}\theta & \geq 0 \\
& -\bar{b}^T y & + & \bar{c}^T x & - & \bar{z}\tau & & = -x_0^T s_0 - \kappa_0 \\
& y \text{ free} & x \geq 0 & \tau \geq 0 & & & \theta \text{ free,} &
\end{array}$$

where

$$\bar{b} = b - Ax_0, \quad \bar{c} = c - A^T y_0 - s_0, \quad \text{and} \quad \bar{z} = c^T x_0 + \kappa_0 - b^T y_0.$$

Let s and κ be the surplus variables for the second and third constraints. For any optimal solution $(y^*, x^*, \tau^*, \theta^*, s^*, \kappa^*)$ of the (HSDP) $\theta^* = 0$ (see [21]), and in this case the variables satisfy the same equations as in (HLF), with the addition of a normalizing constraint. Thus, the comments above apply to the solutions of the (HSDP) as well. Furthermore, $y = y_0$, $x = x_0$, $\tau = 1$, $\theta = 1$, $s = s_0$, and $\kappa = \kappa_0$ is a strictly feasible solution for this problem, and therefore, any (feasible) interior-point algorithm, and in particular algorithms with $O(\sqrt{n} \ln \frac{1}{\epsilon})$ complexity, can be used to solve this problem. Note that Ye et al. choose $\kappa_0 = 1$, but we find it useful to work with arbitrary positive κ_0 values.

More recently, Xu, Hung, and Ye observed that same results can be achieved working with the simpler system in (HLF) and using infeasible-interior-point algorithms [18]. They also reported very favorable computational results from an implementation of their algorithm. However, their

implementation makes several modifications to the algorithm they analyzed theoretically, creating a gap between theory and practice.

These formulations offer significant advantages over the usual primal-dual systems for the analysis of interior-point algorithms. Using the (HSDP) formulation global polynomial convergence of interior-point methods can be established without making any regularity assumptions regarding the data (e.g. existence of feasible solutions, strictly feasible solutions, non-degeneracy, etc.) or using any big M penalty parameters or lower bounds. More remarkably, as opposed to the previous artificial problems that combine the primal and dual problems and have a known feasible solution, the self-dual LP does not necessitate the doubling of the dimension of linear systems to be solved at each iteration. If a primal-dual algorithm uses the same initial point for the primal (HSDP) and its identical dual, then the equations defining the primal and dual search directions are equivalent, and it suffices to solve just one of them. Furthermore, an optimal solution $(y^*, x^*, \tau^*, \theta^*, s^*, \kappa^*)$ with $\tau^* + \kappa^* > 0$ gives conclusive information about the status of the problem, i.e., one can either get primal and dual optimal solutions or a proof of primal or dual infeasibility. This condition is satisfied by *strictly self-complementary* solutions, i.e., solutions with the property that $\begin{pmatrix} x^* \\ \tau^* \end{pmatrix} + \begin{pmatrix} s^* \\ \kappa^* \end{pmatrix} > 0$. When such a solution is available we have the following result:

Theorem 2.1 (Ye-Todd-Mizuno [21]) : *Let $(y^*, x^*, \tau^*, \theta^* = 0, s^*, \kappa^*)$ be a strictly self-complementary solution for the (HSDP).*

- (i) *(LP) has an optimal solution if and only if $\tau^* > 0$. In this case, x^*/τ^* is an optimal solution for (LP) and $(y^*/\tau^*, s^*/\tau^*)$ is an optimal solution for (LD).*
- (ii) *If $\tau^* = 0$, then $\kappa^* > 0$, which implies that $c^T x^* - b^T y^* < 0$, i.e., at least one of $c^T x^*$ and $-b^T y^*$ is strictly less than zero. If $c^T x^* < 0$ then (LD) is infeasible; if $-b^T y^* < 0$ then (LP) is infeasible; and if both $c^T x^* < 0$ and $-b^T y^* < 0$ then both (LP) and (LD) are infeasible. $\square$*

As established by Güler and Ye [3], most interior-point algorithms, including all path-following methods as well as potential-reduction methods using theoretical step-size selection rules, generate strictly complementary solutions. This result implies that these algorithms will produce strictly

self-complementary solutions for the (HSDP), satisfying the hypothesis of Theorem 2.1. However, this is more than one needs. As stated above, having only $\tau^* + \kappa^* > 0$ suffices, and requiring only the satisfaction of this condition from an algorithm has two potential advantages. First, this will lead to less restricted step-size selection rules in the algorithm, enabling better and faster improvements. Second, for degenerate problems, strictly complementary solutions are not vertex solutions. But, most LPs appearing in applications are degenerate, and vertex solutions are often preferred. Hence, relaxation of the strict-complementarity condition can lead to algorithms that converge to vertex solutions. Motivated by these considerations, in Section 5 we will study an algorithm that drops centrality conditions on x and s .

We now state some basic properties of this formulation:

Proposition 2.1 *Let $(y, x, \tau, \theta, s, \kappa)$ and $(y^+, x^+, \tau^+, \theta^+, s^+, \kappa^+)$ be feasible for the (HSDP) and define $\bar{u} := \frac{u}{\tau}$ and $\bar{u}^+ := \frac{u^+}{\tau^+}$ for variables u and u^+ that may be vectors or scalars. Then the following hold:*

$$(i) \quad x^T s + \tau \kappa = \theta(x_0^T s_0 + \kappa_0).$$

$$(ii) \quad \frac{\|A\bar{x} - b\|}{\|A\bar{x}^+ - b\|} = \frac{\|A^T\bar{y} + \bar{s} - c\|}{\|A^T\bar{y}^+ + \bar{s}^+ - c\|} = \frac{c^T\bar{x} + \bar{\kappa} - b^T\bar{y}}{c^T\bar{x}^+ + \bar{\kappa}^+ - b^T\bar{y}^+} = \frac{\bar{\theta}}{\bar{\theta}^+}.$$

PROOF: The proof of (i) is given in [21] and can be obtained by multiplying the first set of equations by y^T , the second set by x^T , the third equation by τ and the last by θ and adding all of them. The proof of (ii) is straightforward.

□

Another identity related to (ii) above is:

$$\bar{x}^T \bar{s} + \bar{\kappa} = -(A\bar{x} - b)^T \bar{y} + (A^T \bar{y} + \bar{s} - c)^T \bar{x} + (c^T \bar{x} + \bar{\kappa} - b^T \bar{y}). \quad (3)$$

Therefore, if the iterates $(\bar{x}, \bar{y}, \bar{s})$ approach feasibility and stay bounded, the expression $(c^T \bar{x} + \bar{\kappa} - b^T \bar{y})$ approaches total uncomplementarity, $\bar{x}^T \bar{s} + \bar{\kappa}$.

3 A family of potential functions

The use of potential functions in conjunction with interior-point methods for mathematical programming has been popular since Karmarkar's original work in 1984 [4]. The function Karmarkar uses is a primal-only potential

function meaning that it only measures the quality of a primal iterate. In 1988, Tanabe [12], and Todd and Ye [15] independently introduced and analyzed a primal-dual potential function that turned out to be a very useful tool for construction and analysis of efficient LP algorithms. For the primal-dual pair of problems (1) and (2), this function is defined as

$$\Phi_\rho(x, s) = \rho \ln(x^T s) - \sum_{i=1}^n \ln x_i s_i, \text{ for every } (x, s) > 0. \quad (4)$$

