Les Cahiers du GERAD ISSN: 0711–2440

A Primal-Dual Interior-Point Algorithm for Linear Programming with Selective Addition of Inequalities

A. Engau
M.F. Anjos

G–2011–44

August 2011

A Primal-Dual Interior-Point Algorithm for Linear Programming with Selective Addition of Inequalities

Alexander Engau

Department of Mathematical and Statistical Sciences
University of Colorado Denver
Denver, CO, United States

aengau@alumni.clemson.edu

Miguel F. Anjos

GERAD & Département de mathématiques et génie industriel
École Polytechnique de Montréal
Montréal (Québec) Canada

anjos@stanfordalumni.org

August 2011

Les Cahiers du GERAD

G–2011–44

Abstract

We present a new primal-dual interior-point algorithm for linear programming problems with equality and inequality constraints. The inequality constraints are initially removed from the problem and selectively added only if they become active during the solution process. After adding an inequality, the algorithm continues with a series of corrector steps to restore centrality and feasibility. We show that the algorithm converges to an optimal solution at which all inequalities are satisfied regardless of whether they are added to the problem or not. We also establish conditions under which the complexity of the algorithm is polynomial in the problem dimension.

Key Words: Linear programming, interior-point algorithms, selective addition of inequalities.

Résumé

Nous présentons un nouvel algorithme primal-dual du point intérieur pour des problèmes de programmation linéaire avec des contraintes d'égalité et d'inégalité. Les contraintes d'inégalité sont d'abord retirées du problème puis ajoutées de manière sélective lorsqu'elles deviennent actives. Après l'ajout d'une inégalité, l'algorithme se poursuit avec une série de pas correcteurs pour restaurer la centralité et l'admissibilité. Nous montrons que l'algorithme converge à une solution optimale qui satisfait toutes les inégalités, indépendamment du fait qu'elles aient été ajoutées ou non au problème. Nous établissons également des conditions sous lesquelles la complexité de l'algorithme est polynomiale en terme de la dimension du problème linéaire.

Mots clés : programmation linéaire, algorithmes du point intérieur, addition sélective d'inégalités.

Acknowledgments: The research of the first author was partially completed while visiting GERAD/École Polytechnique de Montréal in June/July 2011. The support of UC Denver, POLY Montréal, GERAD (Groupe d'études et de recherche en analyze des décisions), and MITACS (Mathematics of Information Technology and Complex Systems) is gratefully acknowledged. The research of the second author was partially supported by MITACS and NSERC, the Natural Sciences and Engineering Research Council of Canada.

1 Introduction

In this paper, we consider linear programming (LP) problems with equality and inequality constraints in the following non-standard primal-dual form:

$$\begin{aligned}
&\min\ c^T x && \max\ b^T y + q^T z && \text{(1a)}\\
&\text{s.t. } Ax = b && \text{s.t. } A^T y + P^T z + s = c && \text{(1b)}\\
&\quad Px \geq q && \quad z \geq 0 && \text{(1c)}\\
&\quad x \geq 0 && \quad s \geq 0. && \text{(1d)}
\end{aligned}$$

Here $c \in \mathbb{R}^n$ is the objective vector for n primal variables, $(A, b) \in \mathbb{R}^{m\times n} \times \mathbb{R}^m$ correspond to m primal equality constraints, and $(P, q) \in \mathbb{R}^{\ell\times n} \times \mathbb{R}^\ell$ correspond to ℓ primal inequalities. Algorithms for LP typically assume that the primal problem is given in standard form in which all constraints are linear equalities and all variables are nonnegative. Several well-known tricks exist to convert any LP to standard form. For example, using nonnegative slack variables r and t to convert the inequalities (1c) into equality constraints, problem (1) can equivalently be written as

$$\begin{aligned}
&\min\ c^T x && \max\ b^T y + q^T z && \text{(2a)}\\
&\text{s.t. } Ax = b && \text{s.t. } A^T y + P^T z + s = c && \text{(2b)}\\
&\quad Px - r = q && \quad z - t = 0 && \text{(2c)}\\
&\quad (x, r) \geq 0 && \quad (s, t) \geq 0 && \text{(2d)}
\end{aligned}$$

which has standard form with

$$\hat{A} = \begin{bmatrix} A & 0 \\ P & -I \end{bmatrix},\ \hat{b} = \begin{bmatrix} b \\ q \end{bmatrix},\ \hat{c} = \begin{bmatrix} c \\ 0 \end{bmatrix},\ \hat{x} = \begin{bmatrix} x \\ r \end{bmatrix},\ \hat{y} = \begin{bmatrix} y \\ z \end{bmatrix},\ \text{ and } \hat{s} = \begin{bmatrix} s \\ t \end{bmatrix}. \tag{3}$$

To solve problem (1), the reformulation (2) is reasonable if the number of inequalities is not much larger than the number of primal variables. Specifically, if $\ell = \mathcal{O}(n)$, then $n + \ell = \mathcal{O}(n)$ as well and the additional slack variables do not impact the asymptotic complexity of solving problem (1) or (2). If the number of inequalities is much larger however, this approach is generally inefficient especially if only few inequalities are active at optimality. Typical examples for this situation are continuous relaxations of combinatorial or discrete optimization problems where the inequalities are large classes of cutting planes. In this scenario, it is often true that a large majority of inequalities is unnecessary to compute a feasible optimal solution, although it is usually not known a priori which inequalities are needed and which ones are not..

Several algorithms have been proposed to deal with this situation, including column-generation or cutting-plane methods [15] and augmented-Lagrangian or other dual approaches [3, 6]. All of these techniques have in common that they repeatedly solve and update successive relaxations of an original problem with many constraints until an optimal solution is found that is also feasible for the original problem. The different relaxation and update schemes of these methods are well known: they have been studied extensively in theory and successfully used for many applications.

As an alternative to solving repeated relaxations, fewer algorithms have been implemented that solve linear and more recently semidefinite programs by dynamically adding (and possibly also removing) inequalities as an integral part of the solution process [1, 2, 5, 7, 13, 14, among others]. Unlike conventional cutting-plane methods, these algorithms do not necessarily generate violated inequalities at infeasible optimal solutions but may predict such inequalities at feasible, intermediate iterates. This gives the opportunity to add selected constraints before they are actually violated and may allow to resume the algorithm from the current and still feasible iterate.

While these approaches have shown some promise in implementation, to the best of our knowledge there are no theoretical convergence or worst-case complexity proofs in the literature for methods with selective addition of inequalities. Hence, our objective in this paper is to formally state such an algorithm and to provide such a proof for it. Specifically, our algorithm for problem (1) is based on a feasible primal-dual path-following short-step predictor-corrector interior-point method (IPM) and thus uses one of the basic LP frameworks with established polynomial-time convergence.

We should mention that without additional assumptions on the inequalities that are added, however, it is not realistic to expect better or even equal worst-case complexity than that of a standard method. In particular, some extra work is inevitable to choose and properly integrate selected inequalities, and of course all inequalities may be necessary in the worst case. Hence, the main contribution in this paper is the new algorithmic formulation and its theoretical analysis; the actual performance of this algorithm will likely be of practical interest only if the number of inequalities to be added is reasonably small. In particular, our analysis provides clear insight into the conditions under which the algorithm is polynomial.

This paper is structured as follows. Section 2 reviews preliminaries including our basic notation and terminology. In Section 2.1, we collect some known results and derive several generalizations for the different predictor and corrector steps that are used in our method. Section 2.2 describes a standard feasible predictor-corrector IPM and outlines its known complexity proof. Our new algorithm and its significantly more intricate analysis are presented in Section 3. To facilitate some of our discussion and motivate the key steps of the new algorithm, we give a brief general outline of the algorithm and a short example in the remaining part of this introduction. Some additional remarks and ideas for possible further work are given in Section 4 which also concludes the paper.

1.1 Outline of Algorithm and Main Result

Like a standard primal-dual path-following interior-point algorithm, we compute intermediate iterates along a trajectory of approximate solutions that are characterized by four quantities: a centrality measure γ of its proximity to the central path that we set to $1/4$, a barrier parameter μ indicating its proximity to optimality, and primal and dual residuals ξ_b and ξ_c measuring infeasibility. The following are our basic assumptions.

Assumption 1 *Problem (1) has an optimal solution (x^*, y^*, z^*, s^*).*

The primary reason for this assumption is to keep the paper shorter in length and focused on our objective to demonstrate convergence to an optimal solution, if it exists. To extend the algorithm and its analysis to also address the more practical aspects of detecting unboundedness or infeasibility, we can adopt similar indicators and termination criteria to those used by similar IPMs [10, 17].

Assumption 2 *Problem (1) has a feasible solution (x, y, z, s) that satisfies $(x, s) > 0$, $z = 0$, $\|Xs - \mu e\| \leq (1/4)\mu$ for $\mu = x^T s/n$, and $Px > q$, or equivalently, $Px - q \geq (1/\tau)\mu$ for some $\tau > 0$.*

Several papers describe how to modify a primal-dual LP such that an initial (strictly) feasible pair can be found [11, 12, 19, 20, among others]. However, we do not assume that problem (1) has a strictly feasible solution because $z = 0$. In particular, if it is strictly feasible and there was some entry of z that was strictly positive for all feasible points, then the corresponding inequality of $Px > q$ must be active at optimality and could immediately be handled as an equality constraint. For all other inequalities that satisfy $Px > q$ with x and s given, the inequality $Px - q \geq (1/\tau)x^T s/n$ can always be satisfied as long as $\tau > 0$ is chosen sufficiently large.

Assumption 3 *There exists a sufficiently large number $M < \infty$ such that the primal residual $r = Px - q$ is bounded from above by M.*

Without loss of generality, we can enforce this bound by adding the additional inequalities $Px \leq q + Me$ for any $M < \infty$ such that $Px^* \leq q + Me$. This doubles the number of primal inequalities but has no impact on the asymptotic terms $\mathcal{O}(\ell)$ for the numbers of inequalities in (1) and $\mathcal{O}(n + \ell)$ for the number of primal variables in (2).

For clarity, we now refer to the full problem (1) as the *original problem*, and we call each reduced problem in which some or all of the inequalities are removed an *instance* of the original problem. Analogous to the hat-notation in (3), we consider each instance in primal-dual standard form with data $(\hat{A}, \hat{b}, \hat{c}) \in \mathbb{R}^{(m+l)\times(n+l)} \times \mathbb{R}^{m+l} \times \mathbb{R}^{n+l}$, where $l \leq \ell$ is the number of added inequalities that are written as slacked equalities. Similarly, we denote the system of inequalities that remain dropped by $(\hat{P}, \hat{q}) \in \mathbb{R}^{(\ell-l)\times(n+l)} \times \mathbb{R}^{\ell-l}$, where the l extra columns of $\hat{P}$ are all zero and correspond to the slack variables of those inequalities that have been added. Finally, we denote the variable vectors for each instance by $(\hat{x}, \hat{y}, \hat{s})$, or $(\hat{x}_r, \hat{y}_z, \hat{s}_t)$ if it is useful to highlight the augmented entries (r, z, t) when adding new inequalities. For convenience, we also occasionally drop the hat-notation when discussing a specific instance after saying so; otherwise (A, b, c, P, q) is the data of the original problem that we associate with the initial instance without any inequalities.

The algorithm now starts from this initial instance with a feasible iterate that satisfies the conditions in Assumption 3 so that all inequalities at this point are inactive with a residual of at least $\rho = \mu/\tau$. Like a standard IPM, we then continue to alternate between predictor steps that reduce the barrier parameter μ, and corrector steps that recenter the iterates and keep them in sufficiently close proximity to the central path. Unlike standard IPMs, however, whenever taking a primal step we also check the removed inequalities and select new constraints for addition if their corresponding residuals fall below the threshold value $\rho = \mu/\tau$, indicating that they tend to become active at optimality. The dependency of this threshold on the barrier parameter μ makes it an adaptive threshold that decreases in proportion to the barrier parameter so that only active inequalities would be added at optimality. When a new inequality is added, the algorithm augments both problem and iterate in such a way that it maintains centrality, the barrier parameter, and primal feasibility. It is well known, however, that it is generally not possible to also preserve feasibility in the dual so that we continue with a series of three additional corrector steps that work together to fully absorb the new dual infeasibility:

1. The first corrector step is a pure centering step and restores the iterate's proximity to the central path; it does not change the barrier parameter nor primal or dual residuals.
2. The second corrector step is a feasibility-restoring step that reduces dual infeasibility; it does not modify the primal iterate but generally changes both centrality and the barrier parameter.
3. The third corrector is a modified centering step that does not change centrality or residuals, but restores the barrier parameter after it has been changed in the second correcter step.

