Chapter 2

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

2.1 Introduction

This chapter studies a new potential function and an infeasible-interior-point method based on this function for the solution of linear programming problems. Infeasible-interior-point potential-reduction methods (IIPPRMs) form an understudied branch of the interior-point methods family. However, they provide a nice framework to derive polynomial convergence results for flexible and practical algorithms that use several implementation tricks. This property of IIPPRMs motivates the work in this chapter that is intended to narrow the apparent gap between those algorithms with the best worst-case complexity and those with the most successful implementations.

Analyses of interior-point 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 analyses 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.

We consider linear programming problems (LPs) given in the following standard form:

$$
(\mathcal{LP}) \qquad \begin{array}{rcl} \min_x & c^T x & \\ Ax & = & b \\ x & \geq & 0, \end{array} \tag{2.1}
$$

where $A \in I\!\!R^{m \times n}$, $b \in I\!\!R^m$, $c \in I\!\!R^n$ are given, and $x \in I\!\!R^n$ is the variable to be determined as the solution to the problem. The matrix A is assumed to have full row rank. This is done without loss of generality because if A does not have full row rank, the augmented matrix $[A|b]$ can be row reduced, which either reveals that the problem is infeasible or that one can continue with the reduced full-rank matrix.

$\mathcal{LP}$ is said to be *feasible* if its constraints are consistent and it is called *unbounded* if there exists a sequence of feasible vectors $\{x^k\}$ such that $c^T x^k \to -\infty$. When $\mathcal{LP}$ is feasible but not unbounded it has an *optimal solution*, i.e. a vector x that satisfies the constraints and minimizes the objective value among all feasible vectors.

There is a dual problem which is associated with $(\mathcal{LP})$:

$$
\begin{array}{llrcrcl}
(\mathcal{LD}) & \max_{y,s} & b^T y & & & & \\
 & & A^T y & + & s & = & c \\
 & & & & s & \geq & 0,
\end{array}
\tag{2.2}
$$

where $y \in I\!\!R^m$ and $s \in I\!\!R^n$. The objective value of any primal feasible solution is at least as large as the objective value of any feasible dual solution; this fact is known as the *weak duality theorem.* Furthermore, when both problems are feasible, their optimal values coincide. This latter fact is the *strong duality theorem.* We refer the reader to, e.g., [Chv83] for further details on the intriguing relationship between $(\mathcal{LP})$ and $(\mathcal{LD})$.

Although Karmarkar's original work on interior-point methods [Kar84] used a potential-reduction algorithm, in later years, especially following Renegar's work [Ren88], the focus of research on interior-point methods shifted to path-following algorithms. Meanwhile, potential-reduction algorithms received consistent but sporadic attention. See recent papers by Todd [Tod95] and Anstreicher [Ans96] for reviews of potential-reduction algorithms for mathematical programming problems.

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 developments on path-following methods [Gon92]. 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" [Tod95]. Our purpose in this chapter is to develop an algorithm that is as close to the ones used in practice as possible and to establish polynomial convergence results for this algorithm; therefore we focus on potential-reduction methods, which seem to serve this purpose better.

Tanabe [Tan88] and Todd and Ye [TY90] independently introduced and analyzed a primal-dual potential function $\Phi_\rho(x, s)$ that has been the main tool of recent potential-reduction algorithms. We will give a definition of $\Phi_\rho(x, s)$ in Section 2.3. A succession of articles showed that this potential function can be reduced by at least a constant amount in one iteration starting from any feasible interior point, for large enough values of ρ [TY90,Ye91,KMY91]. Since feasible iterates approach optimality as the potential function approaches $-\infty$, recipes for guaranteeing reduction in the potential function also guarantee global polynomial convergence. However, there is a drawback of these methods; they are required to start at a strictly feasible initial solution and obtaining such a solution is often a hard task. Even worse, the feasible region (or its interior) may be empty for particular instances of the LP problem. Indeed, finding a feasible initial solution is a fundamental problem of many different families of algorithms for LPs.

Several techniques have been developed to address the issue of starting interior-point methods. One classical method is to augment the problem by introducing artificial variables in such a way that the new problem has a known feasible starting solution. In these methods artificial variables are priced expensively, and they are expected to eventually disappear from the formulation if the problem is feasible. Such methods are known as *big M* methods. The disadvantages of these methods include the numerical problems caused by the huge prices (big Ms) on the artificial variables and slow convergence. On the other hand, in *two-phase* methods one first solves an artificial problem which again has a known feasible solution. Then, the solution to this *Phase I* problem either indicates that the LP instance is infeasible or produces a feasible solution which is then used to start the *Phase II*, where the original problem is solved. The main disadvantage of these methods is the need to solve two LP problems sequentially.

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. As discussed in Chapter 1, these methods are known as infeasible-interior-point methods and seem to be the most popular choice for practical implementations.

Interestingly, most of the infeasible-interior-point methods in the literature are path-following algorithms [KMM93,Miz94,Zha94,Wri94], although there also exist some infeasible-interior-point potential-reduction algorithms. This imbalance can be attributed to the intractability of the theoretical properties of infeasible-interior-point potential-reduction methods. One notable example of such methods is Mizuno, Kojima, and Todd's $O(nL)$-iteration primal-dual algorithm [MKT95]. 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 difficult to obtain. Here, we describe a new potential function and an infeasible-interior-point algorithm based on this function suggesting a way around these difficulties. A related work that deserves mentioning is the constant-potential infeasible-interior-point method of Seifi and Tunçel [ST96].

Their method generalizes the work of Tunçel [Tun94] where a framework was introduced for solving LPs by staying on the contours of a particular potential function. [ST96] also provides results of computational experiments.

One important issue that we deal with in this chapter is the conclusion that can be obtained from an output of an interior-point method. An algorithm working with exact arithmetic is said to *process* an LP instance if it either determines that there is no solution x satisfying the constraints of (2.1), or finds a direction in the feasible region along which the objective function can be decreased indefinitely (equivalently, shows that (2.1) is feasible but (2.2) is infeasible), or finds an x satisfying the constraints and minimizing the objective value. Practical methods do not use exact arithmetic and therefore only a strong indication of one of these three conclusions can be expected from an implementation of interior-point methods.