Using a primal-dual update Kojima, Mizuno, and Yoshise showed that when $\rho \geq n + \sqrt{n}$ the Tanabe-Todd-Ye (TTY) potential function can be reduced by at least 0.2 from any feasible point (x, y, s) with $(x, s) > 0$ [5]. Interestingly, the search direction that they use in their analysis comes from primal-dual path-following algorithms. This direction can be interpreted as a steepest-descent direction for $\Phi_\rho(\cdot)$ with respect to an appropriate norm and is the unique solution to the following system:

$$\begin{aligned} A\Delta x &= 0 \\ A^T \Delta y + \Delta s &= 0 \\ S\Delta x + X\Delta s &= \alpha \left(\frac{x^T s}{\rho'} e - XSe \right), \end{aligned} \quad (5)$$

where $X = \text{diag}(x)$, $S = \text{diag}(s)$, e is a vector of ones of appropriate dimension,

$$\alpha = \frac{\beta \sqrt{\min(XSe)}}{\left\| (XS)^{-\frac{1}{2}} \left(\frac{x^T s}{\rho'} e - XSe \right) \right\|}, \text{ with } \beta \in (0, 1), \quad (6)$$

and $\min(v) = \min\{v_1, v_2, \dots, v_n\}$ for $v \in \mathbb{R}^n$. Kojima, Mizuno, and Yoshise choose $\rho' = \rho$. Ye et al. generalized their result to the case where ρ' is not restricted to equal ρ and can vary from one iteration to the other [20]. If ρ' is chosen carefully, it is possible to maintain the guarantee of constant reduction in every iteration and to improve the practical performance. We reproduce their result to use later in our analysis:

Theorem 3.1 (Ye et al. [20]) : *Let x and (y, s) be feasible for (1) and (2) respectively, and let $(\Delta x, \Delta y, \Delta s)$ satisfy (5) and (6). Then $x^+ = (x + \Delta x)$*

and $(y^+, s^+) = (y + \Delta y, s + \Delta s)$ are feasible for (1) and (2), and

$$\Phi_\rho(x^+, s^+) - \Phi_\rho(x, s) \leq -\sqrt{\frac{\min(XSe)}{x^T s}} \frac{(\rho' - n)(\rho - n) + h_{x,s}}{\sqrt{(\rho' - n)^2 + h_{x,s}}} \beta + \frac{\beta^2}{2(1 - \beta)}, \quad (7)$$

where $h_{x,s} = \sum_i \frac{x^T s}{x_i s_i} - n^2 \geq 0$. $\square$

When $\rho' = \rho \geq n + \sqrt{n}$, Kojima et al. [5] show that

$$\Phi_\rho(x + \Delta x, s + \Delta s) - \Phi_\rho(x, s) \leq -\frac{\sqrt{3}}{2} \beta + \frac{\beta^2}{2(1 - \beta)}, \quad (8)$$

and when $\rho' = n$ and $\frac{\min(XSe)}{x^T s} \leq n^{-2}$, we have that

$$\Phi_\rho(x + \Delta x, s + \Delta s) - \Phi_\rho(x, s) \leq -\sqrt{\frac{n-1}{n+1}} \beta + \frac{\beta^2}{2(1 - \beta)} \quad (9)$$

using Lemma 5 in [3] which shows that $\min(XSe)h_{x,s} \geq \frac{n-1}{n+1} x^T s$. In either case, the right-hand-side is at most -0.2 for $\beta \in [0.32, 0.46]$ (assuming $n \geq 7$). In an implementation, one can employ heuristic methods to reduce the potential function Φ as much as possible. Should heuristics fail to provide a sufficient decrease, one can always use the direction given by (5) and (6) as a backup. This leads to flexible interior-point methods that are guaranteed to converge in polynomial time. However, the TTY potential function is designed to be used with feasible algorithms, i.e., those that generate feasible iterates. To use this potential function with infeasible algorithms, one needs to either impose explicit restrictions on step-sizes, or modify the function (see [9]). Our aim here is to develop a class of potential-functions based on the TTY potential function that are more suitable for infeasible-interior-point methods, and to mimic the result of Theorem 3.1 for such functions. Motivated by this goal, we now study the TTY potential function for the homogeneous self-dual formulation.

If a primal-dual algorithm starts at the same initial point for the primal and dual HSDPs and uses the same steps in primal and dual spaces, all the primal iterates will be identical to the dual ones. Then, for the (HSDP)

the primal and dual nonnegative vectors, say $\tilde{x}$ and $\tilde{s}$, are (x, τ, s, κ) and (s, κ, x, τ) . Therefore, the TTY potential function for the (HSDP) is:

$$\begin{aligned}
\Phi_{\tilde{\rho}}(\tilde{x}, \tilde{s}) &= \tilde{\rho} \ln(2(x^T s + \tau \kappa)) - 2 \sum_{i=1}^n \ln x_i s_i - 2 \ln \tau \kappa \\
&= \tilde{\rho} \ln(2(x_0^T s_0 + \kappa_0) \theta) - 2 \left(\sum_{i=1}^n \ln x_i s_i + \ln \tau \kappa \right) \\
&= 2(\rho \ln \theta - \sum_{i=1}^n \ln x_i s_i - \ln \tau \kappa) + \tilde{\rho} \ln(2(x_0^T s_0 + \kappa_0)).
\end{aligned}$$

Here, $\rho = \frac{\tilde{\rho}}{2}$ and the second equality follows from Proposition 2.1(i). Because of self-duality, the analogue of the system (5) for the homogeneous self-dual formulation reduces to the following:

$$\begin{array}{rclcl}
& A \Delta x & - & b \Delta \tau & + & \bar{b} \Delta \theta & & = & 0 \\
- & A^T \Delta y & & + & c \Delta \tau & - & \bar{c} \Delta \theta & - & \Delta s & = & 0 \\
& b^T \Delta y & - & c^T \Delta x & & + & \bar{z} \Delta \theta & & - & \Delta \kappa & = & 0 \\
- & \bar{b}^T \Delta y & + & \bar{c}^T \Delta x & - & \bar{z} \Delta \tau & & & & & = & 0, \\
& & & S \Delta x & & & + & X \Delta s & & & = & \alpha u \\
& & & \kappa \Delta \tau & & & & + & \tau \Delta \kappa & = & \alpha \nu.
\end{array} \quad (10)$$

Here, we have

$$\alpha = \frac{\beta \sqrt{\min([XSe, \tau \kappa])}}{\sqrt{2} \left\| \begin{bmatrix} XS & 0 \\ 0 & \tau \kappa \end{bmatrix}^{-\frac{1}{2}} \left(\frac{2(x^T s + \tau \kappa)}{\rho'} e - \begin{bmatrix} XS & 0 \\ 0 & \tau \kappa \end{bmatrix} e \right) \right\|}, \quad (11)$$

with $\beta \in (0, 1)$, $u = (\frac{2(x^T s + \tau \kappa)}{\rho'} e - XSe)$, $\nu = (\frac{2(x^T s + \tau \kappa)}{\rho'} - \tau \kappa)$, and ρ' may or may not be different from $\tilde{\rho}$, as above. The factor $\sqrt{2}$ appears in (11) because for any $(n+1)$ -vector p we have

$$\left\| (\text{diag}(XSe, \tau \kappa, SXe, \kappa \tau))^{-\frac{1}{2}} \begin{pmatrix} p \\ p \end{pmatrix} \right\| = \sqrt{2} \left\| (\text{diag}(XSe, \tau \kappa))^{-\frac{1}{2}} p \right\|.$$

We now define

$$\Psi_{\rho}(x, s, \kappa, \tau, \theta) := \rho \ln \theta - \sum_{i=1}^n \ln x_i - \sum_{i=1}^n \ln s_i - \ln \tau - \ln \kappa. \quad (12)$$

Clearly, $\Psi_\rho(x, s, \kappa, \tau, \theta) = \frac{1}{2}\Phi_{2\rho}(\tilde{x}, \tilde{s}) + C$, where C is a constant independent of the other variables. Therefore, the Kojima-Mizuno-Yoshise direction would guarantee a constant reduction in the value of $\Psi_\rho(x, s, \kappa, \tau, \theta)$ if $2\rho \geq 2n + 2 + \sqrt{2n + 2}$, or $\rho \geq n + 1 + \sqrt{\frac{n+1}{2}}$.