Whenever we change the primal iterate in the first and third corrector steps, we also check the residuals of dropped inequalities and, if necessary, add a new inequality before all dual infeasibilities caused by earlier added inequalities have been fully absorbed. In this case, the algorithm absorbs all infeasibility inequality-wise starting from the most recent inequality, in a last-in-first-feasible fashion. This recursive nature of the algorithm leads to a possibly exponential worst-case complexity, which is proved in detail in Section 3.

Theorem 1 *Let problem (1) be given, and (x, y, s) be a strictly feasible point that satisfies $x^T s \le (1/\epsilon)^\kappa$ with $\kappa > 0$ and Assumption 1 with $\tau > 0$. If the problem satisfies Assumptions 2 and 3, then the new algorithm finds an ϵ-optimal solution in $\mathcal{O}(((\kappa + \tau + 1)/\epsilon) l (n + l)^{3/2} e^{\theta/11})$ iterations, where $\theta = \mathcal{O}(l/\sqrt{n + l})$ and l is the number of inequalities that are added to the problem.*

In particular, if $l = \mathcal{O}(\sqrt{n})$, then $\theta = \mathcal{O}(1)$ and $e^{\theta/11} = \mathcal{O}(1)$ so that the iteration bound reduces to $\mathcal{O}(((\kappa + \tau + 1)/\epsilon) l (n + l)^{3/2})$ and the complexity is polynomial. Based on better insight and additional parameters defined in our algorithm, we will establish much weaker conditions for polynomiality in Theorem 4 and Section 3.2 below.

1.2 An Example

We begin by illustrating our method on an example. To keep this discussion relatively short and easier to follow, we intentionally use the trivial problem

$$\min x \text{ s.t. } x \ge 1 \text{ and } x \ge 0. \tag{4}$$

For this small example, it is easy to see that the inequality constraint $x \ge 1$ is active at the optimal solution $x^* = 1$ while the nonnegativity constraint $x \ge 0$ is redundant. In particular, this problem is simple enough to be quickly solved without any centering corrector steps so that we can focus primarily on the augmentation mechanism to add the necessary inequality. Hence, let us first drop this constraint and write the initial instance of problem (4) in primal-dual standard form

$$\min x \text{ s.t. } x \geq 0, \qquad \max 0 \text{ s.t. } s = 1 \text{ and } s \geq 0. \tag{5}$$

Despite its simplicity, note that this problem is a well-defined LP with strict relative interiors $x > 0$ and $s = 1 > 0$, and has standard form with an empty matrix $A \in \mathbb{R}^{0\times 1}$ so that the dual variable $y \in \mathbb{R}^m$ does not appear in the problem. Hence, we can start from a strictly feasible initial point $(x^1, s^1) = (x_0, 1)$ at which the dropped inequality is satisfied, say $x_0 = 4$ and $\mu^1 = x^1 s^1 = x_0 = 4$ so that $x_0 - 1 = 3 \geq (1/\tau)4$ for any $\tau \geq 4/3$. Let us choose $\tau = 2$. Independently of τ, the Newton direction at our initial point is $(\Delta x, \Delta s) = (-x_0, 0) = -(4, 0)$ which targets the optimal point $(0, 1)$ of (5) in a single step. Because a full step into this direction violates the dropped inequality $x \geq 1$, we will reduce the step length and add the primal constraint $x - r = 1$ with a strictly feasible primal slack $r = x - 1 > 0$. At the same time, we also modify the dual problem to maintain proper primal-dual (standard) form:

$$\begin{aligned} &\min x && \max z \\ &\text{s.t. } x - r = 1 && \text{s.t. } z + s = 1, \ -z + t = 0 \\ &\quad (x, r) \geq 0 && \quad (s, t) \geq 0. \end{aligned}$$

Note that $\hat{A} = [1, -1]$, $\hat{b} = 1$, $\hat{c} = [1, 0]^T$, $\hat{x}_r = [x, r]^T$, $\hat{y}_z = z$, and $\hat{s}_t = [s, t]^T$ using the hat-notation of (3).

Next, we explain the computation of the reduced step length α and the initialization of r and t to stay primal feasible and preserve the barrier parameter $(1 - \alpha)\mu = (1 - \alpha)x_0$ at $(x(\alpha), s(\alpha)) = (x^1, x^1) + \alpha(\Delta x, \Delta s) = ((1 - \alpha)x_0, 1)$ after adding the inequality. For the latter, it follows that we need to set r and t such that

$$rt = \mu(\alpha) = x(\alpha)^T s(\alpha)/1 = (1 - \alpha)x_0 = 4(1 - \alpha)$$

because then $\hat{\mu} = (\hat{x}_r)^T \hat{s}_t/2 = (x(\alpha)s(\alpha) + rt)/2 = \mu$. Hence, by choice of $\tau = 2$, we find $r = \mu(\alpha)/\tau = (1 - \alpha)x_0/\tau = 2(1 - \alpha)$ and then solve

$$x(\alpha) - r = (1 - \alpha)x_0 \left(1 - \frac{1}{\tau}\right) = 2(1 - \alpha) = 1$$

for $\alpha = 1 - \tau/((\tau - 1)x_0) = 1/2$ yielding the new primal iterate

$$\begin{aligned} x &= x_0 + \alpha\Delta x = (1 - \alpha)x_0 = \frac{\tau}{\tau - 1} = 2 \\ r &= x - 1 = \frac{\tau}{\tau - 1} - 1 = \frac{1}{\tau - 1} = 1. \end{aligned}$$

Note that this iterate is primal feasible and satisfies $xs = rt = \tau/(\tau - 1) = 2$ so that the augmented iterate is perfectly centered with the same barrier parameter $\hat{\mu} = (xs + rt)/2 = \tau/(\tau - 1) = 2$ as before. Similarly, for the dual problem we set $z = 0$ to maintain feasibility of $s = 1$ for the constraint $s + z = 1$. However, note that the resulting iterate $(z, s, t) = (0, 1, 2)$ is not feasible for the new dual constraint $-z + t = 0$ but has the negative residual $\zeta = 0 - (-z + t) = -\tau = -2$. Hence, we now continue with a dual feasibility step into the dual direction computed from the augmented system

$$\begin{bmatrix} \hat{A} & 0 & 0 \\ 0 & \hat{A}^T & I \\ \hat{S}_t & 0 & \hat{X}_r \end{bmatrix} \begin{bmatrix} \Delta\hat{x}_r \\ \Delta\hat{y}_z \\ \Delta\hat{s}_t \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ s & 0 & 0 & x & 0 \\ 0 & t & 0 & 0 & r \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta r \\ \Delta z \\ \Delta s \\ \Delta t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \zeta \\ 0 \\ 0 \end{bmatrix}$$

where we set $(x, r) = (2, 1)$, $(s, t) = (1, 2)$, and $\zeta = -2$. The solution

$$(\Delta x, \Delta r, \Delta z, \Delta s, \Delta t) = (4/5, 4/5, 2/5, -2/5, -8/5)$$

is used to only update the dual iterate

$$(z, s, t) + (\Delta z, \Delta s, \Delta t) = (0, 1, 2) + (2/5, -2/5, -8/5) = (2/5, 3/5, 2/5)$$

which is now feasible and in this case has reduced the barrier parameter to $\mu = (xs+rt)/2 = (6/5+2/5)/2 = 4/5$. The search direction from the next Newton system

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 3/5 & 0 & 0 & 2 & 0 \\ 0 & 2/5 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta r \\ \Delta z \\ \Delta s \\ \Delta t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -6/5 \\ -2/5 \end{bmatrix}$$

is $(\Delta x, \Delta r, \Delta z, \Delta s, \Delta t) = (-10/7, -10/7, 6/35, -6/35, 6/35)$ and yields the optimal solution after a primal step with $\alpha_p = 7/10$ and a dual step with $\alpha_d = 7/2$:

$$(x, r) = (2, 1) + (7/10)(-10/7, -10/7) = (1, 0)$$
$$(z, s, t) = (2/5, 3/5, 2/5) + (7/2)(6/35, -6/35, 6/35) = (1, 0, 1).$$

Following a generic IPM, the algorithm may also take the same step size in primal and dual space, and restrict the step size to at most 1 especially if we intend to stay in proximity to the central path. In that case, the algorithm would continue analogously with several shorter steps into the primal and dual directions $(\Delta x, \Delta r) = (-1, -1)$ and $(\Delta z, \Delta s, \Delta t) = (1, -1, 1)$ until terminating within some sufficiently small threshold of the optimal solution.

2 Preliminaries and New Results

In this section, we first collect relevant preliminaries of primal-dual interior-point algorithms and then state several known and some new results. Consider an LP in standard form with n primal variables:

$$\min c^T x \qquad\qquad \max b^T y \tag{6a}$$
$$\text{s.t. } Ax = b \qquad\qquad \text{s.t. } A^T y + s = c \tag{6b}$$
$$x \geq 0, \qquad\qquad s \geq 0. \tag{6c}$$

We denote the set of (strictly) feasible primal-dual solutions (x, s) for problem (6) by

$$\mathcal{S}^n := \left\{(x, s) \in \mathbb{R}^n \times \mathbb{R}^n : (x, s) \geq 0,\ Ax = b,\ A^T y + s = c \text{ for some } y\right\}$$
$$\mathcal{S}_0^n := \{(x, s) \in S : (x, s) > 0\}.$$

We frequently drop the superscript n if the dimension of x and s is clear from the context, but find this more explicit notation helpful to improve clarity in the following section. The associated (primal) logarithmic barrier formulation for problem (6) is

$$\min c^T x - \mu \sum_{i=1}^{n} \log(x_i) \text{ s.t. } Ax = b \tag{7}$$

where $\mu > 0$ is the barrier parameter. The first-order optimality conditions of problem (7) are given by the following system of nonlinear equations:

$$Ax = b \tag{8a}$$
$$A^T y + s = c \tag{8b}$$
$$XSe = \mu e. \tag{8c}$$

We refer to the conditions in (8) as primal feasibility, dual feasibility, and complementarity, respectively. In the complementarity condition, we use X and S as the usual notation for diagonal square matrices built of the elements of x and s, and whenever convenient we write I and $e := (1, 1, \ldots, 1)^T$ for the identity matrix and the vector of all ones of suitable dimension, respectively.

Under the assumption that the set $\mathcal{S}_0^n$ of strictly feasible solutions is nonempty, system (8) has a unique solution $(x(\mu), y(\mu), s(\mu))$ for every $\mu > 0$, and we define the set of primal-dual solutions $(x(\mu), s(\mu))$ as the central path

$$\mathcal{C}^n := \{(x, s) \in S_0^n : Xs = \mu e \text{ for some } \mu > 0\}.$$

Starting from a (strictly) feasible initial point, primal-dual path-following methods in each iteration compute a Newton direction from the nonlinear system (8) for a decreasing sequence of values for the barrier parameter μ. An additional step size condition guarantees that all new iterates belong to some suitable neighborhood of the central path and converge to an optimal solution of problem (6) as μ is reduced to zero. Specifically, we will use the short-step neighborhood

$$\mathcal{N}_2^n(\gamma) := \left\{(x, s) > 0 : \|Xs - \mu e\|_2 \le \gamma\mu \text{ where } \mu = x^T s/n\right\}.$$

As before, we drop the superscript n if the problem dimension is clear from the context, and throughout this paper we will use $\|\cdot\| := \|\cdot\|_2$ without subscript to denote the canonical 2-norms for both vectors or matrices. Note that, unlike definitions in other papers, here $\mathcal{N}_2^n(\gamma)$ only depends on the problem dimension n and the centrality parameter γ and thus is independent of the actual central path or the set of feasible solutions. This is convenient to clearly distinguish between centrality and feasibility which significantly facilitates our notation later. While inherently infeasible IPMs can be designed to move along infeasible iterates and establish feasibility as part of the algorithm, we have chosen to work with a class of feasible IPMs that in every iteration require a (strictly) feasible iterate that belongs to $\mathcal{N}_2^n(\gamma) \cap \mathcal{S}_0^n$.