Current infeasible-interior-point methods (IIPMs) 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 [TY94], 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 [YTM94]. The use of the homogeneous self-dual formulation not only provides a framework to replicate the theoretical results for (feasible) IPMs in the case of IIPMs but also lets us avoid regularity assumptions on the problem and big M parameters. In particular, we show how a rich family of potential functions defined on the 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 [KMY91] 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 chapter is organized as follows. Following this introduction, Section 2.2 reviews the homogeneous self-dual formulation of Ye, Todd, and Mizuno [YTM94]. Section 2.3 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 2.4. Section 2.5 describes our algorithm and derives the complexity results. Section 2.6 discusses possible extensions of the ideas and the algorithm developed in the previous sections to monotone linear complementarity problems and concludes the chapter. 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.

2.2 Homogeneous Self-Dual LP Formulation

A common approach for solving the problems $(\mathcal{LP})$ and $(\mathcal{LD})$ 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 (HLF) model developed by Goldman and Tucker [GT56,Tuc56]:

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

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 $(\mathcal{HLF})$ with $\tau^* > 0$ one can obtain optimal solutions for $(\mathcal{LP})$ and $(\mathcal{LD})$ 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). Indeed, if, for example, $A^T y^* = 0$ and $b^T y^* > 0$ then a direct application of Farkas' lemma [Far94,Far98] indicates that $Ax = b$, $x \geq 0$ is an inconsistent system.

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. This happens when the constraint matrix is square and skew-symmetric, the cost vector is the negative of the right-hand-side vector, equality constraints correspond to free variables, and greater-than-or-equal-to constraints correspond to nonnegative variables. Thus, $(\mathcal{HLF})$ can be viewed as a self-dual LP with 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 [YTM94]. Given any $x_0 > 0$, $s_0 > 0$, $\kappa_0 > 0$, and y_0, they formulate an artificial homogeneous self-dual linear program (HSDP):

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

where

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

Note the appearance of the artificial variable θ in addition to the homogenizing variable τ. But, this is not as bad as in the big M methods since the cost coefficient of θ is a user-determined number rather than a large number determined by the data. Also, note that in this chapter when we refer to a homogeneous problem we mean a problem with possibly one nonhomogeneous (normalizing) constraint.

Let s and κ be the nonnegative surplus variables for the second and third set of constraints. For any optimal solution $(y^*, x^*, \tau^*, \theta^*, s^*, \kappa^*)$ of the $\mathcal{HSDP}$ $\theta^* = 0$ (see [YTM94]), and in this case the variables satisfy the same equations as in $\mathcal{HLF}$, with the addition of a normalizing constraint. Thus, the comments above apply to the solutions of the $\mathcal{HSDP}$ 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 $O(\sqrt{n} \ln \frac{1}{\varepsilon})$ complexity, can be used to solve this problem. Note that Ye, Todd, and Mizuno 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 $\mathcal{HLF}$ and using infeasible-interior-point algorithms [XHY93]. They also reported very favorable computational results from an implementation of their algorithm. However, their theoretical analysis ignores the tricks employed in their implementation such as the use of different step-sizes in primal and dual spaces . This is precisely what we would like to avoid; we intend to give a theoretical justification of the fact that heuristic tricks do not destroy the robustness and polynomial convergence guarantees of certain algorithms.

The homogeneous self-dual formulation offers 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, nondegeneracy, 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 $\mathcal{HSDP}$ 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. The condition $\tau^* + \kappa^* > 0$ 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 2.2.1 (Ye-Todd-Mizuno [YTM94]) *: Let* $(y^*, x^*, \tau^*, \theta^* = 0, s^*, \kappa^*)$ *be a strictly self-complementary solution for the* $\mathcal{HSDP}$*.*

(i) $\mathcal{LP}$ *has an optimal solution if and only if* $\tau^* > 0$*. In this case,* x^*/τ^* *is an optimal solution for* $\mathcal{LP}$ *and* $(y^*/\tau^*, s^*/\tau^*)$ *is an optimal solution for* $\mathcal{LD}$*.*

(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* $\mathcal{LD}$ *is infeasible; if* $-b^T y^* < 0$ *then* $\mathcal{LP}$ *is infeasible; and if both* $c^T x^* < 0$ *and* $-b^T y^* < 0$ *then both* $\mathcal{LP}$ *and* $\mathcal{LD}$ *are infeasible.* □

As established by Güler and Ye [GY93] 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 $\mathcal{HSDP}$, satisfying the hypothesis of Theorem 2.2.1. However, this is more than one needs. As stated above, producing a solution satisfying $\tau^* + \kappa^* > 0$ 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 2.5 we will study an algorithm that drops strict complementarity conditions on x and s.

We now state some basic properties of this formulation:

Proposition 2.2.1 *Let* $(y, x, \tau, \theta, s, \kappa)$ *and* $(y^+, x^+, \tau^+, \theta^+, s^+, \kappa^+)$ *be feasible for the problem (*$\mathcal{HSDP}$*) 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) $x^T s + \tau\kappa = \theta(x_0^T s_0 + \kappa_0)$.

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

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

Proof:

The proof of (i) is given in [YTM94] 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 proofs of (ii) and (iii) are straightforward. $\square$

From Proposition 2.2.1(iii), 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 the total *uncomplementarity*, $\bar{x}^T\bar{s} + \bar{\kappa}$. Both of these expressions can be considered as measures of optimality.

2.3 A family of potential functions

The use of potential functions in conjunction with interior-point methods for mathematical programming has been popularized by Karmarkar's original work in 1984 [Kar84]. The function Karmarkar uses is a primal-only potential function meaning that it only measures the quality of a primal iterate. In 1988, Tanabe [Tan88], and Todd and Ye [TY90] independently introduced and analyzed a primal-dual potential function that turned out to be a very useful tool for construction and analysis of feasible-interior-point algorithms for linear programming and linear complementarity problems. For the primal-dual pair of problems (2.1) and (2.2), this function is defined as

$$\Phi_\rho(x, s) = \rho \ln(x^T s) - \sum_{i=1}^{n} \ln x_i s_i, \tag{2.3}$$

for every $(x, s) > 0$.