In the (HSDP) formulation, the variable τ is used for homogenization of the primal and dual feasibility constraints. Therefore, as illustrated in Theorem 2.1, values of the variables $\frac{1}{\tau^k}(x^k, y^k, s^k)$ are more directly related to the solutions of the original problems (LP) and (LD) than (x^k, y^k, s^k) . Thus, we prefer a potential function that will be used in association with the (HSDP) formulation to be a function of the scaled variables. Following Todd [14], we choose $\rho = 2n + 2$ so that Ψ_ρ is invariant under scaling of the variables:

$$\begin{aligned}\Psi_{2n+2}(x, s, \kappa, \tau, \theta) &= (2n + 2) \ln \theta - \sum_{i=1}^n \ln x_i - \sum_{i=1}^n \ln s_i - \ln \tau - \ln \kappa \\ &= (2n + 2) \ln \frac{\theta}{\tau} - \sum_{i=1}^n \ln \frac{x_i}{\tau} - \sum_{i=1}^n \ln \frac{s_i}{\tau} - \ln \frac{\kappa}{\tau} \\ &= \Psi_{2n+2}\left(\frac{x}{\tau}, \frac{s}{\tau}, \frac{\kappa}{\tau}, 1, \frac{\theta}{\tau}\right).\end{aligned}$$

We will denote the scaled vectors $(\frac{x}{\tau}, \frac{y}{\tau}, \frac{s}{\tau}, \frac{\kappa}{\tau}, \frac{\theta}{\tau})$ by $(\bar{x}, \bar{y}, \bar{s}, \bar{\kappa}, \bar{\theta})$, and carry over subscripts and superscripts similarly. Since $2n + 2 > n + 1 + \sqrt{\frac{n+1}{2}}$, $\Psi(\bar{x}, \bar{s}, \bar{\kappa}, 1, \bar{\theta})$ can also be reduced by a constant amount at each iteration. More generally we have the following result:

Lemma 3.1 *Consider a feasible solution $(y, x, \tau, \theta, s, \kappa)$ of the (HSDP), and let $(y^+, x^+, \tau^+, \theta^+, s^+, \kappa^+) = (y, x, \tau, \theta, s, \kappa) + (\Delta y, \Delta x, \Delta \tau, \Delta \theta, \Delta s, \Delta \kappa)$, where the correction vector is a solution to (10) with $\rho' = 4n + 4$ and $\beta \in [0.32, 0.46]$. Also assume that $f(\cdot)$ is a positive real-valued function of the scaled vectors satisfying*

$$\frac{f(\bar{x}, \bar{y}, \bar{s}, \bar{\kappa}, \bar{\theta})}{f(\bar{x}^+, \bar{y}^+, \bar{s}^+, \bar{\kappa}^+, \bar{\theta}^+)} = \frac{\bar{\theta}}{\bar{\theta}^+} = \frac{\frac{\theta}{\tau}}{\frac{\theta^+}{\tau^+}}.$$

Define

$$\phi_f(\bar{x}, \bar{y}, \bar{s}, \bar{\kappa}, \bar{\theta}) := (2n + 2) \ln f(\bar{x}, \bar{y}, \bar{s}, \bar{\kappa}, \bar{\theta}) - \sum_{i=1}^n \ln \bar{x}_i - \sum_{i=1}^n \ln \bar{s}_i - \ln \bar{\kappa}$$

Then, the following holds:

$$\phi_f(\bar{x}^+, \bar{y}^+, \bar{s}^+, \bar{\kappa}^+, \bar{\theta}^+) - \phi_f(\bar{x}, \bar{y}, \bar{s}, \bar{\kappa}, \bar{\theta}) \leq -0.1 \quad (13)$$

PROOF:

$$\begin{aligned} & \phi_f(\bar{x}^+, \bar{y}^+, \bar{s}^+, \bar{\kappa}^+, \bar{\theta}^+) - \phi_f(\bar{x}, \bar{y}, \bar{s}, \bar{\kappa}, \bar{\theta}) \\ &= (2n+2) \ln \frac{f(\bar{x}^+, \bar{y}^+, \bar{s}^+, \bar{\kappa}^+, \bar{\theta}^+)}{f(\bar{x}, \bar{y}, \bar{s}, \bar{\kappa}, \bar{\theta})} - \sum_{i=1}^n \ln \frac{\bar{x}_i^+}{\bar{x}_i} - \sum_{i=1}^n \ln \frac{\bar{s}_i^+}{\bar{s}_i} - \ln \frac{\bar{\kappa}^+}{\bar{\kappa}} \\ &= (2n+2) \ln \frac{\frac{\theta^+}{\tau^+}}{\frac{\theta}{\tau}} - \sum_{i=1}^n \ln \frac{\frac{x_i^+}{\tau^+}}{\frac{x_i}{\tau}} - \sum_{i=1}^n \ln \frac{\frac{s_i^+}{\tau^+}}{\frac{s_i}{\tau}} - \ln \frac{\frac{\kappa^+}{\tau^+}}{\frac{\kappa}{\tau}} \\ &= \Psi_{2n+2}(x^+, s^+, \kappa^+, \tau^+, \theta^+) - \Psi_{2n+2}(x, s, \kappa, \tau, \theta) \\ &= \frac{1}{2}(\Phi_{4n+4}(\tilde{x}^+, \tilde{s}^+) - \Phi_{4n+4}(\tilde{x}, \tilde{s})) \leq \frac{1}{2}(-\frac{\sqrt{3}}{2}\beta + \frac{\beta^2}{2(1-\beta)}) \leq -0.1 \quad \square \end{aligned}$$

Remark 3.1 If $(\Delta y, \Delta x, \Delta \tau, \Delta \theta, \Delta s, \Delta \kappa)$ above solves (10) with $\rho' = 2n+2$ and, in addition, if $\frac{\min([XSe, \kappa\tau])}{2(x^T s + \kappa\tau)} \leq (2n+2)^{-2}$ holds then the argument above and (9) indicate that

$$\phi_f(\bar{x}^+, \bar{y}^+, \bar{s}^+, \bar{\kappa}^+, \bar{\theta}^+) - \phi_f(\bar{x}, \bar{y}, \bar{s}, \bar{\kappa}, \bar{\theta}) \leq -\sqrt{\frac{2n+1}{2n+3}}\beta + \frac{\beta^2}{2(1-\beta)} \quad (14)$$

By Proposition 2.1(ii), $f_1 = \|A\bar{x} - b\|$, $f_2 = \|A^T \bar{y} + \bar{s} - c\|$, or $f_3 = c^T \bar{x} + \bar{\kappa} - b^T \bar{y}$ will be appropriate to be used as f . Moreover, one can use any positively homogeneous function (of degree 1) of f_1 , f_2 , and f_3 such as linear combinations and order statistics. When used to solve the (HSDP) problem directly, all of these functions will decrease equally in each iteration. However, it is essential to observe that a function that combines f_1 , f_2 , and f_3 is independent of the variables τ and θ . Thus, we can use such potential functions without having to use the (HSDP) formulation, as long as we generate the usual iterates (x^k, y^k, s^k) and the auxiliary variable κ^k . At each iteration of the algorithm, one can set up the (HSDP) formulation with the current iterate serving as the $(x_0, y_0, s_0, \kappa_0)$ and choosing $\tau_0 = \theta_0 = 1$. (This is why we choose subscript 0 for the initial solution of the (HSDP) but superscripts for our iterates.) Then, the above analysis shows that any potential function satisfying the hypothesis of Lemma 3.1, can be reduced by at least a constant amount from any infeasible-interior-point by taking a step in the Kojima-Mizuno-Yoshise direction.

4 An infeasible-point primal-dual potential function

In the rest of the paper we will focus on the following potential function that comes from the family of functions described in the previous section:

$$\begin{aligned} \psi(x, y, s, \kappa) \quad := \quad & (2n + 2) \ln \max\{\|Ax - b\|, \|A^T y + s - c\|, c^T x + \kappa - b^T y\} \\ & - \sum_{i=1}^n \ln x_i - \sum_{i=1}^n \ln s_i - \ln \kappa. \end{aligned} \tag{15}$$

Here and henceforth unbarred variables correspond to trial solutions to (LP) and (LD). Let us first comment on why we have chosen this particular function from the family described in the previous section. This function brings together three of the most important quantities for an infeasible-interior-point method: primal and dual infeasibilities, and the duality gap. We want to decrease these three measures to zero simultaneously and if the iterates stay bounded, that is exactly what we obtain by driving ψ to $-\infty$.