Both feasible and infeasible IPMs can solve LP problems in polynomial time. Specifically, given an initial well-centered point with $x^T s = (1/\epsilon)^\kappa$, the feasible short-step method finds an ϵ-optimal solution for problem (6) in at most $\mathcal{O}((\kappa+1)\sqrt{n}\log(1/\epsilon))$ iterations [18]. In comparison, the classical long-step method needs at most $\mathcal{O}((\kappa+1)n\log(1/\epsilon))$ iterations whereas the best infeasible method has a worst-case complexity of $\mathcal{O}(nL)$ [17] where

$$L = \log\left(\max\left\{x^T s, \|Ax - b\|, \left\|A^T y + s - c\right\|\right\}/\epsilon\right).$$

Note that for solving problem (1) in standard form (2), this corresponds to an overall best complexity bound of $\mathcal{O}((\kappa+1)\sqrt{n+\ell}\log(1/\epsilon))$ for a feasible short-step method.

2.1 Predictor and Corrector Steps

Let (x, y, s) be a feasible primal-dual iterate with $Ax = b$ and $A^T y + s = c$. Set $\beta \ge 0$, $\mu = x^T s/n$, and define

$$\xi_{\beta\mu} := \beta\mu e - Xs = \beta(x^T s/n)e - Xs.$$

For $\beta = 0$ or $\beta = 1$ respectively, the standard feasible predictor and corrector directions are the Newton directions of the nonlinear system (8) that can be computed from the system of linear equations

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \xi_{\beta\mu} \end{bmatrix}. \tag{9}$$

Note that this system has the closed-form solutions

$$\Delta y = -EAS^{-1}\xi_{\beta\mu} \tag{10a}$$

$$\Delta s = -A^T \Delta y = A^T EAS^{-1}\xi_{\beta\mu} \tag{10b}$$

$$\Delta x = S^{-1}\xi_{\beta\mu} - D\Delta s = \left(I - DA^T EA\right) S^{-1}\xi_{\beta\mu} \tag{10c}$$

where we write $D := S^{-1}X = XS^{-1}$ and $E := \left(ADA^T\right)^{-1}$. From (9), we know that

$$x_i \Delta s_i + \Delta x_i s_i = \beta x^T s/n - x_i s_i$$
$$x^T \Delta s + \Delta x^T s = -(1-\beta) x^T s$$
$$\Delta x^T \Delta s = -\Delta x^T A^T \Delta y = 0.$$

Furthermore, given a step size α and writing

$$(x(\alpha), y(\alpha), s(\alpha)) := (x, y, s) + \alpha(\Delta x, \Delta y, \Delta s)$$

it follows that $Ax(\alpha) = Ax = b$ and $A^T y(\alpha) + s(\alpha) = A^T y + s = c$ for all α because $A\Delta x = 0$ and $A^T \Delta y + \Delta s = 0$ for all directions computed from system (9), and

$$x(\alpha)^T s(\alpha) = (1 - \alpha(1-\beta)) x^T s \tag{11a}$$
$$X(\alpha) s(\alpha) = (1-\alpha) X s + \alpha\beta(x^T s/n) e + \alpha^2 \Delta X \Delta s. \tag{11b}$$

In particular, equation (11a) implies that

$$\mu(\alpha) := x(\alpha)^T s(\alpha)/n = (1-\alpha(1-\beta))\mu = \begin{cases} (1-\alpha)\mu & \text{if } \beta = 0; \\ \mu & \text{if } \beta = 1. \end{cases} \tag{12}$$

Moreover, for $\alpha \in [0,1]$ and general $\beta \geq 0$, we also have that $\mu(\alpha) = \beta\mu$ if $\alpha = 1$ so that

$$\left| 1 - \frac{\mu(1)}{\mu(\alpha)} \right| = \left| 1 - \frac{\beta}{1-\alpha(1-\beta)} \right| = \left| \frac{(1-\alpha)(1-\beta)}{1-\alpha+\alpha\beta} \right| \leq |1-\beta|. \tag{13}$$

The following result by Mizuno et al. [18] provides bounds on the term $\Delta X \Delta s$ in equation (11b). Note that the use of β and γ in [18] is exactly opposite to their use in this paper.

Lemma 2.1 (Lemma 1(a)/2 in [18]) *Let $(x,s) > 0$, $\mu = x^T s/n$, and $(\Delta x, \Delta s)$ be obtained from system (9). Then*

$$\|\Delta X \Delta s\| \leq (\sqrt{2}/4) \left\| (XS)^{-1/2} \xi_{\beta\mu} \right\|^2 ..$$

(a) *If $\beta = 0$, then $\left\| (XS)^{-1/2} X s \right\|^2 = x^T s = n\mu$.*

(b) *If $\beta = 1$, $\gamma \in (0,1)$ and $(x,s) \in \mathcal{N}_2(\gamma)$, then $\left\| (XS)^{-1/2}(\mu e - Xs) \right\|^2 \leq \left(\gamma^2/(1-\gamma)\right)\mu$.*

(c) *If $\beta \in (0,1)$, $\gamma \in (0,1)$, $\beta \leq 2(1-\gamma)$, and $(x,s) \in \mathcal{N}_2(\gamma)$, then*

$$\left\| (XS)^{-1/2}(\beta\mu e - Xs) \right\|^2 \leq n\mu.$$

The original statement in [18] proves statement (c) in Lemma 2.1 for an iterate (x,s) that belongs to a wider neighborhood than $\mathcal{N}_2(\gamma)$. Because $\mathcal{N}_2(\gamma)$ is contained in this neighborhood, the bound is still valid although not necessarily very accurate. We prove a new, modified version of part (c) that generalizes part (b) and is used in our new results in Section 2.1.4.

Lemma 2.2 *Let $\gamma \in (0,1)$, $(x,s) \in \mathcal{N}_2(\gamma)$, $\mu = x^T s/n$, and $(\Delta x, \Delta s)$ be obtained from system (9). If $|1-\beta| \leq \delta/\sqrt{n}$, then $\left\| (XS)^{-1/2}(\beta\mu e - Xs) \right\|^2 \leq \left((\gamma+\delta)^2/(1-\gamma)\right)\mu$.*

Proof. From $(x,s) \in \mathcal{N}_2(\gamma)$ we know that $\|Xs - \mu e\| \leq \gamma\mu$ so that $(1-\gamma)\mu \leq x_i s_i \leq (1+\gamma)\mu$ for all $i = 1, 2, \ldots, n$. This implies that

$$\left\| (XS)^{-1/2} \right\|^2 = \max\{1/(x_i s_i) : i = 1, 2, \ldots, n\} \leq 1/((1-\gamma)\mu)$$

and yields the desired result as follows:

$$\begin{aligned}
&\left\|(XS)^{-1/2}(\beta\mu e - XSe)\right\|^2 \\
\leq\ & \left\|(XS)^{-1/2}\right\|^2 \cdot \|\beta\mu e - XSe\|^2 \\
\leq\ & \frac{1}{(1-\gamma)\mu} \cdot (\|\mu e - XSe\| + \|(1-\beta)\mu e\|)^2 \\
\leq\ & \frac{1}{(1-\gamma)\mu} \cdot \left(\gamma\mu + |1-\beta|\mu\sqrt{n}\right)^2 \leq\ (\gamma+\delta)^2 \,\frac{\mu}{1-\gamma}.
\end{aligned}$$

□

Note that if $\delta = 0$, then this result reduces to statement (b) in Lemma 2.1.

2.1.1 The Predictor Step: Reducing the Barrier Parameter

If we set $\beta < 1$, then we call the direction obtained from system (9) a predictor direction. In particular, if we set $\beta = 0$ so that $\xi_{\beta\mu} = -Xs$, then we obtain the affine-scaling direction, and a full step into this direction reduces μ to 0 according to (12). However, especially for small values of β, it is usually not possible to take a full step and at the same time maintain centrality and nonnegativity of x and s. The following result gives a lower bound on the guaranteed minimum step size into the affine-scaling direction, and is well known and widely used in the IPM literature.

Lemma 2.3 (Lemma 4 in [18]) *Let $\beta = 0$, $\gamma = 1/4$, $(x, s) \in \mathcal{N}_2(\gamma)$, and $(\Delta x, \Delta y, \Delta s)$ be the solution of system (9) with $\xi_{\beta\mu} = -Xs$. Let $\bar{\alpha}$ be the largest step size such that $(x(\alpha), s(\alpha)) \in \mathcal{N}_2(2\gamma)$ for every $\alpha \in [0, \bar{\alpha}]$. Then*

$$\bar{\alpha} \geq \min\left\{0.5,\ 8^{-1/4}n^{-1/2}\right\}.$$

After the step, we see from (12) that the barrier parameter $\mu(\bar{\alpha})$ at the new iterate $(x(\bar{\alpha}), y(\bar{\alpha}), s(\bar{\alpha}))$ is reduced proportionally to the step length so that $x(\bar{\alpha})^T s(\bar{\alpha}) = (1-\bar{\alpha})x^T s$. Moreover, Lemma 2.3 implies that the new iterate still belongs to the wider neighborhood $\mathcal{N}_2(2\gamma)$, and a centering corrector step can be taken to restore centrality of the iterate in the original neighborhood $\mathcal{N}_2(\gamma)$.

2.1.2 The First Corrector Step: Restoring Centrality

If we set $\beta = 1$ so that $\xi_{\beta\mu} = \mu e - Xs$, then we call the direction obtained from system (9) the pure centering direction. It is pure in the sense that a step into this direction restores only centrality but does not change the barrier parameter, because $x(\alpha)^T s(\alpha) = x^T s$ for all α from (11a).

Lemma 2.4 (Lemma 3 in [18] / Lemma 4.2 in [17]) *Let $\beta = 1$, $\gamma = 1/4$, $(x, s) \in \mathcal{N}_2(2\gamma)$ with $\mu = x^T s/n$, and $(\Delta x, \Delta y, \Delta s)$ be the solution of system (9) with $\xi_{\beta\mu} = \mu e - Xs$. Then the full Newton step is feasible and the new point $(\bar{x}, \bar{y}, \bar{s}) = (x, y, s) + (\Delta x, \Delta y, \Delta s)$ satisfies $(\bar{x}, \bar{s}) \in \mathcal{N}_2(\gamma)$ with barrier parameter $\bar{\mu} = \bar{x}^T \bar{s}/n = \mu$.*

Hence, given an iterate in the wider neighborhood $\mathcal{N}_2(2\gamma)$, it suffices to take a single full step into the affine-scaling direction to recenter the iterate and restore its centrality in the narrower neighborhood $\mathcal{N}_2(\gamma)$ with no change in its barrier parameter.

2.1.3 The Second Corrector Step: Restoring Dual Feasibility

Let us now consider a situation where we have an infeasible point (x, y, s).. The following analysis can be conducted for both primal and dual infeasibilities, but because our method is always primal feasible we omit the discussion of how to restore primal feasibility. Again, let $\mu = x^T s/n$ and define the dual residual

$$\xi_c := c - A^T y - s.$$

We call the direction $(\Delta x, \Delta y, \Delta s)$ obtained from the system

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ \xi_c \\ 0 \end{bmatrix} \tag{14}$$

a feasibility-restoring direction. Note that this system has the closed-form solutions

$$\begin{aligned} \Delta y &= EAD\xi_c \\ \Delta s &= \xi_c - A^T\Delta y = \left(I - A^T EAD\right)\xi_c \\ \Delta x &= -D\Delta s = \left(DA^T EA - I\right) D\xi_c \end{aligned}$$

where we use D and E as defined in (10). Given a step size α, from (14) we know that

$$\xi_c(\alpha) := c - A^T y(\alpha) - s(\alpha) = (1-\alpha)\xi_c \tag{15}$$

so that a full dual step with step size $\alpha = 1$ fully restores dual feasibility. However, it is not guaranteed that a full step maintains centrality and nonnegativity unless the infeasibility is sufficiently small. The following lemma is important for our analysis and a slightly extended version of the original result in [4] that explicitly includes the lower and upper bounds on the new barrier parameter in (a) and upper bounds on the centrality norm in (b) that are derived within the original proof. Note that the parameters θ and β in the original statement are replaced by γ and δ in this paper.