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 $(x, s) > 0$ [KMY91]. It is interesting to note that the search direction that they use in their analysis to guarantee reduction in the potential function comes from primal-dual path-following algorithms. This direction can also 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:

$$\begin{array}{rcrcl} & A\Delta x & & = & 0 \\ A^T\Delta y & & + \quad \Delta s & = & 0 \\ & S\Delta x & + \quad X\Delta s & = & \alpha(\frac{x^T s}{\rho'}e - XSe), \end{array} \tag{2.4}$$

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

$$\alpha = \frac{\beta\sqrt{\min(XSe)}}{\left\|(XS)^{-\frac{1}{2}}(\frac{x^T s}{\rho'}e - XSe)\right\|}, \text{ with } \beta \in (0,1), \tag{2.5}$$

$\min(v) = \min\{v_1, v_2, \ldots, v_n\}$ for $v \in I\!R^n$, and ρ' is a positive parameter related to (or equal to) ρ. Let us first explain how the above system arises. The first and second equations of the system (2.4) are introduced to maintain primal and dual feasibility. The third equation can be seen as a linear combination of

$$S\Delta x \quad + \quad X\Delta s \quad = \quad -XSe, \tag{2.6}$$

and

$$S\Delta x \quad + \quad X\Delta s \quad = \quad \frac{x^T s}{n}e - XSe, \tag{2.7}$$

The feasibility equations combined with (2.6) define the so-called *affine-scaling* direction that corresponds to the Newton equation for estimating a solution of the nonlinear system

$$\begin{array}{rcrcl} & A\hat{x} & & = & b \\ A^T\hat{y} & & + \quad \hat{s} & = & c \\ & \hat{X}\hat{S}e & & = & 0 \end{array} \tag{2.8}$$

from the current point (x, y, s). The system (2.8) augmented with nonnegativity constraints on $\hat{x}$ and $\hat{s}$ defines the optimality conditions for $\mathcal{LP}$. On the other hand, combining the feasibility equations with (2.7) yields a system defining the *centering* direction. Similarly, this system corresponds to the Newton equation for estimating a solution of the system

$$\begin{array}{rcccl} & A\hat{x} & & = & b \\ A^T\hat{y} & & + \ \hat{s} & = & c \\ & \hat{X}\hat{S}e & & = & \mu e. \end{array} \tag{2.9}$$

where $\mu = \dfrac{x^T s}{n}$. Consider the system obtained from (2.9) by adding nonnegativity constraints on $\hat{x}$ and $\hat{s}$. For different values of the parameter $\mu > 0$, the solutions to this system form a nice differentiable curve in the interior of the feasible region. This curve is what is known as the *central path*. Finally, the constant α is introduced to make sure that $(x, s) + (\Delta x, \Delta s) > 0$. Therefore, (2.4) can be seen as a centered and damped Newton equation.

The solution to the system (2.4) is unique since the left hand side matrix $\begin{bmatrix} 0 & A & 0 \\ A^T & 0 & I \\ 0 & S & X \end{bmatrix}$ is nonsingular when A has full row rank and X and S are positive diagonal matrices.

Kojima, Mizuno, and Yoshise choose the direction parameter ρ' to be equal to the potential function parameter ρ. 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 [YKKH93]. The significance of this last result comes from the fact that when 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 2.3.1 (Ye et al. [YKKH93]) : *Let x and (y, s) be feasible for $\mathcal{LP}$ and $\mathcal{LD}$ respectively, and let $(\Delta x, \Delta y, \Delta s)$ satisfy (2.4) and (2.5). Then $x^+ = (x + \Delta x)$ and $(y^+, s^+) = (y + \Delta y, s + \Delta s)$ are feasible for $\mathcal{LP}$ and $\mathcal{LD}$, and*

$$\Phi_\rho(x^+, s^+) - \Phi_\rho(x, s) \leq -\sqrt{\frac{\min(XSe)}{x^T s}} \frac{(\rho' - n)(\rho - n) + h_{x,s}}{\sqrt{(\rho' - n)^2 + h_{x,s}}} \beta + \frac{\beta^2}{2(1-\beta)}, \quad (2.10)$$

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

It is possible to simplify the right-hand-side of the above inequality when ρ' is large enough, or when $\rho' = n$ but the iterate is not well-centered. More precisely, when $\rho' = \rho \geq n + \sqrt{n}$, Kojima et al. [KMY91] show that

$$\Phi_\rho(x^+, s^+) - \Phi_\rho(x, s) \leq -\frac{\sqrt{3}}{2}\beta + \frac{\beta^2}{2(1-\beta)}, \quad (2.11)$$

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

$$\Phi_\rho(x^+, s^+) - \Phi_\rho(x, s) \leq -\sqrt{\frac{n-1}{n+1}}\beta + \frac{\beta^2}{2(1-\beta)} \quad (2.12)$$

using Lemma 5 in [GY93] 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 the heuristics fail to provide a sufficient decrease, one can always use the direction given by (2.4) and (2.5) as a backup. This leads to flexible interior-point methods that are guaranteed to converge in polynomial time. However, as we mentioned above, 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 IIPPRMs, one needs to either impose explicit restrictions on step-sizes, or modify the function (see [MKT95]). Our aim here is to develop a class of potential-functions based on the TTY potential function that are more suitable for IIPPRMs, and to mimic the result of Theorem 2.3.1 for such functions. Motivated by this goal, we now study the TTY potential function for the homogeneous self-dual formulation.

As we observed above, 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 $\mathcal{HSDP}$ the primal and dual nonnegative vectors, say $\tilde{x}$ and $\tilde{s}$, are (x, τ, s, κ) and (s, κ, x, τ), and the TTY potential function is:

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

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 (2.4) for the homogeneous self-dual formulation reduces to the following:

$$
\begin{array}{rrrrrrrrrrrr}
 & & A\Delta x & - & b\Delta\tau & + & \bar{b}\Delta\theta & & & & & = & 0 \\
- & A^T\Delta y & & + & c\Delta\tau & - & \bar{c}\Delta\theta & - & \Delta s & & & = & 0 \\
 & b^T\Delta y & - & c^T\Delta x & & + & \bar{z}\Delta\theta & & & - & \Delta\kappa & = & 0 \\
- & \bar{b}^T\Delta y & + & \bar{c}^T\Delta x & - & \bar{z}\Delta\tau & & & & & & = & 0, \\
 & & S\Delta x & & & & & + & X\Delta s & & & = & \alpha u \\
 & & & & \kappa\Delta\tau & & & & & + & \tau\Delta\kappa & = & \alpha\nu.
\end{array}
\qquad (2.13)
$$

Here, we have

$$\alpha = \frac{\beta\sqrt{\min([XSe,\tau\kappa])}}{\sqrt{2}\left\|(\mathrm{diag}(XSe,\tau\kappa))^{-\frac{1}{2}}(\frac{2(x^Ts+\tau\kappa)}{\rho'}e-(\mathrm{diag}(XSe,\tau\kappa))e)\right\|}, \tag{2.14}$$

with $\beta \in (0,1)$, $u = (\frac{2(x^Ts+\tau\kappa)}{\rho'}e - XSe)$, $\nu = (\frac{2(x^Ts+\tau\kappa)}{\rho'} - \tau\kappa)$, and ρ' may or may not be different from $\tilde{\rho}$, as above. The factor $\sqrt{2}$ appears in (2.14) because for any $(n+1)$-dimensional vector p we have