Loosely speaking, as we drive the maximum of these three measures to zero, if any of them were to get too small prematurely, say if the iterates were to approach primal feasibility without approaching dual feasibility or optimality, the algorithm can back off from that point to get a better reduction in the potential function, i.e., these measures need not decrease monotonically. One immediate practical implication of this flexibility is that an algorithm based on such a potential function can be expected to perform better on warm starts where the infeasibilities may be large compared to the duality gap.

If we were to introduce positive constant multipliers for both of the infeasibility terms in (15), the resulting function would still be of the form in Lemma 3.1. Since we can use this flexibility to scale the three measures, we can assume without loss of generality that they are initially equal. This scaling combined with a termination criteria of the form $\max\{\|Ax - b\|, \|A^T y + s - c\|, c^T x + \kappa - b^T y\} < \epsilon$ corresponds to terminating when the infeasibilities and the duality gap are sufficiently small *relative* to their initial values. Alternatively, rows and columns of A , and b , c , x^0 , y^0 , and s^0 can be scaled to make infeasibilities and the duality gap equal at the initial point.

Now, we establish some important properties of the potential function ψ that form the basis of its use in interior-point methods. First, we show that ψ can actually be driven to $-\infty$.

Theorem 4.1 *The potential function ψ is unbounded below as a function of (x, y, s, κ) .*

PROOF: As argued in the previous section, any potential function that satisfies the hypothesis of Lemma 3.1, and in particular ψ , can be reduced by at least 0.1 from any infeasible-interior-point. Starting from any point, one can do this repeatedly to obtain a point with arbitrarily small potential value. Therefore, ψ must be unbounded below. $\square$

Now, it remains to investigate the behavior of the iterates as the potential function approaches $-\infty$. First, the following well-known result is presented.

Lemma 4.1 *Let x and s be n -dimensional positive vectors and τ and κ be positive numbers. Then*

$$\sum_{i=1}^n \ln x_i + \sum_{i=1}^n \ln s_i + \ln \tau + \ln \kappa \leq (n+1) \ln(x^T s + \tau \kappa) - (n+1) \ln(n+1).$$

PROOF: By the arithmetic mean-geometric mean inequality,

$$\left(\left(\prod_{i=1}^n x_i s_i \right) \cdot \tau \kappa \right)^{\frac{1}{n+1}} \leq \frac{(\sum_{i=1}^n x_i s_i) + \tau \kappa}{n+1},$$

Now, taking logarithms on both sides produces the result. $\square$

We will also need the following approximate version of the Farkas lemma due to Todd and Ye [16]. For this result we conventionally set $0 \cdot +\infty$ and $+\infty \cdot 0$ to 1, and the minimum over an empty set to $+\infty$.

Lemma 4.2 (Todd and Ye [16]) : *Let*

$$\begin{aligned} \alpha_x &:= \min\{\|x\| : Ax = b, x \geq 0\}, \\ \alpha_y &:= \min\{\|y\| : A^T y + s = c, s \geq 0\} \end{aligned}$$

and

$$\begin{aligned} \beta_u &:= \min\{\|u\| : A^T y + s = u, b^T y = 1, s \geq 0\}, \\ \beta_v &:= \min\{\|v\| : Ax = v, c^T x = -1, x \geq 0\}. \end{aligned}$$

Then $\alpha_x \beta_u = \alpha_y \beta_v = 1$. $\square$

By the Farkas lemma, the primal problem (the dual problem, respectively) is infeasible if and only if β_u (respectively, β_v) is zero. Using this result Todd and Ye argue that if β_u is small but nonzero, then there is a neighboring problem that is primal infeasible. A similar result holds for the dual problem when β_v is small.

We now present the main result about the potential function ψ . Here, we prove that as long as a minimal centrality condition is satisfied as the potential function is being decreased, the iterates will either approach an approximate optimal solution or an approximate proof of infeasibility:

Theorem 4.2 *Let $\epsilon < \min\{\|b\|, \|c\|\}$ and λ be small positive numbers and let $\gamma > 1$ be a large positive number. Assume that an algorithm generates iterates $z^k = (x^k, y^k, s^k, \kappa^k)$ such that $\psi^k := \psi(z^k) \rightarrow -\infty$ and $\kappa^k \geq \lambda \frac{(x^k)^T s^k + \kappa^k}{n+1}$. When*

$$\psi^k \leq (2n+2) \ln \epsilon + (n+1) \ln \frac{\lambda}{5\gamma \max\{\|b\|, \|c\|\}} \quad (16)$$

either

$$\max\{\|Ax^k - b\|, \|A^T y^k + s^k - c\|, c^T x^k - b^T y^k\} \leq \epsilon, \quad (17)$$

or for any primal feasible x we have that

$$\|x\| > \gamma, \quad (18)$$

or for any dual feasible (y, s) we have that

$$\|(y, s)\| > \gamma. \quad (19)$$

PROOF: Define $\mu^k := \frac{(x^k)^T s^k + \kappa^k}{n+1}$. By the definition of ψ , we have that

$$\begin{aligned} (2n+2) \ln \max\{\|Ax^k - b\|, \|A^T y^k + s^k - c\|, c^T x^k + \kappa^k - b^T y^k\} \\ = \psi^k + \sum_{i=1}^n \ln x_i^k + \sum_{i=1}^n \ln s_i^k + \ln \kappa^k \\ \leq \psi^k + (n+1) \ln \mu^k. \end{aligned}$$

Therefore,

$$\max\{\|Ax^k - b\|, \|A^T y^k + s^k - c\|, c^T x^k + \kappa^k - b^T y^k\} \leq \sqrt{\mu^k} \exp\left\{\frac{\psi^k}{2n+2}\right\}. \quad (20)$$

Now, we consider two cases. When

$$\mu^k \leq \exp\left\{-\frac{\psi^k}{n+1}\right\} \epsilon^2 \quad (21)$$

(20) implies (17). Note that this conclusion does not require (16). Now, we assume that the iterate z^k does not satisfy (21) and let β_u and β_v be as in Lemma 4.2. In this case, we will prove that at least one of these two numbers is very small. Combining the inequality $\kappa^k \geq \lambda\mu^k$ and (20) we see that one of the following two inequalities must be satisfied:

$$-c^T x^k \geq \frac{1}{2}(\lambda\mu^k - \sqrt{\mu^k} \exp\left\{\frac{\psi^k}{2n+2}\right\}) > 0, \quad (22)$$

$$b^T y^k \geq \frac{1}{2}(\lambda\mu^k - \sqrt{\mu^k} \exp\left\{\frac{\psi^k}{2n+2}\right\}) > 0. \quad (23)$$

To see the positivity of these expressions we proceed as follows. From (16),

$$\exp\left\{\frac{\psi^k}{n+1}\right\} \leq \frac{\lambda\epsilon^2}{5\gamma \max\{\|b\|, \|c\|\}} < \frac{\lambda\epsilon}{5\gamma}. \quad (24)$$

Thus, since (21) fails,

$$\begin{aligned} \lambda\mu^k - \sqrt{\mu^k} \exp\left\{\frac{\psi^k}{2n+2}\right\} &= \sqrt{\mu^k}(\lambda\sqrt{\mu^k} - \exp\left\{\frac{\psi^k}{2n+2}\right\}) \\ &\geq \sqrt{\mu^k}(\lambda\epsilon \exp\left\{-\frac{\psi^k}{2n+2}\right\} - \exp\left\{\frac{\psi^k}{2n+2}\right\}) \\ &= \sqrt{\mu^k} \exp\left\{\frac{\psi^k}{2n+2}\right\}(\lambda\epsilon \exp\left\{-\frac{\psi^k}{n+1}\right\} - 1) \\ &> \sqrt{\mu^k} \exp\left\{\frac{\psi^k}{2n+2}\right\}(5\gamma - 1) > 0, \end{aligned} \quad (25)$$

where the second inequality uses (24). We first assume that (22) holds. From (20) we also get:

$$\|Ax^k\| \leq \|b\| + \sqrt{\mu^k} \exp\left\{\frac{\psi^k}{2n+2}\right\} = \sqrt{\mu^k} \left(\frac{\|b\|}{\sqrt{\mu^k}} + \exp\left\{\frac{\psi^k}{2n+2}\right\}\right)$$

$$\leq \sqrt{\mu^k} \left(\frac{\|b\|}{\epsilon} + 1 \right) \exp\left\{ \frac{\psi^k}{2n+2} \right\}. \quad (26)$$