Lemma 2.5 (Lemma 3.3 in [4]) *Let $(x, s) \in \mathcal{N}_2(\gamma)$, $\mu = x^T s/n$, $\delta < \sqrt{n}$, and $(\Delta y, \Delta s)$ be the dual direction obtained from system (14).. If*

$$\left\|S^{-1}\xi_c\right\|_2 \le \frac{\delta}{\sqrt{n}} \cdot \left(\frac{1+\gamma}{1-\gamma}\right)^{1/2}$$

then the full Newton step in the dual space is feasible, and the new point $(\bar{x}, \bar{y}, \bar{s}) = (x, y+\Delta y, s+\Delta s)$ absorbs the total infeasibility ξ_c. Furthermore, the new barrier parameter $\bar{\mu} = \bar{x}^T\bar{s}/n$ satisfies

(a) $(1 - \delta/\sqrt{n})\mu \le \bar{\mu} \le (1 + \delta/\sqrt{n})\mu$;

(b) $\left\|\bar{X}\bar{s} - \bar{\mu}e\right\| \le ((1+\gamma)\delta + \gamma + \delta)\mu$.

The original lemma also states that $(\bar{x}, \bar{s}) \in \mathcal{N}_2(2\gamma)$ for $\gamma = 1/4$, $\delta = 1/10$, and $\sqrt{n} \ge 100$. Similar to this observation, from (a) we first notice that $\mu \le \bar{\mu}/(1 - \delta/\sqrt{n}) \le \bar{\mu}/(1-\delta)$ so that

$$\left\|\bar{X}\bar{s} - \bar{\mu}e\right\| \le \left(((1+\gamma)\delta + \gamma + \delta)/(1-\delta)\right)\bar{\mu}$$

from (b). In particular, because $((1+\gamma)\delta + \gamma + \delta)/(1-\delta) = 1/2$ for $\gamma = 1/4$ and $\delta = 1/11$, this implies that $(\bar{x}, \bar{s}) \in \mathcal{N}_2(2\gamma)$ independent of n for $\gamma = 1/4$ and $\delta \le 1/11$. These two observations form the basis of our new Lemma 2.6.

Lemma 2.6 (Corollary to Lemma 2.5) *Let $\gamma = 1/4$, $(x, s) \in \mathcal{N}_2(\gamma)$, $\mu = x^T s/n$, $\delta \le 1/11$, $\lambda = (\delta/\sqrt{n})((1+\gamma)/(1-\gamma))^{1/2}$, $\sigma = \lambda/\left\|S^{-1}\xi_c\right\|$ and $(\Delta y, \Delta s)$ be the dual direction obtained from system (14). Let $\bar{\alpha} \le 1$ be the largest step size such that*

$$(x, s(\alpha)) \in \mathcal{N}_2(2\gamma) \tag{16a}$$

$$\left(1 - \frac{\delta}{\sqrt{n}}\right) x^T s \le x^T s(\alpha) \le \left(1 + \frac{\delta}{\sqrt{n}}\right) x^T s \tag{16b}$$

hold for all $\alpha \in [0, \bar{\alpha}]$, and denote $(\bar{x}, \bar{y}, \bar{s}) = (x, y(\bar{\alpha}), s(\bar{\alpha}))$ and $\bar{\xi}_c = c - A^T\bar{y} - \bar{s}$.

(a) *If $\sigma \ge 1$, then $\bar{\alpha} = 1$ and $\bar{\xi}_c = 0$.*

(b) *If $\sigma \le 1$, then $\bar{\alpha} \ge \sigma$ and $\left\|\bar{\xi}_c\right\| \le (1-\sigma)\|\xi_c\|$.*

Proof. If $\sigma \geq 1$, or equivalently, if $\left\|S^{-1}\xi_c\right\| \leq \lambda$, then the above observation and Lemma 2.5 imply that the full dual step $(\Delta y, \Delta s)$ satisfies the conditions in (16) and absorbs the total dual infeasibility ξ_c, so $\bar{\alpha} = 1$ and $\bar{\xi}_c = 0$. If $\sigma \leq 1$, then let $(\Delta\tilde{y}, \Delta\tilde{s}) = \sigma(\Delta y, \Delta s)$ correspond to the dual solution of system (14) with a scaled right-hand side $\tilde{\xi}_c = \sigma\xi_c$. Because $\left\|S^{-1}\tilde{\xi}_c\right\| = \lambda$, by the same arguments as above this shortened step satisfies the conditions in (16) and absorbs the partial infeasibility $\tilde{\xi}_c$, so $\bar{\alpha} \geq \sigma$ and $\left\|\bar{\xi}_c\right\| = \|\xi_c(\bar{\alpha})\| = (1-\bar{\alpha})\|\xi_c\| \leq (1-\sigma)\|\xi_c\|$ from (15). □

We highlight that the primal iterate in Lemmata 2.5 and 2.6 does not change, which is important later because it implies that no new primal inequalities need to be added when restoring dual feasibility in the second corrector step. Furthermore, because the dual iterate may only change within the neighborhood $\mathcal{N}_2(2\gamma)$, it again suffices to eventually take a single full first corrector step to restore centrality in $\mathcal{N}_2(\gamma)$. First, however, we introduce a new third corrector step that restores the barrier parameter after it has changed from μ to $\bar{\mu}$.

2.1.4 The Third Corrector Step: Restoring the Barrier Parameter

Similar to the first corrector step, the third corrector step is a modified centering step that does not restore centrality but adjusts the barrier parameter from μ to $\beta\mu$ for some value of β that is close to 1, and that generally can be either smaller or larger. The next lemma gives a new, general result that is based on our Lemma 2.2.

Lemma 2.7 *Let $\gamma = 1/4$, $\delta \leq 1/5$, $(x,s) \in \mathcal{N}_2(2\gamma)$, $\mu = x^Ts/n$, and $(\Delta x, \Delta s)$ be the direction from system (9) with $\xi_{\beta\mu} = \beta\mu e - Xs$. If $|1-\beta| \leq \delta/\sqrt{n}$, then the full Newton step is feasible and $(x(\alpha), s(\alpha)) \in \mathcal{N}_2(2\gamma)$ for all $\alpha \in [0,1]$.*

Proof. From (11), Lemma 2.1, Lemma 2.2, (12), and the stated assumptions we see that

$$\begin{aligned}
&\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| \\
&\leq (1-\alpha)\|Xs - \mu e\| + \alpha^2\|\Delta X \Delta s\| \\
&\leq (1-\alpha)\|Xs - \mu e\| + \alpha^2(\sqrt{2}/4)\left\|(XS)^{-1/2}\xi_{\beta\mu}\right\|^2 \\
&\leq (1-\alpha)(2\gamma)\mu + \alpha^2(\sqrt{2}/4)\left((2\gamma+\delta)^2/(1-2\gamma)\right)\mu \\
&\leq \left((1-\alpha)/2 + \alpha^2(\sqrt{2}/2)(1/2+\delta)^2\right)\mu(\alpha)/(1-\alpha(1-\beta)).
\end{aligned}$$

With $\delta \leq 1/5$, it follows that $\sqrt{2}(1/2+\delta)^2 \leq 49\sqrt{2}/100 < 4/5$, and $\beta \geq 1-\delta/\sqrt{n} \geq 1-\delta \geq 4/5$. Because $\alpha \geq \alpha^2$ for all $\alpha \in [0,1]$, the above yields

$$\begin{aligned}
\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| &\leq \frac{1-\alpha+\alpha^2\sqrt{2}(1/2+\delta)^2}{1-\alpha(1-\beta)} \cdot \frac{\mu(\alpha)}{2} \\
&\leq \frac{1-\alpha\left(1-\sqrt{2}(1/2+\delta)^2\right)}{1-\alpha(1-\beta)} \cdot \frac{\mu(\alpha)}{2} \leq \frac{\mu(\alpha)}{2}
\end{aligned}$$

and thus $X(\alpha)s(\alpha) \geq (1/2)\mu(\alpha)e > 0$ for all such α. This also implies that $(x(\alpha), s(\alpha)) > 0$ by continuity and thus shows that a full step is feasible with $(x(\alpha), s(\alpha)) \in \mathcal{N}_2(2\gamma)$ for all $\alpha \in [0,1]$. □

Our final result in this section follows as a corollary from the more general Lemma 2.7 specifically for restoring a barrier parameter that has been changed from μ to $\bar{\mu}$ in the second corrector step of Lemma 2.6, by setting $\beta = \mu/\bar{\mu}$. The new upper bound of $\delta \leq 1/6$ compared to $\delta \leq 1/5$ in Lemma 2.7 is intentional as will become clear from the proof.

Lemma 2.8 (Corollary to Lemma 2.7) *Let $\beta = \mu/\bar{\mu}$, $\gamma = 1/4$, $\delta \leq 1/6$, $(\bar{x}, \bar{s}) \in \mathcal{N}_2(2\gamma)$ with $\bar{x}^T\bar{s} = \bar{\mu}$, and $(\Delta x, \Delta s)$ be the solution of system (9) at $(\bar{x}, \bar{s})$ with $\xi_{\beta\mu} = \mu e - \bar{X}\bar{s}$. If*

$$\left(1-\delta/\sqrt{n}\right)\mu \leq \bar{\mu} \leq \left(1+\delta/\sqrt{n}\right)\mu \tag{17}$$

then the full Newton step is feasible and $(\bar{x}(\alpha), \bar{s}(\alpha)) \in \mathcal{N}_2(2\gamma)$ *for all* $\alpha \in [0,1]$.

Proof. Let $\beta_1 = (1+\delta/\sqrt{n})^{-1}$ and $\beta_2 = (1-\delta/\sqrt{n})^{-1}$ so that (17) can equivalently be written as $\beta_1 \le \beta \le \beta_2$. From

$$1-\beta_1 = 1-\left(1+\frac{\delta}{\sqrt{n}}\right)^{-1} = 1-\frac{\sqrt{n}}{\sqrt{n}+\delta} = \frac{\delta}{\sqrt{n}+\delta} > 0$$
$$1-\beta_2 = 1-\left(1-\frac{\delta}{\sqrt{n}}\right)^{-1} = 1-\frac{\sqrt{n}}{\sqrt{n}-\delta} = \frac{-\delta}{\sqrt{n}-\delta} < 0$$

we see that $|1-\beta_2| = \delta/(\sqrt{n}-\delta) > \delta/(\sqrt{n}+\delta) = |1-\beta_1|$ which implies that

$$|1-\beta| \le \max\{|1-\beta_1|, |1-\beta_2|\} = \delta/\left(\sqrt{n}-\delta\right).$$

With $\delta \le 1/6$, it now follows that $|1-\beta| \le 1/(6\sqrt{n}-1) \le 1/(5\sqrt{n})$ and the result follows immediately from Lemma 2.7. □

Finally, from (12) and (13) we also observe that a full step with $\alpha = 1$ fully recovers the old barrier parameter $\mu = \beta\bar{\mu}$ whereas a general step with $\alpha \in [0,1]$ still achieves a new barrier parameter $\bar{\mu}(\alpha)$ whose difference to μ is no less than the difference between $\bar{\mu}$ and μ.