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

We now define

$$\Psi_\rho(x,s,\kappa,\tau,\theta) := \rho\ln\theta - \sum_{i=1}^{n}\ln x_i - \sum_{i=1}^{n}\ln s_i - \ln\tau - \ln\kappa. \tag{2.15}$$

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 2.2.1, the 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 $\mathcal{LP}$ and $\mathcal{LD}$ than (x^k,y^k,s^k). Consequently, 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 [Tod94], we choose $\rho = 2n+2$ so that Ψ_ρ is invariant under scaling of the variables:

$$\Psi_{2n+2}(x,s,\kappa,\tau,\theta) \;=\; (2n+2)\ln\theta - \sum_{i=1}^{n}\ln x_i - \sum_{i=1}^{n}\ln s_i - \ln\tau - \ln\kappa$$

$$= \quad (2n+2)\ln\frac{\theta}{\tau} - \sum_{i=1}^{n}\ln\frac{x_i}{\tau} - \sum_{i=1}^{n}\ln\frac{s_i}{\tau} - \ln\frac{\kappa}{\tau}$$
$$= \quad \Psi_{2n+2}(\frac{x}{\tau}, \frac{s}{\tau}, \frac{\kappa}{\tau}, 1, \frac{\theta}{\tau}).$$

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 2.3.1 *Consider a feasible solution $(y, x, \tau, \theta, s, \kappa)$ of $\mathcal{HSDP}$, 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 (2.13) 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*

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

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:

$$\phi_f(\bar{x}^+, \bar{y}^+, \bar{s}^+, \bar{\kappa}^+, \bar{\theta}^+) - \phi_f(\bar{x}, \bar{y}, \bar{s}, \bar{\kappa}, \bar{\theta}) \leq -0.1. \tag{2.16}$$

Proof:

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

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

Remark 2.3.1 *If $(\Delta y, \Delta x, \Delta\tau, \Delta\theta, \Delta s, \Delta\kappa)$ above solves (2.13) with $\rho' = 2n+2$ and, in addition, if $\frac{\min([XSe, \kappa\tau])}{2(x^Ts+\kappa\tau)} \le (2n+2)^{-2}$ holds then the argument above and (2.12) indicate that*

$$
\phi_f(\bar{x}^+, \bar{y}^+, \bar{s}^+, \bar{\kappa}^+, \bar{\theta}^+) - \phi_f(\bar{x}, \bar{y}, \bar{s}, \bar{\kappa}, \bar{\theta}) \quad \le \quad -\sqrt{\frac{2n+1}{2n+3}}\frac{\beta}{2} + \frac{\beta^2}{4(1-\beta)}. \tag{2.17}
$$

By Proposition 2.2.1(ii),

$$
f_1 = \|A\bar{x} - b\|,
$$

$$
f_2 = \|A^T\bar{y} + \bar{s} - c\|,
$$

$$
f_3 = c^T\bar{x} + \bar{\kappa} - b^T\bar{y}
$$

are appropriate to be used as the function f of Lemma 2.3.1 and inequality (2.17). Moreover, one can use any positively homogeneous function (of degree 1) of f_1, f_2, and f_3 such as their linear combinations, maximum, minimum, etc. When used to solve the $(\mathcal{HSDP})$ 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 $\mathcal{HSDP}$ but superscripts for our iterates.) Then, the above analysis shows that any potential function satisfying the hypothesis of Lemma 2.3.1 can be reduced by at least a constant amount from any infeasible interior point by taking a step in the Kojima-Mizuno-Yoshise direction.

2.4 An infeasible-point primal-dual potential function

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

$$\begin{aligned} \psi(x, y, s, \kappa) \;\; &:= \;\; (2n+2)\ln\max\{\|Ax-b\|, \|A^Ty+s-c\|, c^Tx+\kappa-b^Ty\} \\ &\qquad - \textstyle\sum_{i=1}^n \ln x_i - \sum_{i=1}^n \ln s_i - \ln\kappa. \end{aligned} \tag{2.18}$$

Here and henceforth unbarred variables correspond to trial solutions to $\mathcal{LP}$ and $\mathcal{LD}$. 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 -∞.

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 (2.18), the resulting function would still be of the form in Lemma 2.3.1. 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 $\max\{\|Ax-b\|, \|A^Ty+s-c\|, c^Tx+\kappa-b^Ty\} < \varepsilon$ 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 for its use in interior-point methods. First, we show that ψ can actually be driven to $-\infty$.

Theorem 2.4.1 *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 2.3.1, 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. □

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 2.4.1 *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,

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

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 [TY94]. 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 2.4.2 (Todd and Ye [TY94]) *: 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. This is a useful result for the study of IPMs since the solutions or proofs of infeasibilities generated by IPMs are often approximate values and/or are contaminated by round-off errors, and the information they provide is often of the form, e.g., β_u is small, instead of $\beta_u = 0$.

We now present the main result about the potential function ψ. Here, we prove that if 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. One point that needs to be emphasized is that results we develop in this section depend only on the potential function ψ, and not on any particular algorithm that may use it.

Theorem 2.4.2 *Let $\varepsilon < \min\{\|b\|, \|c\|\}$ 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 $\psi^k := \psi(z^k) \to -\infty$ and $\kappa^k \geq \lambda \frac{(x^k)^T s^k + \kappa^k}{n+1}$. When*

$$\psi^k \leq (2n+2)\ln\varepsilon + (n+1)\ln\frac{\lambda}{5\gamma\max\{\|b\|, \|c\|\}} \tag{2.19}$$

either

$$\max\{\|Ax^k - b\|, \|A^T y^k + s^k - c\|, c^T x^k - b^T y^k\} \leq \varepsilon, \tag{2.20}$$

or for any primal feasible x we have that

$$\|x\| > \gamma, \tag{2.21}$$

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

$$\|y\| \;>\; \gamma. \tag{2.22}$$

Proof:

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

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

The inequality follows from Lemma 2.4.1. Therefore,

$$\max\{\|Ax^k - b\|, \|A^T y^k + s^k - c\|, c^T x^k + \kappa^k - b^T y^k\} \le \sqrt{\mu^k}\exp\{\frac{\psi^k}{2n+2}\}. \tag{2.23}$$