Putting (22), (26) and (25) together we obtain:

$$\begin{aligned} \beta_v \leq \|A \frac{x^k}{-c^T x^k}\| &\leq \frac{(\frac{\|b\|}{\epsilon} + 1) \exp\left\{ \frac{\psi^k}{2n+2} \right\}}{\frac{1}{2}(\lambda \sqrt{\mu^k} - \exp\left\{ \frac{\psi^k}{2n+2} \right\})} \\ &\leq \frac{(\frac{\|b\|}{\epsilon} + 1) \exp\left\{ \frac{\psi^k}{2n+2} \right\}}{\frac{1}{2}(\lambda \epsilon \exp\left\{ -\frac{\psi^k}{2n+2} \right\} - \exp\left\{ \frac{\psi^k}{2n+2} \right\})} \\ &= \frac{2(\frac{\|b\|}{\epsilon} + 1)}{\lambda \epsilon \exp\left\{ -\frac{\psi^k}{n+1} \right\} - 1} \\ &< \frac{5\|b\|}{\lambda \epsilon^2 \exp\left\{ -\frac{\psi^k}{n+1} \right\}} \\ &\leq \frac{1}{\gamma}. \end{aligned}$$

The strict inequality above can be verified by cross-multiplying the two ratios, using $\epsilon < \min\{\|b\|, \|c\|\}$, $\gamma > 1$, (16), and $\mu^k > \exp\left\{ -\frac{\psi^k}{n+1} \right\} \epsilon^2$. The last inequality follows from (16). Now, Lemma 4.2 indicates that $\alpha_y > \gamma$. So, any feasible solution to the dual problem must have large norm, i.e., (19) holds indicating that either the dual problem is infeasible or close to being infeasible. Similarly, if (23) holds, then we can show that $\beta_u < \frac{1}{\gamma}$ and that $\alpha_x > \gamma$. In this case (18) holds. This completes the proof. $\square$

Corollary 4.1 *Assume that an algorithm generates iterates $z^k = (x^k, y^k, s^k, \kappa^k)$ such that $\kappa^k \geq \lambda \frac{(x^k)^T s^k + \kappa^k}{n+1}$ and*

$$\psi(z^{k+1}) - \psi(z^k) \leq -\frac{\delta}{(n+1)^r}$$

in each iteration k , for constants $\delta > 0$ and $r \geq 0$. Then, either (17) or (18) or (19) will be satisfied within

$$K = \frac{(n+1)^{1+r}}{\delta} \left(2 \ln \frac{1}{\epsilon} + \ln \frac{5\gamma \max\{\|b\|, \|c\|\}}{\lambda} + \frac{\psi^0}{n+1} \right)$$

iterations. $\square$

In Theorem 4.2, (17) is similar to the Stopping Rule $3a_p$ in Section 3 of [16], (18) and (19) are similar to the conclusions implied by Stopping Rule $3a_d$ of the same paper. Ideally, we would like to replace $c^T x - b^T y$ in (17) with $x^T s$, i.e. to obtain an iterate that satisfies

$$\max\{\|Ax^k - b\|, \|A^T y^k + s^k - c\|, x^T s\} \leq \epsilon, \quad (27)$$

or a strong indication of infeasibility, such as (18) or (19). As Todd and Ye argue, it is hard to get a polynomial bound on the number of iterations in this case. As a matter of fact, they provide only pseudo-polynomial bounds on the number of iterations for Potra's [10] and Mizuno's [8] algorithms, when they are not terminated until such conclusions can be obtained. In our case, following their recipe we can do the following: If we have an iterate that satisfies (16) and violates (21), we conclude that any feasible solution for (LP) and (LD) must have large norm. If, however, (21) is satisfied then $c^T x^k - b^T y^k < \epsilon$, i.e. reverse weak duality almost holds. If μ^k is large in this case, then we must have an indication of infeasibility. Indeed, if

$$\frac{2\gamma(\max\{\|b\|, \|c\|\} + \epsilon) + \epsilon}{\lambda} < \mu^k \leq \exp\left\{-\frac{\psi^k}{n+1}\right\} \epsilon^2,$$

then $\sqrt{\mu^k} \exp\left\{\frac{\psi^k}{2n+2}\right\} \leq \epsilon$, so that

$$\lambda\mu^k - \sqrt{\mu^k} \exp\left\{\frac{\psi^k}{2n+2}\right\} \geq \lambda\mu^k - \epsilon > 0.$$

So, either (22) or (23) holds, and we can then show that (18) or (19) must be satisfied. If (22) holds, then

$$\begin{aligned} \beta_v \leq \left\| A \frac{x^k}{-c^T x^k} \right\| &\leq \frac{\|b\| + \sqrt{\mu^k} \exp\left\{\frac{\psi^k}{2n+2}\right\}}{\frac{1}{2}(\lambda\mu^k - \sqrt{\mu^k} \exp\left\{\frac{\psi^k}{2n+2}\right\})} \\ &\leq \frac{2(\|b\| + \epsilon)}{\lambda\mu^k - \epsilon} < \frac{1}{\gamma}, \end{aligned}$$

and this implies (19). If $-c^T x^k$ does not satisfy (22), then $b^T y^k$ satisfies (23), and similarly we get that $\beta_u < \frac{1}{\gamma}$, which implies (18). On the other hand, if

$\mu^k \leq \frac{2\gamma(\max\{\|b\|, \|c\|\} + \epsilon) + \epsilon}{\lambda}$, then we can switch to the perturbed problem (with $\bar{b} = Ax^k$ and $\bar{c} = A^T y^k + s^k$) and continue with a feasible-interior-point algorithm. Then in $\mathcal{O}(\sqrt{n} \ln \frac{\gamma \max\{\|b\|, \|c\|\}}{\lambda \epsilon})$ additional iterations we can obtain an iterate that satisfies (27).

5 A potential-reduction algorithm

Although the homogeneous self-dual formulation described in Section 2 provides important advantages and simplifications for the analysis of interior-point methods, solving the (HSDP) problem directly by a feasible-interior-point method introduces some limitations as well. In particular, the search direction $(\Delta y, \Delta x, \Delta \tau, \Delta \theta, \Delta s, \Delta \kappa)$ must satisfy the following equations to maintain feasibility:

$$\begin{array}{rclcl} & A\Delta x & - & b\Delta\tau & + & \bar{b}\Delta\theta & & = & 0 \\ - & A^T\Delta y & & + & c\Delta\tau & - & \bar{c}\Delta\theta & - & \Delta s & = & 0 \end{array} \quad (28)$$

The variables τ and θ link the directions in the primal and dual spaces. This dictates the use of equal step sizes for primal and dual updates in feasible algorithms, since otherwise the next iterate would not be feasible. However, most of the state-of-the-art implementations use different step sizes for primal and dual updates [6, 7].

Xu, Hung, and Ye use an infeasible-interior-point method to solve the simplified (HLF) problem. Their formulation eliminates the θ variable and relaxes feasibility so that different steps in primal and dual spaces are allowed. This flexibility is incorporated into their implementation. Nevertheless, their theoretical analysis assumes that equal steps are taken for the primal and dual updates in each iteration. In our case, this assumption is not necessary. Our method is similar to that of Xu et al. in the sense that $\bar{b}$, $\bar{c}$, and $\bar{z}$ are modified at the beginning of each iteration. We construct a new (HSDP) with the current iterate being a strictly feasible solution for the problem. If the κ is not too far from the “center”, we use the Kojima-Mizuno-Yoshise update, otherwise we choose a centering direction. We note that these directions guarantee reductions in the potential function and maintain the λ -condition given in Theorem 4.1; however, one can use heuristic methods in order to get

better reductions in the potential function and use the directions above only when heuristics fail.