2.2 Polynomiality of a Feasible Predictor-Corrector Method

After the first polynomiality proofs for LP using the projective-scaling method by Karmarkar [8] and the ellipsoid method by Khachiyan [9], a variety of conceptually and notationally much simpler interior-point algorithms was proposed and subsequently proved to converge in polynomial time by Kojima et al. [10] and Mizuno [16, 17], among many others. We base the new algorithm in this paper on the feasible predictor-corrector algorithm described by Mizuno et al. [18] that alternates between affine-scaling predictor steps within the neighborhood $\mathcal{N}_2(1/2)$ and pure-centering corrector steps to recenter each iterate within the initial neighborhood $\mathcal{N}_2(1/4)$.

Algorithm 1 (Feasible Predictor-Corrector Algorithm in [18]) *Let problem* (6) *be given.*

Step 1 (Initialization): *Set* $\gamma = 1/4$ *and* $\epsilon > 0$. *Let* (x^1, y^1, s^1) *be a strictly feasible iterate with* $(x^1, s^1) \in \mathcal{N}_2(\gamma)$. *Set* $\kappa > 0$ *such that* $(x^1)^T s^1 \le (1/\epsilon)^\kappa$. *Set* $\mu^1 = (x^1)^T s^1/n$ *and* $k = 1$.

Step 2 (Termination): *If* $(x^k)^T s^k \le \epsilon$ *then stop with*

$$(x^*, y^*, s^*) = (x^k, y^k, s^k).$$

Step 3 (Predictor Step): *Set* $(x, y, s, \mu) = (x^k, y^k, s^k, \mu^k)$, $\beta = 0$, *and compute the unique solution* $(\Delta x, \Delta y, \Delta s)$ *from system* (9) *at* (x, s) *with* $\xi_{\beta\mu} = -Xs$. *Let* $\bar{\alpha}$ *be the largest step size such that* $(x(\alpha), s(\alpha)) \in \mathcal{N}_2(2\gamma)$ *holds for all* $\alpha \in [0, \bar{\alpha}]$, *and set*

$$(\bar{x}, \bar{y}, \bar{s}, \bar{\mu}) = (x(\bar{\alpha}), y(\bar{\alpha}), s(\bar{\alpha}), \mu(\bar{\alpha})).$$

Step 4 (Corrector Step): *Set* $\beta = 1$, *compute the unique solution* $(\Delta\bar{x}, \Delta\bar{y}, \Delta\bar{s})$ *from system* (9) *at* $(\bar{x}, \bar{s})$ *with* $\xi_{\beta\mu} = \bar{\mu}e - \bar{X}\bar{s}$, *and set*

$$(x^{k+1}, y^{k+1}, s^{k+1}, \mu^{k+1}) = (\bar{x}, \bar{y}, \bar{s}, \bar{\mu}) + (\Delta\bar{x}, \Delta\bar{y}, \Delta\bar{s}, 0).$$

Step 5: (Reiteration): *Increase* k *by 1 and go to Step 2.*

To analyze the above algorithm, we recall that the predictor step maintains intermediate iterates $(\bar{x}, \bar{s}) \in \mathcal{N}_2(1/2)$ so that Lemma 2.4 implies that every iteration achieves a new point (x^k, y^k, s^k) that satisfies $(x^k, s^k) \in \mathcal{N}_2(1/4)$. Hence, the predictor step size is bounded from below by

$$\bar{\alpha} \ge \alpha^* = \min\{0.5, 8^{-1/4} n^{-1/4}\}$$

according to Lemma 2.3 and reduces the barrier parameter by a factor of at least $1-\alpha^*$. From equation (12) together with the result of Lemma 2.4 that the barrier parameter remains unchanged during the corrector step, it follows that

$$(x^{k+1})^T s^{k+1} \leq (1 - 8^{-1/4} n^{-1/2})(x^k)^T s^k \leq (1 - 8^{-1/4} n^{-1/2})^k (x^1)^T s^1 \tag{18}$$

by induction, where $(x^1)^T s^1 \leq (1/\epsilon)^\kappa$. Hence, we can solve the inequality $(1 - 8^{-1/4} n^{-1/2})^k (1/\epsilon)^\kappa \leq \epsilon$ for $k \geq (\kappa+1)\log(\epsilon)/\log(1 - 8^{-1/4} n^{-1/2})$ and use the relationship that

$$\lim_{n\to\infty} \log(1 + Ln^{-K})/n^{-K} = L$$

for $K > 0$ to write $\log(1 - 8^{-1/4} n^{-1/2}) = -\mathcal{O}(n^{-1/2})$. Combining terms, the following result is shown and corresponds to Theorem 1 in [18].

Theorem 2 (Theorem 1 in [18]) *Algorithm 1 finds an ϵ-optimal solution to problem (6) in $\mathcal{O}\big((\kappa+1)\log(1/\epsilon)\sqrt{n}\big)$ iterations, if it exists.*

Note that for problem (1), this corresponds to a worst-case iteration bound of $\mathcal{O}\left((\kappa+1)\log(1/\epsilon)\sqrt{n+\ell}\right)$.

3 New Algorithm

In this section, we formulate and analyze our new algorithm for solving LP problems in the non-standard form (1). The description of the algorithm is given in two parts.

1. All feasible predictor-corrector steps are taken in Algorithm 2 and basically mimick those in Algorithm 1. However, we make sure that we do not violate any inequalities that are currently dropped from the problem by shortening the step size, if necessary. We use the counter k for the number of those iterations that guarantee the minimum step length of Lemma 2.4 and thus reduce μ according to (18)..
2. Once a new inequality has been selected, the algorithm calls the separate function `Augment` in Algorithm 3 that consists primarily of the three corrector steps previously discussed in Section 2.1. We use the counter l for the total number of calls to this function, or equivalently, for the total number of inequalities added.

In addition, we use a third counter h for the number of nested recursive calls to `Augment` after the third corrector step (from Step 4.5) in Algorithm 3, which is useful for our subsequent complexity discussion.

Algorithm 2 (Predictor-Corrector Algorithm with Selective Addition of Inequalities) *Let problem (1) be given.*

Step 1 (Initialization): *Set $\gamma = 1/4$, $\delta = 1/11$, and $\epsilon > 0$. Let (x^1, y^1, s^1) be a strictly feasible iterate for the initial instance (A, b, c, P, q) with $(x^1, s^1) \in \mathcal{N}_2^n(\gamma)$.. Set $\mu^1 = (x^1)^T s^1/n$, and $\kappa > 0$ and $\tau > 0$ be such that $(x^1)^T s^1 \leq (1/\epsilon)^\kappa$ and $Px^1 - q \geq (1/\tau)\mu^1 e$. Set $k = 1$ and $l = h = 0$.*

Step 2 (Termination): *If $(x^k)^T s^k \leq \epsilon$, then stop with*

$$(x^*, y^*, z^*) = (x^k, y^k, s^k).$$

Step 3 (Predictor Step): *Set $(x, y, s, \mu) = (x^k, y^k, s^k, \mu^k)$, $\beta = 0$, `Pred` = `True`, and compute the unique solution $(\Delta x, \Delta y, \Delta s)$ from system (9) at (x, s) with $\xi_{\beta\mu} = -Xs$. Let $\bar{\alpha}$ be the largest step size such that $(x(\alpha), s(\alpha)) \in N_2^{n+l}(2\gamma)$ and*

$$Px(\alpha) - (1/\tau)\mu(\alpha)e \geq q \tag{19a}$$

$$x(\alpha)^T s(\alpha) \geq \epsilon \tag{19b}$$

hold for all $\alpha \in [0, \bar{\alpha}]$. Set `Pred` = `False` if $\bar{\alpha}$ was decided by one of the conditions in (19), and let

$$(\bar{x}, \bar{y}, \bar{s}, \bar{\mu}) = (x(\bar{\alpha}), y(\bar{\alpha}), s(\bar{\alpha}), \mu(\bar{\alpha})).$$

Step 3.5 (Augmentation): *If $\bar{\alpha}$ was decided by (19a), then call Algorithm 3 and set*

$$(A, b, c, P, q, \bar{x}, \bar{y}, \bar{s}, l, h) = \texttt{Augment}(A, b, c, P, q, \bar{x}, \bar{y}, \bar{s}, l, h).$$

Step 4 (Corrector Step): *Set $\beta = 1$ and compute the unique solution $(\Delta\bar{x}, \Delta\bar{y}, \Delta\bar{s})$ from system (9) at $(\bar{x}, \bar{s})$ with $\xi_{\beta\mu} = \bar{\mu}e - \bar{X}\bar{s}$. Let $\tilde{\alpha} \leq 1$ be the largest step size such that*

$$P\bar{x}(\alpha) - (\bar{\mu}/\tau)e \geq q \tag{20}$$

holds for all $\alpha \in [0, \tilde{\alpha}]$, and let

$$(\tilde{x}, \tilde{y}, \tilde{s}) = (\bar{x}(\tilde{\alpha}), \bar{y}(\tilde{\alpha}), \bar{s}(\tilde{\alpha})).$$

Step 4.5 (Augmentation): *If $\tilde{\alpha} < 1$ was decided by (20), then call Algorithm 3 and repeat Step 4 with*

$$(A, b, c, P, q, \bar{x}, \bar{y}, \bar{s}, l, h) = \texttt{Augment}(A, b, c, P, q, \tilde{x}, \tilde{y}, \tilde{s}, l, h).$$

Step 5 (Reiteration): *Increase k by 1 if* `Pred` = `True`*, and go back to Step 2 with*

$$(x^k, y^k, s^k, \mu^k) = (\tilde{x}, \tilde{y}, \tilde{s}, \bar{\mu}).$$

We point out the main differences between Algorithms 1 and 2. First, unless we start Algorithm 2 with an ϵ-optimal solution so that it stops already in Step 2, condition (19b) implies that Algorithm 2 never reduces $x^T s$ below ϵ in the predictor step but terminates after a final corrector step with an exact ϵ-solution (x^*, y^*, s^*) with $(x^*)^T s^* = \epsilon$ and $(x^*, s^*) \in \mathcal{N}_2(\gamma)$. We will utilize this condition later in the proof of Lemma 3.2. Similarly, conditions (19a) and (20) imply that the residuals $r = Px - q$ are never reduced below the residual threshold $\rho = \mu/\tau$: whenever an inequality reaches that threshold and thus determines the maximum step size, we call the function `Augment` in Algorithm 3 to add that inequality. In particular, this implies that in Algorithm 3 we can always select at least one inequality $(p^T, \pi) \in \mathbb{R}^{n+l} \times \mathbb{R}$ from $(P, q) \in \mathbb{R}^{(\ell-l)\times(n+l)} \times \mathbb{R}^{\ell-l}$ that satisfies $p^T x - \rho = \pi$ with equality at the current primal iterate. Besides ρ, the other parameters γ, δ, and τ are as in Algorithm 2. The variables ι and ζ to refer to a specific inequality and its residual, whose relevance will become clear in our further explanation throughout this section.

Algorithm 3 **(**`Augment`**$(A, b, c, P, q, x, y, s, l, h)$)** *Let (x, y, s) be the current iterate for problem instance (A, b, c, P, q) with l inequalities, and let h be the current number of nested recursive calls to the function* `Augment` *after the third corrector step (from Step 4.5 below) in previous instances of Algorithm 3.*

Step 1 (Augmentation): *Let $\mu = x^T s/(n + l)$, $\rho = \mu/\tau$, and select a single inequality (p^T, π) from (P, q) such that $p^T x - \rho = \pi$. Let $(\check{P}, \check{q})$ be the subsystem of (P, q) without this inequality and augment*

$$\left(\hat{A}, \hat{b}, \hat{c}, \hat{P}, \hat{q}\right) := \left(\begin{bmatrix} A & 0 \\ p^T & -1 \end{bmatrix}, \begin{bmatrix} b \\ \pi \end{bmatrix}, \begin{bmatrix} c \\ 0 \end{bmatrix}, \begin{bmatrix} \check{P} & 0 \end{bmatrix}, \check{q}\right) \tag{21a}$$

$$(\hat{x}, \hat{y}, \hat{s}) = (\hat{x}_r, \hat{y}_z, \hat{s}_t) := \left(\begin{bmatrix} x \\ r \end{bmatrix}, \begin{bmatrix} y \\ z \end{bmatrix}, \begin{bmatrix} s \\ t \end{bmatrix}\right) = \left(\begin{bmatrix} x \\ \rho \end{bmatrix}, \begin{bmatrix} y \\ 0 \end{bmatrix}, \begin{bmatrix} s \\ \tau \end{bmatrix}\right). \tag{21b}$$

Increase l by 1 and (locally) set $\zeta = -\tau$ and $\iota = l$.