Now, we consider two cases. We will see that when the potential function is small, if μ^k is also small then we must have approximate solutions, and otherwise either the primal or the dual problem must be almost infeasible. More precisely, when

$$\mu^k \;\le\; \exp\{-\frac{\psi^k}{n+1}\}\varepsilon^2 \tag{2.24}$$

(2.23) implies (2.20) indicating that approximate solutions are obtained for the primal and dual problems. Note that this conclusion does not require (2.19) and an algorithm may be terminated once (2.24) is satisfied without waiting for the potential function to get small enough. On the other hand, if we assume that the iterate z^k does not satisfy (2.24) and let β_u and β_v be as in Lemma 2.4.2 we can prove that at least one of these two numbers is very small. Combining the inequality $\kappa^k \geq \lambda\mu^k$ and (2.23) we see that one of the following two inequalities must be satisfied:

$$\begin{aligned} -c^T x^k &\geq \frac{1}{2}(\lambda\mu^k - \sqrt{\mu^k}\exp\{\frac{\psi^k}{2n+2}\}) > 0, && (2.25)\\ b^T y^k &\geq \frac{1}{2}(\lambda\mu^k - \sqrt{\mu^k}\exp\{\frac{\psi^k}{2n+2}\}) > 0. && (2.26)\end{aligned}$$

To see the positivity of these expressions we proceed as follows. From (2.19),

$$\exp\{\frac{\psi^k}{n+1}\} \leq \frac{\lambda\varepsilon^2}{5\gamma\max\{\|b\|,\|c\|\}} < \frac{\lambda\varepsilon}{5\gamma}. \qquad (2.27)$$

Thus, if (2.24) fails,

$$\begin{aligned} \lambda\mu^k - \sqrt{\mu^k}\exp\{\frac{\psi^k}{2n+2}\} &= \sqrt{\mu^k}(\lambda\sqrt{\mu^k} - \exp\{\frac{\psi^k}{2n+2}\})\\ &\geq \sqrt{\mu^k}(\lambda\varepsilon\exp\{-\frac{\psi^k}{2n+2}\} - \exp\{\frac{\psi^k}{2n+2}\})\\ &= \sqrt{\mu^k}\exp\{\frac{\psi^k}{2n+2}\}(\lambda\varepsilon\exp\{-\frac{\psi^k}{n+1}\} - 1)\\ &> \sqrt{\mu^k}\exp\{\frac{\psi^k}{2n+2}\}(5\gamma - 1) > 0, && (2.28)\end{aligned}$$

where the second inequality uses (2.27). We first assume that (2.25) holds. From (2.23) we also get:

$$\begin{aligned} \|Ax^k\| &\leq \|b\| + \sqrt{\mu^k}\exp\{\frac{\psi^k}{2n+2}\} = \sqrt{\mu^k}(\frac{\|b\|}{\sqrt{\mu^k}} + \exp\{\frac{\psi^k}{2n+2}\})\\ &\leq \sqrt{\mu^k}(\frac{\|b\|}{\varepsilon} + 1)\exp\{\frac{\psi^k}{2n+2}\}. && (2.29)\end{aligned}$$

Putting (2.25), (2.29), and (2.28) together we obtain:

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

$$
\begin{aligned}
&= \frac{2(\frac{\|b\|}{\varepsilon}+1)}{\lambda\varepsilon\exp\{-\frac{\psi^k}{n+1}\}-1} \\
&< \frac{5\|b\|}{\lambda\varepsilon^2\exp\{-\frac{\psi^k}{n+1}\}} \\
&\leq \frac{1}{\gamma}.
\end{aligned}
$$

The strict inequality above can be verified by cross-multiplying the two ratios, using $\varepsilon < \min\{\|b\|, \|c\|\}$, $\gamma > 1$, (2.19), and (2.27). The last inequality follows from (2.27). Now, Lemma 2.4.2 indicates that $\alpha_y > \gamma$. So, any feasible solution to the dual problem must have large norm, i.e., (2.22) holds indicating that either the dual problem is infeasible or close to being infeasible. Similarly, if (2.26) holds, then we can show that $\beta_u < \dfrac{1}{\gamma}$ and that $\alpha_x > \gamma$. In this case (2.21) holds. This completes the proof. ∎

Corollary 2.4.1 *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 (2.20) or (2.21) or (2.22) will be satisfied within

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

iterations. ∎

In Theorem 2.4.2, (2.20) is similar to the Stopping Rule 1a in Section 3 of [TY94], while (2.21) and (2.22) are similar to the conclusions implied by Stopping Rules 3a_p and 3a_d of the same paper. Ideally, we would like to replace $c^T x - b^T y$

in (2.20) with x^Ts, i.e., to obtain an iterate that satisfies

$$\max\{\|Ax^k - b\|, \|A^Ty^k + s^k - c\|, x^Ts\} \;\; \leq \;\; \varepsilon, \tag{2.30}$$

or a strong indication of infeasibility, such as (2.21) or (2.22). 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 [Pot92] and Mizuno's [Miz94] 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 (2.19) and violates (2.24), we conclude that any feasible solution for $\mathcal{LP}$ and $\mathcal{LD}$ must have large norm. If, however, (2.24) is satisfied then $c^Tx^k - b^Ty^k < \varepsilon$, 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

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

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

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

So, either (2.25) or (2.26) holds, and we can then show that (2.21) or (2.22) must be satisfied. If (2.25) holds, then

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

and this implies (2.22). If $-c^Tx^k$ does not satisfy (2.25), then b^Ty^k satisfies (2.26), and similarly we get that $\beta_u < \dfrac{1}{\gamma}$, which implies (2.21). On the other hand, if

$$\mu^k \quad \leq \quad \frac{2\gamma(\max\{\|b\|, \|c\|\} + \varepsilon) + \varepsilon}{\lambda}, \tag{2.31}$$

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 if we use Kojima, Mizuno, and Yoshise's feasible potential-reduction algorithm [KMY91], and if the current potential value is $\mathcal{O}(\sqrt{n} \ln \dfrac{\gamma \max\{\|b\|, \|c\|\}}{\lambda\varepsilon})$, then in as many additional iterations we can obtain an iterate that satisfies (2.30).