Now we give a formal description of the algorithm:

Algorithm 1 Choose a tolerance level $\epsilon < \min\{\|b\|, \|c\|\}$, a centrality parameter $\lambda < (2n + 2)^{-2}$, and a large number $\gamma > 1$. Choose an initial point $x^0 > 0$, $y^0, s^0 > 0$, and $\kappa^0 > 0$ that satisfies $\kappa^0 \geq \lambda \frac{(x^0)^T s^0 + \kappa^0}{n+1}$. Set $\tau^0 = 1$, $\theta^0 = 1$, and $k = 0$.

- 1) If $\max\{\|Ax^k - b\|, \|A^T y^k + s^k - c\|, c^T x^k - b^T y^k\} < \epsilon$, then STOP.
- 2) If $c^T x^k < 0$ and $\frac{\|Ax^k\|}{-c^T x^k} < \frac{1}{\gamma}$, or if $b^T y^k > 0$ and $\frac{\|A^T y^k + s^k\|}{b^T y^k} < \frac{1}{\gamma}$, then STOP.
- 3) Construct the (HSDP) formulation with $x_0 = x^k$, $y_0 = y^k$, $s\kappa^k \geq 2\lambda \frac{(x^k)^T s^k + \kappa^k}{n+1}$ then let $(\Delta y, \Delta x, \Delta \tau, \Delta \theta, \Delta s, \Delta \kappa)$ be the unique solution to (10) with $\rho' = 4n+4$ and $\beta = \frac{1}{3}$. Otherwise let $(\Delta y, \Delta x, \Delta \tau, \Delta \theta, \Delta s, \Delta \kappa)$ be the unique solution to (10) with $\rho' = 2n+2$ and $\beta = \frac{0.9}{\sqrt{2n+2}}$. Define

$$\begin{aligned}\delta x &= \frac{x_0 + \Delta x}{1 + \Delta \tau} - x_0, \\ \delta y &= \frac{y_0 + \Delta y}{1 + \Delta \tau} - y_0, \\ \delta s &= \frac{s_0 + \Delta s}{1 + \Delta \tau} - s_0, \\ \delta \kappa &= \frac{\kappa_0 + \Delta \kappa}{1 + \Delta \tau} - \kappa_0.\end{aligned}$$

- 4) Find (or approximate) numbers $\bar{\alpha}_P, \bar{\alpha}_D, \bar{\alpha}_\kappa$ that

$$\begin{aligned}\text{minimize} \quad & \psi(x^k + \alpha_P \delta x, y^k + \alpha_D \delta y, s^k + \alpha_D \delta s, \kappa^k + \alpha_\kappa \delta \kappa) \\ \text{subject to:} \quad & (x^k + \alpha_P \delta x, s^k + \alpha_D \delta s, \kappa^k + \alpha_\kappa \delta \kappa) > 0, \\ & \kappa^k + \alpha_\kappa \delta \kappa \geq \lambda \frac{(x^k + \alpha_P \delta x)^T (s^k + \alpha_D \delta s) + \kappa^k + \alpha_\kappa \delta \kappa}{n+1}.\end{aligned}$$

Set

$$(x^{k+1}, y^{k+1}, s^{k+1}, \kappa^{k+1}) = (x^k + \bar{\alpha}_P \delta x, y^k + \bar{\alpha}_D \delta y, s^k + \bar{\alpha}_D \delta s, \kappa^k + \bar{\alpha}_\kappa \delta \kappa).$$

5) Set $k=k+1$, go to step 1.

If the algorithm stops in step 1, then we obtain approximate optimal solutions for (LP) and (LD). If the algorithm stops in step 2, then we either conclude that any primal feasible x must satisfy $\|x\| > \gamma$ or that any dual feasible (y, s) must satisfy $\|(y, s)\| > \gamma$. In order to get a complexity bound for the algorithm, we need to estimate the minimum decrease in the potential function in step 4.

Choosing $\alpha_P = \alpha_D = \alpha_\kappa = 1$ one gets $x^{k+1} = \frac{x^k + \Delta x}{1 + \Delta\tau}$, etc. But $(y^k + \Delta y, x^k + \Delta x, 1 + \Delta\tau, 1 + \Delta\theta, s^k + \Delta s, \kappa^k + \Delta\kappa)$ is the next iterate for the current (HSDP) generated either using the Kojima-Mizuno-Yoshise direction ($\rho' = 4n + 4$) or the centering direction ($\rho' = 2n + 2$). In what follows we will prove that if the λ -condition is satisfied even when λ is replaced by 2λ , then the Kojima-Mizuno-Yoshise direction will generate a next iterate that still satisfies the λ -condition and decreases the potential function by at least 0.1. Also, we will establish that if the λ -condition is satisfied tightly, i.e. if $\lambda\mu \leq \kappa < 2\lambda\mu$, then a centering direction will generate a new iterate that decreases the potential function by at least $\frac{0.1}{\sqrt{n+1}}$ and still satisfies the λ -condition. This argument and Corollary 4.1 indicate that the algorithm will terminate within $\mathcal{O}(n^{\frac{3}{2}} \ln \frac{n\gamma}{\epsilon})$ iterations. In an implementation of this algorithm, one can start a 3-dimensional search for a “good” step-size vector $(\bar{\alpha}_P, \bar{\alpha}_D, \bar{\alpha}_\kappa)$ from $(1, 1, 1)$.

Before presenting our results, we introduce the following notation, for a given iterate (x, y, s, κ) :

$$V = \begin{bmatrix} (XS)^{\frac{1}{2}} \\ \kappa^{\frac{1}{2}} \end{bmatrix}, \quad v = Ve, \quad v_{\min} = \min_i(v_i) = \sqrt{\min([XSe, \kappa])},$$

$$w = \frac{x^T s + \kappa}{2n + 2} (V^{-1}e - \frac{2n + 2}{x^T s + \kappa} v), \quad \text{and} \quad w' = \frac{x^T s + \kappa}{n + 1} (V^{-1}e - \frac{n + 1}{x^T s + \kappa} v).$$

We will assume throughout that $n \geq 2$. First, we present two simple lemmas:

Lemma 5.1 *Given $(y, x, \tau = 1, \theta = 1, s, \kappa)$ let $(\Delta y, \Delta x, \Delta\tau, \Delta\theta, \Delta s, \Delta\kappa)$ be the unique solution to (10). Then*

$$|\Delta x_i \Delta s_i| \leq \frac{\beta^2 v_{\min}^2}{2}, \quad i = 1, \dots, n, \quad (29)$$

$$|\Delta\kappa\Delta\tau| \leq \frac{\beta^2 v_{\min}^2}{2}. \quad (30)$$

PROOF: Premultiplying the last two inequalities in (10) by V^{-1} we obtain

$$\begin{bmatrix} X^{-\frac{1}{2}} S^{\frac{1}{2}} \Delta x \\ \kappa^{-\frac{1}{2}} \tau^{\frac{1}{2}} \Delta \kappa \end{bmatrix} + \begin{bmatrix} S^{-\frac{1}{2}} X^{\frac{1}{2}} \Delta s \\ \tau^{-\frac{1}{2}} \kappa^{\frac{1}{2}} \Delta \tau \end{bmatrix} = \frac{\beta v_{\min}}{\sqrt{2}} \frac{V^{-1}e - \frac{\rho'}{2(x^T s + \tau \kappa)}v}{\|V^{-1}e - \frac{\rho'}{2(x^T s + \tau \kappa)}v\|}. \quad (31)$$

Furthermore, the rest of the equations in (10) indicate that

$$\begin{bmatrix} X^{-\frac{1}{2}} S^{\frac{1}{2}} \Delta x \\ \kappa^{-\frac{1}{2}} \tau^{\frac{1}{2}} \Delta \kappa \end{bmatrix}^T \begin{bmatrix} S^{-\frac{1}{2}} X^{\frac{1}{2}} \Delta s \\ \tau^{-\frac{1}{2}} \kappa^{\frac{1}{2}} \Delta \tau \end{bmatrix} = \Delta x^T \Delta s + \Delta \kappa \Delta \tau = 0.$$

Then (31) can be rewritten as $p + q = r$, with $p^T q = 0$ and $\|r\| = \frac{\beta v_{\min}}{\sqrt{2}}$.