Step 2 (First Corrector/Recentering Step): *Set $\beta = 1$ and compute the unique solution $(\Delta\hat{x}, \Delta\hat{y}, \Delta\hat{s})$ of the augmented system (9) at $(\hat{x}, \hat{s})$ with $\xi_{\beta\mu} = \mu e - \hat{X}\hat{s}$. Let $\alpha' \leq 1$ be the largest step size such that*

$$\hat{P}\hat{x}(\alpha) - (\mu/\tau)e \geq \hat{q} \tag{22}$$

holds for all $\alpha \in [0, \alpha']$, and set

$$(\hat{x}', \hat{y}', \hat{s}') = (\hat{x}(\alpha'), \hat{y}(\alpha'), \hat{s}(\alpha'))\,. \tag{23}$$

Step 2.5 (Augmentation): *If $\alpha' < 1$, then call Algorithm 3 and repeat Step 2 with*

$$(\hat{A}, \hat{b}, \hat{c}, \hat{P}, \hat{q}, \hat{x}, \hat{y}, \hat{s}, l, h) = \texttt{Augment}(\hat{A}, \hat{b}, \hat{c}, \hat{P}, \hat{q}, \hat{x}', \hat{y}', \hat{s}', l, h).$$

Step 3 (Second Corrector/Feasibility Step): *Let $\xi_c \in \mathbb{R}^{n+l}$ be the vector with entry ζ in component $n+\iota$ and zero everywhere else, and compute the unique solution $(\Delta\hat{x}', \Delta\hat{y}', \Delta\hat{s}')$ of the augmented system (14) at $(\hat{x}', \hat{s}')$ with ξ_c as defined above. Let $\alpha'' \le 1$ be the largest step size such that*

$$(\hat{x}', \hat{s}'(\alpha)) \in \mathcal{N}_2^{n+l}(2\gamma) \tag{24a}$$

$$\left(1 - \frac{\delta}{\sqrt{n+l}}\right)(\hat{x}')^T\hat{s}' \le (\hat{x}')^T\hat{s}'(\alpha) \le \left(1 + \frac{\delta}{\sqrt{n+l}}\right)(\hat{x}')^T\hat{s}'. \tag{24b}$$

hold for all $\alpha \in [0, \alpha'']$, reduce ζ to $(1-\alpha'')\zeta$, and set

$$(\hat{x}'', \hat{y}'', \hat{s}'') = (\hat{x}', \hat{y}'(\alpha''), \hat{s}'(\alpha'')) \,. \tag{25}$$

Step 4 (Third Corrector/Barrier Step): *Set $\hat{\mu} = (\hat{x}'')^T\hat{s}''/(n+l)$, $\beta = \mu/\hat{\mu}$, and compute the unique solution $(\Delta\hat{x}'', \Delta\hat{y}'', \Delta\hat{s}'')$ of the augmented system (9) with $\xi_{\beta\mu} = \mu e - \hat{X}''\hat{s}''$. Let $\alpha''' \le 1$ be the largest step size such that*

$$\hat{P}\hat{x}''(\alpha) - (1/\tau)\hat{\mu}(\alpha)e \ge \hat{q} \tag{26}$$

hold for all $\alpha \in [0, \alpha''']$, and let

$$(\hat{x}, \hat{y}, \hat{s}) = (\hat{x}''(\alpha'''), \hat{y}''(\alpha'''), \hat{s}''(\alpha''')) \,.$$

Step 4.5 (Augmentation): *If $\alpha''' < 1$, then increase h by 1, call Algorithm 3, decrease h by 1, and repeat Step 4 with*

$$(\hat{A}, \hat{b}, \hat{c}, \hat{P}, \hat{q}, \hat{x}'', \hat{y}'', \hat{s}'', l, h) = \texttt{Augment}(\hat{A}, \hat{b}, \hat{c}, \hat{P}, \hat{q}, \hat{x}'', \hat{y}'', \hat{s}'', l, h).$$

Step 5 (Reiteration/Return): *If $|\zeta| > 0$ go back to Step 2; otherwise return $(\hat{A}, \hat{b}, \hat{c}, \hat{P}, \hat{q}, \hat{x}, \hat{y}, \hat{s}, l, h)$.*

Before we formally analyze Algorithm 3, we briefly explain its basic ideas. In Step 1, we remove the selected inequality $p^Tx \ge \pi$ from (P, q) and add it as slacked equality to the augmented instance $(\hat{A}, \hat{b}, \hat{c})$. At the same time, we also augment the current iterate (x, y, s) with variables $(r, z, t) = (\rho, 0, \tau)$ whose values are chosen so that $rt = \rho\tau = \mu$, $p^Tx - r = p^Tx - \rho = \pi$, and $\xi_c = c - A^Ty - zp^T - s = c - A^Ty - s$. By Lemma 3.1 below, these choices guarantee that the augmented iterate $(\hat{x}_r, \hat{y}_z, \hat{s}_t)$ maintains centrality and the barrier parameter μ as well as primal feasibility and residuals with respect to all previous dual constraints, if any. In particular, it follows that the only infeasibility in the augmented residual $\hat{\xi}_c = [\xi_c, \zeta]^T$ is caused by the new dual constraint $-z + t = 0$ with an initial residual of $\zeta = -\tau$ in the last component of $\hat{\xi}_c$ indexed by $\iota = l$. The three following corrector steps 2, 3, and 4 work together to successively reduce ζ to zero while maintaining centrality, the current barrier parameter, and all other residuals as outlined in the introduction:

1. The first corrector step corresponds to the regular centering step from Section 2.1.2. Like Step 4 in Algorithm 2, this step does not change the barrier parameter or any residuals and serves to guarantee sufficient centrality before taking the second corrector step.
2. The second corrector step corresponds to the feasibility step from Section 2.1.3 to reduce ζ. It updates the dual iterate and keeps all residuals other than ζ the same. However, because the complementarity products are not recentered the barrier parameter μ may generally change to a new $\hat{\mu}$ within the bounds imposed in the step size condition (24b).
3. The third corrector step corresponds to the modified centering step from Section 2.1.4 with $\beta = \mu/\hat{\mu}$ and is designed to restore the barrier parameter without any other changes to centrality, primal feasibility, and dual residuals. The counter h serves to keep track of the current number of nested recursive calls to the function `Augment` after this step (from Step 4.5), which is useful for our later discussion of the algorithm's complexity.

In addition, we also continue to check the need for adding new inequalities from $(\hat{P}, \hat{q})$ whenever taking a primal step in the first and third corrector step in Algorithm 3. In this case, the algorithm calls itself recursively to add new inequalities while preserving the remaining infeasibility ζ in the component ι of $\hat{\xi}_c$ until all new infeasibilities in those components indexed by $n+\iota+1, \ldots, n+l$ are fully absorbed. In other words, with each recursive call of Algorithm 3 a new inequality is added immediately and the infeasibilities are absorbed in a last-in-first-feasible fashion.

3.1 Main Results

To analyze the new algorithm in detail, we begin by considering Algorithm 3 for a general call to the `Augment` function of the form

$$(\hat{A}, \hat{b}, \hat{c}, \hat{P}, \hat{q}, \hat{x}, \hat{y}, \hat{s}, l, H) = \texttt{Augment}(A, b, c, P, q, x, y, s, j, h). \quad (27)$$

Because the value of h is only changed in Step 4.5 in Algorithm 3, it is apparent that $H = h$ unless the function `Augment` is called after the third predictor step of Algorithm 3, in which case $H = h+1$. Theorem 3 and its proof in Subsections 3.1.1 and 3.1.2 are the main contribution of this paper.

Theorem 3 *Let $\gamma = 1/4$, $\delta = 1/11$, and ϵ and τ be as in Algorithms 2 and 3. Let (x, y, s) be the iterate for an instance (A, b, c, P, q) of problem (1) with j inequalities that satisfies $(x, s) \in \mathcal{N}_2^{n+j}(2\gamma)$, $\mu = x^T s/(n + j)$, $Ax = b$, $Px - (1/\tau)\mu \geq q$, and $\xi_c = c - A^T y - s$ when calling the function `Augment` from Step 3.5 or 4.5 in Algorithm 2, or Step 2.5 or 4.5 in Algorithm 3. Algorithm 3 returns the output of the function call (27) in $\mathcal{O}((\tau/\epsilon)(n+l)^{3/2}e^{\delta\theta})$ iterations and $l-j-1$ recursive calls to itself, where $\theta = h/\sqrt{n+l}$, $(\hat{x}, \hat{s}) \in \mathcal{N}_2^{n+l}(2\gamma)$, $\hat{\mu} = \hat{x}^T\hat{s}/(n + l) = \mu$, $\hat{A}\hat{x} = \hat{b}$, $\hat{P}\hat{x} - (1/\tau)\hat{\mu} \geq \hat{q}$, and $\hat{\xi}_c = \hat{c} - \hat{A}^T\hat{y} - \hat{s} = [\xi_c, 0]^T$ with $0 \in \mathbb{R}^{l-j-1}$. In particular, if (y, s) is dual feasible, then $(\hat{y}, \hat{s})$ is dual feasible.*

Starting from a primal-dual feasible point in Algorithm 2, Theorem 3 implies that the integration of new inequalities in Algorithm 3 do not change the current iterate's feasibility, its centrality in a wide neighborhood, and its current barrier parameter. However, if the `Augment` function is called from Step 3.5 in Algorithm 2, then the step size $\bar{\alpha}$ in the predictor step is decided by (19a) so that we cannot guarantee the minimum step length from Lemma 2.4 and the corresponding reduction of the barrier parameter in (18). Hence, we do not increase the iteration counter k if an inequality is added after the predictor step, so that the estimate in (18) is still true, and we count those iterations in which we call Algorithm 3 to add an inequality separately using the inequality counter l. Similarly, if the `Augment` function is called from Step 4.5, then the step size $\bar{\alpha}$ in the corrector step is decided by (20) and although we know from (12) and Lemma 3.3 that we have not changed the barrier parameter, we cannot guarantee that the new iterate is sufficiently recentered. Because Lemma 2.8 with $\delta = 0$ implies that the new iterate (x, y, s) still satisfies $(x, s) \in \mathcal{N}_2^{n+l}(2\gamma)$, we can repeat this step until no new inequality needs to be added and a full corrector step can be taken to recenter the current iterate. Hence, the following theorem follows as a corollary from Theorems 2 and 3.

Theorem 4 *Algorithm 2 finds an ϵ-optimal solution to problem (1) in $\mathcal{O}((\kappa+1)\log(1/\epsilon)(n+l)^{1/2})$ iterations and $\mathcal{O}(l)$ calls to Algorithm 3, where l is the number of inequalities added to the problem at optimality.*

(a) *The combined algorithm terminates in $\mathcal{O}(((\kappa+\tau+1)/\epsilon)l(n+l)^{3/2}e^{\delta\theta})$ iterations, where $\theta = \mathcal{O}\left(h/\sqrt{n+l}\right)$ and h is the maximum number of nested recursive calls to Algorithm 3 after the third corrector step.*

(b) *If $h \in \mathcal{O}\left(\sqrt{n+l}\right)$, then $\theta = \mathcal{O}(1)$ and the combined algorithm terminates in $\mathcal{O}((\kappa+\tau+1)/\epsilon)l(n+l)^{3/2})$ iterations.*

Note that the statements about the combined algorithm use the weak estimate $\log(1/\epsilon) = \mathcal{O}(1/\epsilon)$ to combine the iteration complexities from Algorithms 2 and 3. Also note that in the worst-case, h may be as large as l so that Theorem 4 reduces to Theorem 1 as stated in Section 1.1. Extending that discussion, however, now Theorem 4 implies that the algorithm stays polynomial even for an arbitrarily large number l, as long as the maximum value of h stays sufficiently small. The reason and other implications of this condition are revealed in the following proof and further discussed in Section 3.2.