Remark 2.4.1 *In Theorem 2.4.2 we implicitly assumed that* $\min\{\|b\|, \|c\|\} > 0$*. Note that this is not a restrictive assumption.*

2.5 A potential-reduction algorithm

Although the homogeneous self-dual formulation described in Section 2.2 provides important advantages and simplifications for the analysis of interior-point methods, solving $\mathcal{HSDP}$ directly by a feasible-interior-point method introduces some limitations as well. In particular, recall from (2.13) that the search direction $(\Delta y, \Delta x, \Delta\tau, \Delta\theta, \Delta s, \Delta\kappa)$ must satisfy the following equations to maintain feasibility:

$$\begin{array}{rcccccccccl} & & A\Delta x & - & b\Delta\tau & + & \bar{b}\Delta\theta & & & = & 0 \\ - & A^T\Delta y & & + & c\Delta\tau & - & \bar{c}\Delta\theta & - & \Delta s & = & 0 \end{array} \tag{2.32}$$

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 [LMS92,Meh92].

Xu, Hung, and Ye use an infeasible-interior-point method to solve the simplified homogeneous linear feasibility problem [XHY93]. This formulation eliminates the variable θ and relaxes feasibility constraints 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 homogeneous self-dual formulation with the current iterate being a strictly feasible solution for the problem. If the auxiliary variable 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 2.4.2; however, one can use heuristic methods in order to get better reductions in the potential function and use the directions above only when the heuristics fail.

A formal description of the algorithm is given in Figure 2.1. If the algorithm stops in Step 1, we obtain approximate optimal solutions for both $\mathcal{LP}$ and $\mathcal{LD}$. 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 $\|y\| > \gamma$. In order to get a complexity bound for the algorithm, we need to estimate the minimum decrease in the potential function in Step 4.

Let us first explain how we choose δx, etc. in Step 3. Consider the next iterate

Step 0. Choose a tolerance level $\varepsilon < \min\{\|b\|, \|c\|\}$, 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 $k = 0$.

Step 1. <u>If</u> $(\max\{\|Ax^k - b\|, \|A^T y^k + s^k - c\|, c^T x^k - b^T y^k\} < \varepsilon,)$ <u>then</u>

<u>return</u>(`done`)

<u>End if</u>

Step 2. <u>If</u> ($c^T x^k < 0$ and $\dfrac{\|Ax^k\|}{-c^T x^k} < \dfrac{1}{\gamma}$, or if $b^T y^k > 0$ and $\dfrac{\|A^T y^k + s^k\|}{b^T y^k} < \dfrac{1}{\gamma}$) <u>then</u>

<u>return</u>(`done`)

<u>End if</u>

Step 3. Construct the HSDP formulation with $x_0 = x^k$, $y_0 = y^k$, $s_0 = s^k$, and $\kappa_0 = \kappa^k$. If $\kappa^k \geq 2\lambda \frac{(x^k)^T s^k + \kappa^k}{n+1}$ then let $(\Delta y, \Delta x, \Delta \tau, \Delta \theta, \Delta s, \Delta \kappa)$ be the unique solution to (2.13) 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 (2.13) with $\rho' = 2n + 2$ and $\beta = \frac{0.9}{\sqrt{2n+2}}$. Define

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

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

$$\text{minimize } \psi(x^k + \alpha_P \delta x, y^k + \alpha_D \delta y, s^k + \alpha_D \delta s, \kappa^k + \alpha_\kappa \delta \kappa)$$

$$\text{subject to: } (x^k + \alpha_P \delta x, s^k + \alpha_D \delta s, \kappa^k + \alpha_\kappa \delta \kappa) > 0,$$
$$\kappa^k + \alpha_\kappa \delta \kappa \geq \lambda \frac{(x^k + \alpha_P \delta x)^T (s^k + \alpha_D \delta s) + \kappa^k + \alpha_\kappa \delta \kappa}{n+1}.$$

Let

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

Step 5. Set $k = k + 1$ and go to Step 1.

Figure 2.1: An infeasible-interior-point potential-reduction algorithm

$(y^k+\Delta y, x^k+\Delta x, 1+\Delta\tau, 1+\Delta\theta, s^k+\Delta s, \kappa^k+\Delta\kappa)$ for the current $\mathcal{HSDP}$ generated either using the Kojima-Mizuno-Yoshise direction ($\rho' = 4n+4$) or the centering direction ($\rho' = 2n+2$). In terms of the original variables this solution would correspond to choosing $x^{k+1} = \dfrac{x^k+\Delta x}{1+\Delta\tau}$, etc., i.e., choosing $\alpha_p = \alpha_d = \alpha_\kappa = 1$.

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 $\dfrac{0.1}{\sqrt{n+1}}$ and still satisfies the λ-condition. This argument and Corollary 2.4.1 indicate that the algorithm will terminate within $\mathcal{O}(n^{\frac{3}{2}} \ln \dfrac{n\gamma}{\varepsilon})$ iterations. Therefore, this choice of δx, etc., in Step 3 of the algorithm leads to polynomial complexity results in addition to providing flexibility in choices of α_p, α_d, and α_κ. 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):

$$V := \begin{bmatrix} (XS)^{\frac{1}{2}} & \\ & \kappa^{\frac{1}{2}} \end{bmatrix}, \tag{2.33}$$

$$v := Ve, \tag{2.34}$$

$$v_{\min} := \min_i(v_i) = \sqrt{\min([XSe, \kappa])}, \tag{2.35}$$

$$w := \frac{x^Ts+\kappa}{2n+2}\left(V^{-1}e - \frac{2n+2}{x^Ts+\kappa}v\right), \tag{2.36}$$

$$w' := \frac{x^Ts+\kappa}{n+1}\left(V^{-1}e - \frac{n+1}{x^Ts+\kappa}v\right). \tag{2.37}$$

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

Lemma 2.5.1 *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 (2.13). Then*

$$|\Delta x_i \Delta s_i| \leq \frac{\beta^2 v_{\min}^2}{2}, \quad i = 1, \ldots, n, \tag{2.38}$$

$$|\Delta\kappa\Delta\tau| \leq \frac{\beta^2 v_{\min}^2}{2}. \tag{2.39}$$

Proof:

Premultiplying the last two inequalities in (2.13) by V^{-1} we obtain

$$\begin{bmatrix} X^{-\frac{1}{2}} S^{\frac{1}{2}} \Delta x \\ \kappa^{-\frac{1}{2}} \tau^{\frac{1}{2}} \Delta\kappa \end{bmatrix} + \begin{bmatrix} S^{-\frac{1}{2}} X^{\frac{1}{2}} \Delta s \\ \tau^{-\frac{1}{2}} \kappa^{\frac{1}{2}} \Delta\tau \end{bmatrix} = \frac{\beta v_{\min}}{\sqrt{2}} \frac{V^{-1}e - \frac{\rho'}{2(x^T s + \tau\kappa)} v}{\|V^{-1}e - \frac{\rho'}{2(x^T s + \tau\kappa)} v\|}. \tag{2.40}$$