Therefore,

$$\|p\|^2 + \|q\|^2 = \|r\|^2 = \frac{\beta^2 v_{\min}^2}{2},$$

and

$$\begin{aligned} |\Delta x_i| \frac{\sqrt{x_i s_i}}{x_i} &\leq \|p\|_{\infty} \leq \|p\| \leq \frac{\beta v_{\min}}{\sqrt{2}}, \\ |\Delta s_i| \frac{\sqrt{x_i s_i}}{s_i} &\leq \|q\|_{\infty} \leq \|q\| \leq \frac{\beta v_{\min}}{\sqrt{2}}, \end{aligned}$$

and finally

$$|\Delta x_i \Delta s_i| \leq \frac{\beta^2 v_{\min}^2}{2}, \quad i = 1, \dots, n.$$

Similarly, $|\Delta\kappa\Delta\tau| \leq \frac{\beta^2 v_{\min}^2}{2}$. □

Lemma 5.2 For $0 < \beta \leq 0.35$,

$$\frac{1 - \frac{\beta v_{\min}}{\sqrt{2}\|w\|} - \frac{\beta^2}{2}}{1 - \frac{\beta v_{\min}}{2\sqrt{2}\|w\|}} > \frac{1}{2}.$$

PROOF: Define $\xi = \frac{v_{\min}}{\sqrt{2}\|w\|}$. First, we observe that

$$g(\beta, \xi) := \frac{1 - \beta\xi - \frac{\beta^2}{2}}{1 - \frac{\beta\xi}{2}}$$

is a decreasing function of ξ when $\beta > 0$. But,

$$\begin{aligned} \xi &= \frac{(2n+2)}{\sqrt{2}} \frac{v_{\min}^2}{x^T s + \kappa} \frac{1}{\|V^{-1}e - \frac{2n+2}{v^T v} v\| v_{\min}} \\ &\leq \frac{(2n+2)}{\sqrt{2}(n+1)} \frac{1}{\|V^{-1}e - \frac{2n+2}{v^T v} v\| v_{\min}} \\ &\leq \sqrt{2} \frac{2}{\sqrt{3}} = \frac{4}{\sqrt{6}}, \end{aligned}$$

where the last inequality follows from Lemma 2.5 in Kojima et al. [5]. Thus,

$$g(\beta, \xi) \geq \bar{g}(\beta) := g\left(\beta, \frac{4}{\sqrt{6}}\right) = \frac{1 - \frac{4\beta}{\sqrt{6}} - \frac{\beta^2}{2}}{1 - \frac{2\beta}{\sqrt{6}}}.$$

We now observe that $\bar{g}(\beta)$ is a decreasing function of β when $\beta \in (0, 1)$, and that $\bar{g}(0.35) > 0.5$. Thus, the assertion of the lemma follows. $\square$

Proposition 5.1 *Assume that (x, y, s, κ) satisfying $\kappa \geq 2\lambda \frac{x^T s + \kappa}{n+1}$ is given. Determine $(\delta x, \delta y, \delta s, \delta \kappa)$ as in Step 2 of Algorithm 1, with $\rho' = 4n + 4$ and $\beta = \frac{1}{3}$, and define $(x^+, y^+, s^+, \kappa^+) = (x, y, s, \kappa) + (\delta x, \delta y, \delta s, \delta \kappa)$. Then,*

$$\psi(x^+, y^+, s^+, \kappa^+) - \psi(x, y, s, \kappa) \leq -0.1$$

and

$$\kappa^+ \geq \lambda \frac{(x^+)^T s^+ + \kappa^+}{n+1}.$$

PROOF: Let $(\Delta y, \Delta x, \Delta \tau, \Delta \theta, \Delta s, \Delta \kappa)$ be the solution to (10) with $\rho' = 4n + 4$ and $\beta = \frac{1}{3}$. Define

$$(\hat{y}, \hat{x}, \hat{\tau}, \hat{\theta}, \hat{s}, \hat{\kappa}) := (y, x, 1, 1, s, \kappa) + (\Delta y, \Delta x, \Delta \tau, \Delta \theta, \Delta s, \Delta \kappa).$$

First, note that

$$\frac{\kappa^+}{(x^+)^T s^+ + \kappa^+} = \frac{\hat{\kappa} \hat{\tau}}{\hat{x}^T \hat{s} + \hat{\kappa} \hat{\tau}}. \quad (32)$$

Also,

$$\begin{aligned} \hat{x}^T \hat{s} + \hat{\kappa} \hat{\tau} &= x^T s + \kappa + (x^T \Delta s + s^T \Delta x + \kappa \Delta \tau + \Delta \kappa) \\ &= x^T s + \kappa + \frac{\beta v_{\min}}{\sqrt{2} \|w\|} \left(\frac{x^T s + \kappa}{2n+2} (n+1) - (x^T s + \kappa) \right) \\ &= (x^T s + \kappa) \left(1 - \frac{\beta v_{\min}}{2\sqrt{2} \|w\|} \right). \end{aligned}$$

Note that the first equation uses the identity $\Delta x^T \Delta s + \Delta \kappa \Delta \tau = 0$. On the other hand,

$$\begin{aligned} \hat{\kappa} \hat{\tau} &= \kappa + (\kappa \Delta \tau + \Delta \kappa) + \Delta \kappa \Delta \tau \\ &= \kappa + \frac{\beta v_{\min}}{\sqrt{2} \|w\|} \left(\frac{x^T s + \kappa}{2n+2} - \kappa \right) + \Delta \kappa \Delta \tau \\ &\geq \left(1 - \frac{\beta v_{\min}}{\sqrt{2} \|w\|} \right) \kappa + \Delta \kappa \Delta \tau \\ &\geq \left(1 - \frac{\beta v_{\min}}{\sqrt{2} \|w\|} - \frac{\beta^2}{2} \right) \kappa. \end{aligned}$$

Here the second inequality follows from Lemma 5.1. Now,

$$\frac{\kappa^+}{(x^+)^T s^+ + \kappa^+} \geq \frac{1 - \frac{\beta v_{\min}}{\sqrt{2} \|w\|} - \frac{\beta^2}{2}}{1 - \frac{\beta v_{\min}}{2\sqrt{2} \|w\|}} \frac{\kappa}{x^T s + \kappa} \geq \frac{\lambda}{n+1}. \quad (33)$$

The last part of the inequality above follows from Lemma 5.2. Finally, by Lemma 3.1 we have that

$$\psi(x^+, y^+, s^+, \kappa^+) - \psi(x, y, s, \kappa) \leq -0.1$$

This completes the proof of the proposition. $\square$

Proposition 5.2 Assume that (x, y, s, κ) satisfying $2\lambda \frac{x^T s + \kappa}{n+1} > \kappa \geq \lambda \frac{x^T s + \kappa}{n+1}$ is given. Determine $(\delta x, \delta y, \delta s, \delta \kappa)$ as in Step 2 of Algorithm 1, with $\rho' = 2n+2$ and $\beta = \frac{9}{\sqrt{2n+2}}$, and define $(x^+, y^+, s^+, \kappa^+) = (x, y, s, \kappa) + (\delta x, \delta y, \delta s, \delta \kappa)$. Then,

$$\psi(x^+, y^+, s^+, \kappa^+) - \psi(x, y, s, \kappa) \leq -\frac{0.1}{\sqrt{n+1}}$$

and

$$\kappa^+ \geq \lambda \frac{(x^+)^T s^+ + \kappa^+}{n+1}.$$

PROOF: We define $(\hat{y}, \hat{x}, \hat{\tau}, \hat{\theta}, \hat{s}, \hat{\kappa})$ as in the proof of Proposition 5.1, so that (32) still holds. In this case,

$$\begin{aligned} \hat{x}^T \hat{s} + \hat{\kappa} \hat{\tau} &= x^T s + \kappa + \frac{\beta v_{\min}}{\sqrt{2} \|w'\|} \left(\frac{x^T s + \kappa}{n+1} (n+1) - (x^T s + \kappa) \right) \\ &= x^T s + \kappa. \end{aligned} \quad (34)$$