3.1.1 Proof of Theorem 3 when Adding a Single Inequality

The proof of Theorem 3 is split into two parts and starts with the analysis of Algorithm 3 called from Step 3.5 or 4.5 of Algorithm 2 in which we can take full Newton steps in all first and third corrector steps. In particular, this means that only a single inequality is added so that $j = l - 1$ and $H = h = 0$ in the function call (27), which therefore can be written as

$$(\hat{A}, \hat{b}, \hat{c}, \hat{P}, \hat{q}, \hat{x}, \hat{y}, \hat{s}, l, 0) = \texttt{Augment}(A, b, c, P, q, x, y, s, l - 1, 0). \quad (28)$$

The first result addresses the initial augmentation step.

Lemma 3.1 *Let (x, y, s) be the current iterate for instance (A, b, c, P, q) of problem (1) with $l-1$ inequalities. Denote $\mu = x^T s/(n + l - 1)$, $\rho = \mu/\tau$, and let $p^T x - \rho = \pi$ and $(\hat{x}, \hat{y}, \hat{s}) = (\hat{x}_r, \hat{y}_z, \hat{s}_t)$ be the added inequality and the augmented iterate defined in (21) in Step 1 of Algorithm 3, respectively.*

(a) *If $(r, t) = (\rho, \tau)$, $(x, s) \in \mathcal{N}_2^{n+l-1}(2\gamma)$, $Ax = b$, and $Px - \rho e \geq q$, then $\hat{\mu} = \hat{x}^T \hat{s}/(n + l) = \mu$, $(\hat{x}, \hat{s}) \in \mathcal{N}_2^{n+l}(2\gamma)$, $\hat{A}\hat{x} = \hat{b}$, and $\hat{P}\hat{x} - \hat{\mu}/\tau \geq \hat{q}$.*

(b) *If $\xi_c = c - A^T y - s$, then $\hat{\xi}_c := \hat{c} - \hat{A}^T \hat{y} - \hat{s} = [\xi_c, \zeta]^T$ with $\zeta = -\tau$. In particular, if (y, s) is dual feasible, then $(\hat{y}, \hat{s})$ has a single nonzero dual residual of value $-\tau$.*

Proof. Let (x, s) and (r, t) satisfy the assumptions of the lemma. For (a), we first compute that

$$\hat{\mu} = \frac{\hat{x}_r^T \hat{s}_t}{n + l} = \frac{x^T s + rt}{n + l} = \frac{(n + l - 1)\mu + \mu}{n + l} = \mu.$$

Second, from $(x, s) \in \mathcal{N}_2^{n+l-1}(2\gamma)$ we know that $\|Xs - \mu e\| \leq (2\gamma)\mu$ which then implies that

$$\left\|\hat{X}_r \hat{s}_t - \hat{\mu} e\right\| = \left\|\hat{X}_r \hat{s}_t - \mu e\right\| = \left\|\begin{bmatrix} Xs - \mu e \\ rt - \mu \end{bmatrix}\right\| = \|Xs - \mu e\| \leq (2\gamma)\mu = (2\gamma)\hat{\mu}$$

so that $(\hat{x}, \hat{s}) \in \mathcal{N}_2^{n+l}(2\gamma)$. Third, we can explicitly write down the new system of primal constraints and substitute $r = \rho = p^T x - \pi$ to find that

$$\hat{A}\hat{x}_r = \begin{bmatrix} A & 0 \\ p^T & -1 \end{bmatrix} \begin{bmatrix} x \\ r \end{bmatrix} = \begin{bmatrix} Ax \\ p^T x - r \end{bmatrix} = \begin{bmatrix} Ax \\ p^T x - \rho \end{bmatrix} = \begin{bmatrix} b \\ \pi \end{bmatrix} = \hat{b}$$

which shows that the augmented iterate remains primal feasible for the augmented instance. Fourth, we have

$$\hat{P}\hat{x}_r - \rho = \begin{bmatrix} \check{P} & 0 \end{bmatrix} \begin{bmatrix} x \\ r \end{bmatrix} - \rho e = \check{P}x - \rho e \geq \check{q}$$

because all inequalities in $(\check{P}, \check{q})$ are also contained in (P, q) for which $Px - \rho e \geq q$. Fifth and finally, similar to primal feasibility we can explicitly write down the new dual residual vector and substitute $(z, t) = (0, \tau)$ to verify that

$$\hat{\xi}_c = \begin{bmatrix} c \\ 0 \end{bmatrix} - \begin{bmatrix} A^T & p \\ 0 & -1 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} - \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} c - A^T y - zp - s \\ 0 - (-z) - t) \end{bmatrix} = \begin{bmatrix} \xi_c \\ -\tau \end{bmatrix}.$$

□

Lemma 3.1 shows that for an initial iterate (x, y, s) that satisfies the given assumptions, the augmented iterate $(\hat{x}, \hat{y}, \hat{s})$ maintains centrality, barrier parameter, primal feasibility, dual residuals of all previous constraints, if any, and $\hat{P}\hat{x} - \rho e \geq \hat{q}$. This also means that $\alpha' = 0$ is a feasible step size in the first corrector step that therefore is well-defined during the first iteration. Furthermore, by our current assumption that we can take a full first correcter step, Lemma 2.4 implies that the next iterate $(\hat{x}', \hat{y}', \hat{s}') = (\hat{x}(1), \hat{y}(1), \hat{s}(1))$ defined in (23) satisfies $(\hat{x}', \hat{s}') \in \mathcal{N}_2^{n+l}(\gamma)$ with $(\hat{x}')^T \hat{s}'/(n + l) = \mu$, and $\hat{P}\hat{x}' - (\mu/\tau)e \geq q$ according to (22). Hence, $\alpha'' = 0$ is also feasible for the subsequent second corrector that therefore is well-defined in the first iteration as well. The next lemma gives a lower bound on its step size and the amount of infeasibility absorbed, which depend on the termination criterion $\epsilon > 0$ in Algorithm 2 and the upper bound M from Assumption 3.

Lemma 3.2 (Corollary to Lemma 2.6) *Let $\gamma = 1/4$, $\delta = 1/11$, and $(\hat{x}', \hat{y}', \hat{s}')$ be the iterate in the second corrector step of Algorithm 3 with $(\hat{x}', \hat{s}') \in \mathcal{N}_2^{n+l}(\gamma)$, $\mu = (\hat{x}')^T(\hat{s}')/(n + l)$, $\hat{P}\hat{x}' - (\mu/\tau)e \geq \hat{q}$, and step direction $(\Delta\hat{x}', \Delta\hat{y}', \Delta\hat{s}')$. Let $\xi_c = [0, \zeta]$ with $|\zeta| \leq \tau$, $\alpha'' \leq 1$, and $(\hat{x}'', \hat{y}'', \hat{s}'') = (\hat{x}', \hat{y}'(\alpha''), \hat{s}'(\alpha''))$ be defined as in the second corrector step of Algorithm 3.*

(a) *If $\alpha'' = 1$, then $|\zeta|$ is reduced to zero.*

(b) *If* $\alpha'' < 1$, *then the step size satisfies* $\alpha'' \geq (1-\gamma^2)^{1/2}\delta\epsilon / \big(M(n+\ell)^{3/2}\tau\big)$ *and* $|\zeta|$ *is reduced by at least* $(1-\gamma^2)^{1/2}\delta\epsilon / \big(M(n+\ell)^{3/2}\big)$.

Proof. Part (a) is clear from (15). For (b), we collect several bounds and then apply Lemma 2.6. Writing $(\hat{x}', \hat{y}', \hat{s}') = (\hat{x}'_r, \hat{y}'_z, \hat{s}'_t)$, from $(\hat{x}', \hat{s}') \in \mathcal{N}_2^{n+l}(\gamma)$ with $\mu = (\hat{x}')^T \hat{s}'/(n+l)$ we know that

$$r't' \geq (1-\gamma)\mu.$$

where $r' \leq M$ by Assumption 3 and $\mu = \hat{x}^T \hat{s}/(n+l) \geq \epsilon/(n+l)$ by (19b). It follows that

$$t' \geq (1-\gamma)\epsilon/(M(n+l)), \tag{29}$$

and using $\xi_c = [0, \zeta]^T$ we find

$$\left\|(\hat{S}'_t)^{-1}\xi_c\right\| = \left\|\begin{bmatrix} S' & 0 \\ 0 & t' \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \zeta \end{bmatrix}\right\| = \frac{|\zeta|}{t'} \leq \frac{M(n+l)\tau}{(1-\gamma)\epsilon}. \tag{30}$$

Hence, using Lemma 2.6 the first statement of part (b) follows from

$$\alpha'' \geq \sigma \geq \frac{\lambda}{\left\|(\hat{S}'_t)^{-1}\xi_c\right\|} = \frac{\delta}{\sqrt{n+l}}\left(\frac{1+\gamma}{1-\gamma}\right)^{1/2} \frac{(1-\gamma)\epsilon}{M(n+l)\tau} = \frac{(1-\gamma^2)^{1/2}\delta\epsilon}{M(n+\ell)^{3/2}\tau}.$$

Combining the above with (29), (30), and Lemma 2.6, the second statement is shown similarly by

$$|\zeta| - (1-\alpha'')|\zeta| = \alpha''|\zeta| \geq \sigma|\zeta| = \frac{\lambda|\zeta|}{\|S^{-1}\xi_c\|} = \lambda t' \geq \frac{(1-\gamma^2)^{1/2}\delta\epsilon}{M(n+\ell)^{3/2}}.$$

□

After the second corrector step, we know from (24) that the new iterate $(\hat{x}'', \hat{y}'', \hat{s}'')$ defined in (25) satisfies $(\hat{x}'', \hat{s}'') \in \mathcal{N}_2^{n+l}(2\gamma)$ and the bounds in (17) with $n+l$ for $\bar{\mu} = (\hat{x}'')^T\hat{s}''/(n+l)$. In particular, because $\hat{x}'' = \hat{x}'$ from before, we still have $\hat{P}\hat{x}'' - \rho e = \hat{q}$. Hence, for the third time, the step size $\alpha''' = 0$ is feasible so that the subsequent third corrector step is well-defined. Under our assumption that a full Newton step is feasible, now equation (12) implies that this step recovers the barrier parameter μ from $\bar{\mu}$ and achieves a new iterate $(\hat{x}, \hat{y}, \hat{s})$ that still satisfies $(\hat{x}, \hat{s}) \in \mathcal{N}_2^{n+l}(2\gamma)$ by Lemma 2.8, and thus again satisfies all properties of Lemma 3.1(a). By induction, this shows that in every iteration of Algorithm 3 that does not add a new inequality, we maintain centrality, the barrier parameter, primal feasibility, and dual residuals in all but the new dual constraint, in which we continue to absorb infeasibility according to Lemma 3.2. Because the initial residual is $|\zeta| = \tau$, it follows that the number of steps required to fully restore feasibility is bounded from above by

$$\tau M(n+\ell)^{3/2}/((1-\gamma^2)^{1/2}\delta\epsilon) = \mathcal{O}((\tau/\epsilon)\,(n+l)^{3/2}).$$

This completes the analysis of a single instance of Algorithm 3 and proves the following lemma as a special case of Theorem 3.

Lemma 3.3 *If* (x, s, y) *satisfies the assumptions in Lemma 3.1, then Algorithm 3 returns the output of the function call* (28) *with* $j = l-1$ *and* $H = h = 0$ *in* $\mathcal{O}((\tau/\epsilon)(n+l)^{3/2})$ *iterations and with a new iterate* $(\hat{x}, \hat{y}, \hat{s})$ *that satisfies the properties in Lemma 3.1(a) and has a dual residual* $\hat{\xi_c} = [\xi_c, 0]^T$.