Furthermore, the rest of the equations in (2.13) 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 (2.40) can be rewritten as $p + q = r$, with $p^T q = 0$ and $\|r\| = \dfrac{\beta v_{\min}}{\sqrt{2}}$. Therefore,

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

and

$$|\Delta x_i| \frac{\sqrt{x_i s_i}}{x_i} \leq \|p\|_\infty \leq \|p\| \leq \frac{\beta v_{\min}}{\sqrt{2}},$$

$$|\Delta s_i| \frac{\sqrt{x_i s_i}}{s_i} \leq \|q\|_\infty \leq \|q\| \leq \frac{\beta v_{\min}}{\sqrt{2}},$$

and finally

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

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

Lemma 2.5.2 *For $0 < \beta \leq 0.35$,*

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

Proof:

Let $\xi := \dfrac{v_{\min}}{\sqrt{2}\|w\|}$. 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. [KMY91]. Thus,

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

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. $\square$

Now we present two propositions that estimate the minimum decrease in the potential function for the Kojima-Mizuno-Yoshise and centering directions, respectively, and also show how these two directions affect the satisfaction of the minimal centrality condition.

Proposition 2.5.1 *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 the algorithm, 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 (2.13) 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

$$\frac{\kappa^+}{(x^+)^T s^+ + \kappa^+} = \frac{\hat{\kappa}\hat{\tau}}{\hat{x}^T\hat{s} + \hat{\kappa}\hat{\tau}}. \tag{2.41}$$

Also, recalling the definition of w in (2.36)

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

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 2.5.1. Now,

$$\frac{\kappa^+}{(x^+)^T s^+ + \kappa^+} \geq \frac{1 - \dfrac{\beta v_{\min}}{\sqrt{2}\|w\|} - \dfrac{\beta^2}{2}}{1 - \dfrac{\beta v_{\min}}{2\sqrt{2}\|w\|}} \frac{\kappa}{x^T s + \kappa} \geq \frac{\lambda}{n+1}. \tag{2.42}$$

The last part of the inequality above follows from Lemma 2.5.2. Finally, by Lemma 2.3.1 we have that

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

This completes the proof of the proposition. □

Proposition 2.5.2 *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, for $n \geq 2$,*

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

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 2.5.1, so that (2.41) still holds. In this case,

$$\begin{aligned} \hat{x}^T \hat{s} + \hat{\kappa} \hat{\tau} &= x^T s + \kappa + \frac{\beta v_{\min}}{\sqrt{2}\|w'\|} \left(\frac{x^T s + \kappa}{n+1}(n+1) - (x^T s + \kappa) \right) \\ &= x^T s + \kappa. \end{aligned} \tag{2.43}$$

To bound $\hat{\kappa}\hat{\tau}$ from below, we will need the following inequality related to the norm of w' defined in (2.37):

$$\begin{aligned}\|w'\|^2 &= \mu^2\Big(\sum_{i=1}^{n}\Big(\frac{1}{\sqrt{x_i s_i}} - \frac{\sqrt{x_i s_i}}{\mu}\Big)^2 + \Big(\frac{1}{\sqrt{\kappa}} - \frac{\sqrt{\kappa}}{\mu}\Big)^2\Big)\\ &= \mu^2\Big(\sum_{i=1}^{n}\frac{1}{x_i s_i} + \frac{1}{\kappa} - \frac{n+1}{\mu}\Big) \le \frac{\mu^2(n+1)}{v_{\min}^2}. \end{aligned} \tag{2.44}$$

Now,

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

The last three inequalities follow from (2.44), $\kappa > \lambda\mu$, and $\beta = \frac{.9}{\sqrt{2n+2}} < \frac{1-2\lambda}{\sqrt{2n+2}}$ (assuming $n \ge 2$ and recalling $\lambda < (2n+2)^{-2}$). Now, $\hat{\kappa}\hat{\tau} \ge \kappa$, (2.43), and (2.41) imply that $\kappa^+ \ge \lambda\frac{(x^+)^T s^+ + \kappa^+}{n+1}$.

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

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

(2.17) implies the following:

$$\begin{aligned}\psi(x^+, y^+, s^+, \kappa^+) - \psi(x, y, s, \kappa) &\le \frac{1}{2}\left(-\sqrt{\frac{2n+1}{2n+3}}\beta + \frac{\beta^2}{2(1-\beta)}\right)\\ < \frac{1}{2}\left(-\frac{0.76}{\sqrt{2n+2}} + \frac{0.81}{2(2n+2)(1-\frac{0.9}{\sqrt{2n+2}})}\right) &< -\frac{0.1}{\sqrt{n+1}}.\end{aligned}$$

This concludes the proof. $\square$

To conclude this section, we note that Proposition 2.5.1 and Proposition 2.5.2 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 $\mathcal{HSDP}$ 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 while preserving the ability to achieve conclusive information about the status of the problem, makes it a more attractive choice for implementations.

2.6 Possible Extensions to Linear Complementarity Problems

In this section we discuss the possibility of extending the results and the algorithm of the previous section to the case of linear complementarity problems (LCP). In particular, we consider two homogeneous formulations of the LCP that allow us to use an infeasible-start potential-reduction algorithm which has the flexibility features of the algorithm we presented in Section 2.5. Complexity estimates, however, do not appear to be extendable to this case easily.

The linear complementarity problem (LCP) is defined as the problem of finding vectors x and s satisfying

$$
\begin{array}{ll}
(\mathcal{LCP}) & s - Mx = q \\
& x \geq 0, s \geq 0, x^T s = 0,
\end{array}
\tag{2.45}
$$

where $M \in \Re^{n\times n}$ and $q \in \Re^n$ are the inputs. When M is positive semidefinite, (2.45) is called a monotone linear complementarity problem. Here we only consider monotone LCPs. We will give a more detailed description of the LCP in Chapter 3.

2.6.1 Homogeneous LCP Formulation

As discussed in Section 2.2, Ye, Todd and Mizuno developed a homogeneous self-dual LP formulation that allows one to use interior-point algorithms to solve any LP starting from *any* (possibly infeasible) interior-point for the problem [YTM94]. A very important feature of this artificial problem is that it has strictly feasible primal and dual solutions even when the underlying LP does not. Thus, one can develop and analyze algorithms for the HSDP without needing to make regularity assumptions, such as the existence of interior solutions or nondegeneracy, about the underlying LP problem.