To bound $\hat{\kappa} \hat{\tau}$ from below, we will need the following inequality:

$$\begin{aligned} \|w'\|^2 &= \mu^2 \left(\sum_{i=1}^n \left(\frac{1}{\sqrt{x_i s_i}} - \frac{\sqrt{x_i s_i}}{\mu} \right)^2 + \left(\frac{1}{\sqrt{\kappa}} - \frac{\sqrt{\kappa}}{\mu} \right)^2 \right) \\ &= \mu^2 \left(\sum_{i=1}^n \frac{1}{x_i s_i} + \frac{1}{\kappa} - \frac{n+1}{\mu} \right) \leq \frac{\mu^2 (n+1)}{v_{\min}^2}. \end{aligned} \quad (35)$$

Now,

$$\begin{aligned} \hat{\kappa} \hat{\tau} &= \kappa + \frac{\beta v_{\min}}{\sqrt{2} \|w'\|} \left(\frac{x^T s + \kappa}{n+1} - \kappa \right) + \Delta \kappa \Delta \tau \\ &\geq \kappa + \frac{\beta v_{\min}}{\sqrt{2} \|w'\|} \left(\frac{1}{2\lambda} - 1 \right) \kappa - \frac{\beta^2 v_{\min}^2}{2} \\ &= \kappa + \beta v_{\min}^2 \left(\frac{\kappa}{\sqrt{2} \|w'\| v_{\min}} \left(\frac{1}{2\lambda} - 1 \right) - \frac{\beta}{2} \right) \\ &\geq \kappa + \beta v_{\min}^2 \left(\frac{\lambda}{\sqrt{2n+2}} \left(\frac{1}{2\lambda} - 1 \right) - \frac{\beta}{2} \right) \\ &\geq \kappa. \end{aligned}$$

The last two inequalities follow from $\kappa > \lambda\mu$, (35), and $\beta = \frac{.9}{\sqrt{2n+2}} < \frac{1-2\lambda}{\sqrt{2n+2}}$ (assuming $n \geq 2$ and recalling $\lambda < (2n+2)^{-2}$). Now, $\hat{\kappa}\hat{\tau} \geq \kappa$, (34), and (32) imply that $\kappa^+ \geq \lambda \frac{(x^+)^T s^+ + \kappa^+}{n+1}$.

Finally, we look at the decrease in the potential function. Since

$$\frac{\min([XSe, \kappa])}{2(x^T s + \kappa)} \leq \frac{\kappa}{2(x^T s + \kappa)} < \lambda < (2n+2)^{-2},$$

(14) implies the following:

$$\begin{aligned} \psi(x^+, y^+, s^+, \kappa^+) - \psi(x, y, s, \kappa) &\leq \frac{1}{2} \left(-\sqrt{\frac{2n+1}{2n+3}} \beta + \frac{\beta^2}{2(1-\beta)} \right) \\ &< \frac{1}{2} \left(-\frac{0.69}{\sqrt{2n+2}} + \frac{0.81}{2(2n+2)(1-\frac{0.9}{\sqrt{2n+2}})} \right) < -\frac{0.1}{\sqrt{n+1}}. \end{aligned}$$

This concludes the proof. $\square$

To conclude, we note that Proposition 5.1 and Proposition 5.2 also show how the TTY potential function for the (HSDP) formulation can be reduced by at least $\mathcal{O}(\frac{1}{\sqrt{n}})$ maintaining the weak centrality condition on κ and τ . So, one can use the strategy outlined in the proofs of these two propositions to solve the (HSDP) problem directly, using the TTY potential function, and centering only the artificial variables κ and τ . Although the worst-case complexity of such an algorithm would be worse than just using, say, a path following algorithm, the flexibility of not having to follow the central path, and the ability to achieve conclusive information about the status of the problem, makes it a more attractive choice for implementations.

ACKNOWLEDGMENT

The work that led to this paper was initiated and has developed through several discussions I had with Michael Todd. He also read a previous draft of the paper and made some useful suggestions to improve the presentation. I gratefully acknowledge his input and encouragement. I am solely responsible for any errors.

References

- [1] A. J. Goldman and A. W. Tucker, "Theory of linear programming," in: H.W. Kuhn and A.W. Tucker eds., *Linear Inequalities and Related Systems*, Princeton University Press, Princeton, NJ (1956) pp. 53-97.
- [2] C. C. Gonzaga, "Path following methods for linear programming," *SIAM Review* 34 (1992) pp. 167-227.
- [3] O. Güler and Y. Ye, "Convergence behavior of interior-point algorithms," *Mathematical Programming*, 60 (1993) pp. 215-228.
- [4] N. K. Karmarkar, "A new polynomial time algorithm for linear programming," *Combinatorica*, 4 (1984) pp. 373-395.
- [5] M. Kojima, S. Mizuno, and A. Yoshise, "An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementarity problems," *Mathematical Programming*, 50 (1991) pp. 331-342.
- [6] I. J. Lustig, R. E. Marsten, and D. F. Shanno, "On implementing Mehrotra's predictor-corrector interior-point method for linear programming," *SIAM Journal on Optimization*, 2 (1992) pp 435-449.
- [7] S. Mehrotra, "On the implementation of a primal-dual interior point method," *SIAM Journal on Optimization*, 2 (1992) pp. 575-601.
- [8] S. Mizuno, "Polynomiality of infeasible-interior-point algorithms for linear programming," Technical Report, The Institute of Statistical Mathematics, (Tokyo, Japan, 1993).
- [9] S. Mizuno, M. Kojima, and M. J. Todd, "Infeasible-interior-point primal-dual potential-reduction algorithms for linear programming," *SIAM Journal on Optimization*, 5 (1995) pp. 52-67.
- [10] F. Potra, "An infeasible interior-point predictor-corrector algorithm for linear programming," Reports on Computational Mathematics No. 26, University of Iowa, Iowa City, IA, June 1992, revised August 1994, to appear in *SIAM Journal on Optimization*.

- [11] J. Renegar, "A polynomial-time algorithm based on Newton's method for linear programming," *Mathematical Programming*, 40 (1988) pp. 59-94.
- [12] K. Tanabe, "Centered Newton method for mathematical programming," in: M. Iri and K. Yajima, eds., *Lecture Notes in Control and Information Sciences*, Springer-Verlag, Berlin (1988) pp. 197-206.
- [13] M. J. Todd, "Potential-reduction methods in mathematical programming," TR-1112 (1995), School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY.
- [14] M. J. Todd, Private communication.
- [15] M. J. Todd and Y. Ye, "A centered projective algorithm for linear programming," *Mathematics of Operations Research*, 15 (1990) pp. 508-529.
- [16] M. J. Todd and Y. Ye, "Approximate Farkas lemmas and stopping rules for iterative infeasible-point algorithms for linear programming," TR-1109 (1994), School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, revised January 1996.
- [17] A. W. Tucker, "Dual systems of homogeneous linear relations," in: H.W. Kuhn and A.W. Tucker eds., *Linear Inequalities and Related Systems*, Princeton University Press, Princeton, NJ (1956) pp. 3-18.
- [18] X. Xu, P-F. Hung, and Y. Ye, "A simplified homogeneous and self-dual linear programming algorithm and its implementation," Manuscript, Department of Management Sciences, University of Iowa, (Iowa City, Iowa, 1993).
- [19] Y. Ye, "An $O(n^3L)$ potential-reduction algorithm for linear programming," *Mathematical Programming*, 50 (1991) pp. 239-258.
- [20] Y. Ye, K. O. Kortanek, J. Kaliski, and S. Huang, "Near boundary behavior of primal-dual potential-reduction algorithms for linear programming," *Mathematical Programming*, 58 (1993) pp. 243-255.

- [21] Y. Ye, M. J. Todd, and S. Mizuno, “An $O(\sqrt{n}L)$ -iteration homogeneous and self-dual linear programming algorithm,” *Mathematics of Operations Research*, 19 (1994) pp. 53-67.