3.1.2 Proof of Theorem 3 when Adding Multiple Inequalities

Next, we analyze the case where we also add inequalities from within Algorithm 3 after taking a primal step in either the first or third corrector step, when the step size α is decided by (22) or (26). Because both of these steps maintain iterates in $\mathcal{N}_2(2\gamma)$ and also satisfy all other assumptions of Lemma 3.1, the augmentation step works exactly the same and maintains centrality and barrier parameter at the current iterate, and primal feasibility. In particular, whereas $\xi_c = 0$ when starting from a dual feasible point and adding only a single inequality, now we typically have $\xi_c \neq 0$ due to one or more remaining, only partially absorbed residuals from previously added inequalities.

As discussed earlier, however, we can temporarily ignore these infeasibilities when calling Algorithm 3 recursively and restore feasibility of all added inequalities in a last-in-first-feasible fashion. In particular, given a finite number of inequalities there must be a last inequality that is added for which the analysis is basically identical to that of adding a single inequality, because it maintains centrality, barrier parameter, primal feasibility, and all other dual residuals. This also implies that the process terminates and recursively returns to every inequality that has been added, to eventually restore full dual feasibility and continue in Algorithm 2.

We now look at the basic idea of the previous paragraph in a little more detail. First, we observe that Steps 2 and 2.5 in Algorithm 3 are basically identical to Steps 4 and 4.5 in Algorithm 2, and that the barrier parameter at the iterate defined in (23) is identical to that of the initial iterate (x, y, s) independent of the step size α'. Hence, like before it suffices to repeat this step upon return from the recursive call of Algorithm 3 to ensure that centrality is fully restored.

In the second case in which we call Algorithm 3 from Step 4.5, however, the barrier parameter μ that we wish to restore from $\bar{\mu}$ by setting $\beta = \mu/\bar{\mu}$ may have been only partially restored to $\mu(\alpha)$ and generally be still smaller or larger than μ. In this case, we can still repeat this step upon return from the recursive call of Algorithm 3 because equation (13) implies that the assumptions of Lemma 2.8 remain valid for the new choice of $\beta = \mu/\bar{\mu}(\alpha)$.

However, our analysis of Algorithm 3 now differs from that in Section 3.1.1 because the change in the barrier parameter may also change our estimate in Lemma 3.2. In particular, calling Algorithm 3 from Algorithm 2 with a barrier parameter $\mu \geq \epsilon$, after h nested recursive calls from the augmentation step 4.5 in Algorithm 3 the new barrier parameter could have increased or decreased and tend toward $(1+\delta/\sqrt{n+l})^h\mu$ or $(1-\delta/\sqrt{n+l})^h\mu$, respectively. Whereas the increase is unproblematic because our estimate $\bar{\mu} \geq \epsilon$ in Lemma 3.2 remains valid, the decrease is more critical because this estimate now must be replaced by the bound $\bar{\mu} \geq (1-\delta/\sqrt{n+l})^h\mu = (1-\delta/\sqrt{n+l})^h\epsilon$ to account for the worst case, which vanishes exponentially.

Lemma 3.4 (Corollary to Lemma 3.3) *Let $\theta = h/\sqrt{n+l}$. If (x, y, s) satisfies the assumptions in Lemma 3.1, then Algorithm 3 returns the output of the function call (27) in $\mathcal{O}((\tau/\epsilon)(n+l)^{3/2}e^{\delta\theta})$ iterations, with j recursive functions calls to itself, and with a new iterate $(\hat{x}, \hat{y}, \hat{s})$ that satisfies the properties in Lemma 3.1(a) and has dual residual $\hat{\xi}_c = [\xi_c, 0]^T$ with $0 \in \mathbb{R}^{l-j-1}$.*

Proof. In large parts identical to our above analysis, the proof follows analogously to that of Lemma 3.3 from Lemmas 3.1 and 3.2 if the bound $\mu \geq \epsilon/(n+l)$ in (29), or equivalently, the bound $1/\mu \leq (1/\epsilon)(n+l)$ in the rest of the proof is replaced by

$$\begin{aligned} 1/\mu &\leq (1/\epsilon)(n+l)(1-\delta/\sqrt{n+l})^{-h} \\ &= (1/\epsilon)(n+l)\left((1-\delta/\sqrt{n+l})^{\sqrt{n+l}}\right)^{-h/\sqrt{n+l}} = \mathcal{O}\left((1/\epsilon)(n+l)e^{\delta\theta}\right). \end{aligned}$$

□

To establish the remaining parts of Theorems 3 and 4, it is now sufficient to note that if $h \in \mathcal{O}(\sqrt{n+l})$, then $\theta = \mathcal{O}(1)$ so that the above bound reduces to $\mathcal{O}((1/\epsilon)(n+l)^{3/2})$ which again is polynomial in the problem dimension.

3.2 Polynomiality Conditions and Discussion

To conclude our analysis of Algorithms 2 and 3, we make a few interesting observations about the possible worst-case performance of our method. As anticipated from our discussion in the introduction, our complexity analysis shows that the proposed algorithm does not meet the polynomial time complexity of a standard IPM but has an exponential worst-case iteration bound as a consequence of the possibility of having to add a large number of inequalities in recursive corrector steps at an iterate that is already very close to optimality. This confirms another well-known, similar observation when trying to use IPMs for adding new inequalities at optimal or near-optimal solutions in practice: if the barrier parameter vanishes or becomes very small, then the amount of infeasibility that can be absorbed in every step can also become very small and IPMs tend to jam.

More specifically, our analysis has shown that the exponential worst-case behavior of the algorithm is an implication of a large number of successive, recursive nested calls after the third predictor step of Algorithm 3, that occur close to optimality and in which the preceding feasibility step draws the current barrier parameter μ to its lower permissible bound. While we have decided to not make this distinction and count all recursive calls using the parameter h to shorten the proofs and explanations of Theorems 3 and 4, our theoretical insight implies that the worst-case iteration complexity of the new algorithm will be polynomial under the following decreasingly restrictive assumptions, namely

(i) if the total number of inequalities added is of order $\mathcal{O}(\sqrt{n})$ (stated after Theorem 1);

(ii) if (i) holds only for those inequalities that are added recursively in nested calls to the `Augment` function after the third corrector step (from Step 4.5) of Algorithm 3 (stated in Theorem 4);

(iii) if (ii) holds only for those inequalities for which the preceding second corrector step reduces the barrier parameter toward its lower possible bound (stated above); and

(iv) if (iii) holds only for those inequalities for which the lower possible bound is less than the termination tolerance ϵ.

Whether it is possible to modify the algorithm or its analysis in order establish polynomiality without any of these additional assumptions remains one of our ongoing research questions.

4 Concluding Remarks

This paper describes a new method for solving linear programs in nonstandard form with equality and inequality constraints. The algorithm can be motivated for problems with a large number of known inequalities for which only a relatively small yet a priori unknown subset is active at optimality and must be added to find a feasible optimal solution. This situation occurs frequently especially for linear or more generally conic convex relaxations of combinatorial or discrete optimization problems in which the inequalities are large classes of cutting planes. The objective of this paper is to present for the first time theoretical support for a general algorithmic framework that solves optimization problems by dynamically selecting and adding inequalities as integral part of the algorithm.

To show convergence and analyze the complexity of this approach, we formulate the algorithm using a primal-dual path-following predictor-corrector short-step IPM that differs from both standard IPMs and cutting-plane methods in how we handle the inequality constraints. Unlike most standard IPMs for which the primal problem is written in standard form by adding non-negative slack variables to convert each inequality into an equality constraint, our algorithm starts with an initially reduced problem without any inequalities and selectively adds new constraints only if the reduction in their residuals indicates that they tend to become active at optimality. This also avoids the need to decide how many and, if not all, then which violated inequalities to add by a cutting-plane method, which has no best answer and is typically based on some heuristic. In particular, our algorithm maintains feasibility with respect to all primal constraints throughout, and can be terminated prematurely to find feasible nearly-optimal solutions.

While the implementation of similar algorithms has already shown encouraging results in practice [1, 2, 5, 7, 13, 14, among others], it would be interesting to take a closer look at if and how any of these previous techniques could benefit from the new theoretical insights of the analysis given in this paper. Other research directions that emerge from our work include possible enhancements in the formulation or theoretical analysis of the above algorithm to establish an improved complexity, either with or without any additional assumptions. Specifically, a variation of the feasibility-restoring step that works to restore dual feasibility simultaneously rather than recursively for each newly added inequality may remedy the current curse of recursive corrector steps and potentially lead to a new infeasible method with a possible impact on actual implementations in practice.

References

[1] Engau, A.: Recent progress in interior-point methods: Cutting plane methods and warm starts. In: M.F. Anjos, J.B. Lasserre (eds.) Handbook of Semidefinite, Conic, and Polynomial Optimization, chap. 17. Springer (2012)

[2] Engau, A., Anjos, M.F., Vannelli, A.: On handling cutting-planes in interior-point methods for solving semidefinite relaxations of binary quadratic optimization problems. Optim. Meth. Soft. (forthcoming)

[3] Evtushenko, Y.G., Golikov, A.I., Mollaverdy, N.: Augmented Lagrangian method for large-scale linear programming problems. Optim. Methods Softw. **20**(4-5), 515–524 (2005)

[4] Gondzio, J., Grothey, A.: Reoptimization with the primal-dual interior point method. SIAM J. Optim. **13**(3), 842–864 (electronic) (2003) (2002)

[5] Gruber, G., Rendl, F.: Computational experience with stable set relaxations. SIAM J. Optim. **13**(4), 1014–1028 (2003)

[6] Güler, O.: Augmented Lagrangian algorithms for linear programming. J. Optim. Theory Appl. **75**(3), 445–470 (1992)

[7] Helmberg, C., Rendl, F.: Solving quadratic $(0, 1)$-problems by semidefinite programs and cutting planes. Math. Programming **82**(3, Ser. A), 291–315 (1998)

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

[9] Khachiyan, L.G.: A polynomial algorithm in linear programming. Dokl. Akad. Nauk SSSR **244**(5), 1093–1096 (1979)

[10] Kojima, M., Megiddo, N., Mizuno, S.: A primal-dual infeasible-interior-point algorithm for linear programming. Math. Programming **61**(3, Ser. A), 263–280 (1993)

[11] Kojima, M., Mizuno, S., Yoshise, A.: A polynomial-time algorithm for a class of linear complementarity problems. Math. Programming **44**(1, (Ser. A)), 1–26 (1989)

[12] Kojima, M., Mizuno, S., Yoshise, A.: A primal-dual interior point algorithm for linear programming. In: Progress in mathematical programming (Pacific Grove, CA, 1987), pp. 29–47. Springer, New York (1989)

[13] Mitchell, J.E.: Computational experience with an interior point cutting plane algorithm. SIAM J. Optim. **10**(4), 1212–1227 (2000)

[14] Mitchell, J.E.: Polynomial interior point cutting plane methods. Optim. Methods Softw. **18**(5), 507–534 (2003)

[15] Mitchell, J.E.: Cutting plane methods and subgradient methods. In: M. Oskoorouchi (ed.) Tutorials in Operations Research, chap. 2, pp. 34–61. INFORMS (2009)

[16] Mizuno, S.: A new polynomial time method for a linear complementarity problem. Math. Programming **56**(1, Ser. A), 31–43 (1992)

[17] Mizuno, S.: Polynomiality of infeasible-interior-point algorithms for linear programming. Math. Programming **67**(1, Ser. A), 109–119 (1994)

[18] Mizuno, S., Todd, M.J., Ye, Y.: On adaptive-step primal-dual interior-point algorithms for linear programming. Math. Oper. Res. **18**(4), 964–981 (1993)

[19] Monteiro, R.D.C., Adler, I.: Interior path following primal-dual algorithms. I. Linear programming. Math. Programming **44**(1, (Ser. A)), 27–41 (1989)

[20] Monteiro, R.D.C., Adler, I.: Interior path following primal-dual algorithms. II. Convex quadratic programming. Math. Programming **44**(1, (Ser. A)), 43–66 (1989)