Soon after, Ye generalized this approach to the case of monotone linear complementarity problems [Ye95]. Starting from $x_0 = e$, the vector of all ones, $s_0 = \sigma e$, $\kappa_0 = \sigma$, $\tau_0 = 1$ with

$$\sigma \quad = \quad \max\{0, \max([M + M^T]e)\} + 1$$

he formulates the following homogeneous and self-dual LCP related to (2.45):

$$\begin{array}{lll}
\min & x^T s + \tau\kappa & \\
(\mathcal{HLCP}) & \begin{pmatrix} s \\ \kappa \\ 0 \end{pmatrix} = \begin{pmatrix} M & q & \bar{r} \\ -q^T & 0 & \bar{z} \\ -\bar{r}^T & -\bar{z} & 0 \end{pmatrix} \begin{pmatrix} x \\ \tau \\ \theta \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \bar{n} \end{pmatrix} & (2.46) \\
 & x, s, \kappa, \tau \geq 0, \theta \text{ free} &
\end{array}$$

where

$$\bar{r} = s_0 - Mx_0 - q,\ \bar{z} = q^T x_0 + \kappa_0,$$

and

$$\bar{n} = x_0^T s_0 + \kappa_0 - x_0^T M x_0 \leq (n+1)\sigma.$$

When $M = \begin{bmatrix} 0 & A \\ -A^T & 0 \end{bmatrix}$, $x = \begin{bmatrix} y \\ x \end{bmatrix}$, and $s = \begin{bmatrix} 0 \\ s \end{bmatrix}$ the above formulation reduces to the homogeneous self-dual linear programming formulation of Section 2.2. Note that for this particular choice, the matrix M is skew-symmetric and therefore is positive semidefinite.

We recall that LPs have strictly complementary solutions when they are bounded and feasible. Unfortunately, LCPs do not share this nice property. Therefore, instead of a strictly complementary solution Ye's algorithm converges to a so-called *maximal* complementarity solution. This, however, complicates his analysis since $\tau^* + \kappa^*$ may not be positive (i.e., may equal zero) as in the LP case.

Alternatively, one can use the following homogeneous LCP model:

$$\begin{array}{ll} \min & x^T s + \tau\kappa \\ (\mathcal{HLCP}2) & \begin{pmatrix} s \\ \kappa \\ 0 \end{pmatrix} = \begin{pmatrix} M & q & \bar{r} \\ -((M+M^T)x_0 + q)^T & x_0^T M x_0 & \bar{z} \\ -\bar{r}^T & -\bar{z} & 0 \end{pmatrix} \begin{pmatrix} x \\ \tau \\ \theta \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \bar{n} \end{pmatrix} \end{array} \tag{2.47}$$

$$x, s, \kappa, \tau \geq 0, \theta \text{ free}$$

where

$$\bar{r} = s_0 - Mx_0 - q,\ \bar{z} = q^T x_0 + \kappa_0 + x_0^T M x_0,$$

and

$$\bar{n} = x_0^T s_0 + \kappa_0.$$

Once again, if $M = \begin{bmatrix} 0 & A \\ -A^T & 0 \end{bmatrix}$, then $M + M^T = 0$ and $x_0^T M x_0 = 0$ (by the skew-symmetry of M), and this second formulation is the same as the previous one, and reduces to the $\mathcal{HSDP}$. We note that this alternative homogeneous formulation is inspired by the linear systems arising in the homogeneous formulation of monotone nonlinear complementarity problems due to Anderson and Ye [AY95].

Based on these problems we define the analogue of the infeasible-point potential function ψ as follows:

$$\begin{aligned} \psi_{LCP}(x, s, \kappa) \quad := \quad & (2n+2)\ln\max\{\|Mx + q - s\|, q^T x + \kappa\} \\ & - \textstyle\sum_{i=1}^n \ln x_i - \sum_{i=1}^n \ln s_i - \ln\kappa. \end{aligned} \tag{2.48}$$

This potential function carries some of the theoretical properties of its linear programming counterpart. For example, it is not hard to show that ψ_{LCP} is unbounded below. In addition, as ψ_{LCP} approaches $-\infty$ either the argument of the first logarithm term in the function is driven to zero indicating that the iterates approach a solution, or the barrier terms increase indefinitely suggesting that the LCP is infeasible.

To extend the potential reduction algorithm of the previous section and to carry out a similar convergence analysis in the case of linear complementarity problems we would like to generalize the results of the previous sections estimating the reduction in the potential function by taking steps in the Kojima-Mizuno-Yoshise and centering directions. In particular, we would like to replicate Lemma 2.3.1 for the case of LCPs. We recall that Lemma 2.3.1 was based on Theorem 2.3.1

due to Ye, Kortanek, Kaliski, and Huang [YKKH93]. Güler and Ye generalize this theorem to the case of LCPs:

Theorem 2.6.1 (Güler and Ye [GY93]) *Let (x, s) be feasible for $\mathcal{LCP}$ and let $(\Delta x, \Delta s)$ satisfy*

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

where α is defined in (2.5). Then $(x^+, s^+) = (x + \Delta x, s + \Delta s)$ is feasible for $\mathcal{LCP}$, and

$$\Phi_\rho(x^+, s^+) - \Phi_\rho(x, s) \le -\sqrt{\frac{\min(XSe)}{x^T s}} \frac{(\rho' - n)(\rho - n) + h_{x,s}}{\sqrt{(\rho' - n)^2 + h_{x,s}}} \beta + \max\left\{\beta^2, \frac{\beta^2}{2(1-\beta)}\right\}, \tag{2.49}$$

where $h_{x,s} := \sum_i \frac{x^T s}{x_i s_i} - n$ and $n \le \rho \le 2n$.

$\square$

Despite this encouraging result, we could not generalize Lemma 2.3.1 to the case of LCPs. A straightforward generalization fails because Proposition 2.2.1(i) does not hold for LCPs, that is, the total uncomplementarity $\bar{x}^T \bar{s} + \bar{\kappa}$ does not go down at the same rate as the infeasibility and optimality measures for the Kojima-Mizuno-Yoshise and centering directions. To remedy this problem we consider a different update on the κ variable so as to make the total uncomplementarity and infeasibility and optimality measures decrease at the same rate. But then, Theorem 2.6.1 needs to be modified, too. We leave this concern as a topic for future research.