Convergence and polynomiality of primal-dual interior-point algorithms for linear programming with selective addition of inequalities

Citation for published version:
Engau, A & Anjos, MF 2017, 'Convergence and polynomiality of primal-dual interior-point algorithms for linear programming with selective addition of inequalities', *Optimization*, vol. 66, no. 12, pp. 2063-2086.
https://doi.org/10.1080/02331934.2016.1244268

Digital Object Identifier (DOI):
10.1080/02331934.2016.1244268

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Optimization

To appear in *Optimization*
Vol. 00, No. 00, Month 20XX, 1–26

Convergence and Polynomiality of a Primal-Dual Interior-Point Algorithm for Linear Programming with Selective Addition of Inequalities

Alexander Engau and Miguel F. Anjos

(*Received 00 Month 20XX; accepted 00 Month 20XX*)

This paper presents the convergence proof and complexity analysis of an interior-point framework that solves linear programming problems by dynamically selecting and adding inequalities. First, we formulate a new primal-dual interior-point algorithm for solving linear programs in nonstandard form with equality and inequality constraints. The algorithm uses a primal-dual path-following predictor-corrector short-step interior-point method that starts with a reduced problem without any inequalities and selectively adds a given inequality only if it becomes active on the way to optimality. Second, we prove convergence of this algorithm to an optimal solution at which all inequalities are satisfied regardless of whether they have been added by the algorithm or not. We thus provide a theoretical foundation for similar schemes already used in practice. We also establish conditions under which the complexity of the algorithm is polynomial in the problem dimension.

Keywords: linear programming, interior-point algorithms, selective addition of inequalities

AMS Subject Classification: 65K05, 90C05, 90C51

1. Introduction

Algorithms for linear programming (LP) typically assume that problems are given in standard form in which all constraints are linear equalities and all variables are nonnegative. Although inequalities can be converted to equalities using nonnegative slack variables in principle, this reformulation is generally inefficient if only few inequalities are active at optimality, and impractical if their number is much larger than the number of original variables. Common examples are semi-infinite optimization problems and continuous relaxations of combinatorial problems, where large numbers of inequalities are typically handled most efficiently using column-generation and cutting-plane methods [12, 27, 28] or other dual approaches including augmented Lagrangian relaxations [8, 15, 17]. Starting from an initial relaxation, the conventional scheme of these methods is to repeatedly solve and update successive relaxations by adding new inequalities that are violated at the currently optimal point, until the new optimal solution is also feasible for the original problem.

As an alternative to repeatedly solving relaxations especially if the number of inequalities is not too large, several other methods modify this classical cutting-plane scheme and propose to dynamically add and remove inequalities as an integral part of the solution process, see e.g. [5, 7, 14, 16, 26]. In particular, and quite different in spirit from a cutting-plane method, these algorithms do not necessarily generate violated inequalities at infeasible, optimal solutions but try to predict such inequalities already at feasible, intermediate iterates. They use different heuristic strategies to augment or adjust problems and iterates for newly added constraints or variables. In practice, this may allow to resume the algorithm without having to restart which can significantly reduce both computation times and costs. While these papers report extensive computational results that demonstrate the promise of such approaches when solving linear or semidefinite programs in practice, to the best of our knowledge they have no supporting theoretical analysis with proofs of convergence and worst-case complexity. The objective of this paper is to address this remaining gap.

1.1. *Scope of Paper and Related Literature*

The algorithm and theoretical analysis presented in this paper seek to capture the spirit of interior-point methods (IPMs) that dynamically add inequalities "on the fly" and *before* they are actually violated. The prediction of inequalities to become violated is typically based on one of several indicators [4, 9, 34] and the assumption that candidate inequalities are not too many and known in advance. The main practical challenge of these methods then is to avoid a phenomenon known as jamming, when new inequalities must be added close to optimality so that step sizes become small and the algorithm may fail to converge. The theoretical results of this paper explain this practical challenge and provide conditions under which jamming may occur or can be avoided. Hence, the scope of our paper is quite different from convergence and complexity proofs or related interior-point cutting-plane methods that separate *violated* inequalities at optimality or generate cuts using an oracle. By repeatedly solving improved formulations, and under the assumption that the separation oracle runs in polynomial time, the different challenges of these methods is to prove that the number of required inequalities remains polynomial, as shown under suitable conditions e.g. in [1, 10, 11, 25, 27].

Closer related to our own paper are several "build-up" (inequality-adding) or "build-down" (inequality-removing) methods that employ constraint-reduction techniques to lower the cost of computing search directions in variants of the dual affine-scaling and potential-reduction algorithm [2, 19, 38]. Among these methods, the build-up dual affine-scaling (BDAS) method [2] repeatedly solves a sequence of ellipsoidal subproblems and updates the description of the ellipsoid by adding new dual constraints including those that first block feasible movement into the dual affine-scaling direction. The current iterate is kept constant until enough inequalities are added and a step into the new direction again becomes feasible. In a related second paper [19], the authors propose an alternative decomposition variant based on a potential-reduction algorithm [38] that considers only "promising" constraints with small dual slacks to form the ellipsoid at each iteration and unlike BDAS is guaranteed to terminate in at most $O(\sqrt{n}L)$ steps. Applying this method to transportation problems, it is noted that "although BDAS is guaranteed to converge, there is no known theoretical rate of convergence [...] due to the fact that there is also no worst case complexity result available for [...] dual affine scaling." Moreover, these authors also encounter the problem of jamming and point out the "bouncing" behavior of their method "which caused the algorithm to slow considerably once it neared the optimal face."

It is interesting that although our algorithm uses the different framework of a primal-dual path-following IPM, its analysis suggests that a similar behavior near the optimal solution may occur, and possibly be inherent in methods of this type. Nonetheless, and in support of the favorable computational results in the literature, we can theoretically explain under which conditions such slow convergence may occur, which are likely quite rare in practice. For path-following methods specifically, theoretical insight on other effects of adding, deleting, or shifting constraints has also been utilized for a column generation and deletion variant of the logarithmic barrier method [3]. The constraint selection and addition mechanism that we analyze in our paper is quite different, however: if feasible movement is blocked after we add a new constraint, we do not simply return to one of the previous iterates and try again but continue from the current iterate with a sequence of three newly defined corrector steps. These new steps utilize recent progress in warm-starting IPMs [5, 6, 13] which is challenging in itself and often perceived not to be possible at all. Hence, the role of warm-starting in our analysis is another novel contribution of this paper.

More recent research has also explored other constraint-reduced, adaptive variants of affine-scaling or predictor-corrector methods to reduce the computational cost of generating search directions [18, 37, 39]. Applied to linear and more generally convex quadratic programs with large numbers of variables or constraints, these algorithms do not assemble the exact normal-equation matrix, but rather an approximate matrix which includes only a subset of working constraints that seem to be most critical and may change from iteration to iteration. In addition to significant cost savings, these methods achieve global convergence with a quadratic local convergence rate. Finally, this technique has also been used within a predictor-corrector algorithm for semidefinite programming [35] for which it is shown that the impact on the search direction is sufficiently small so as not to impair the polynomial-time convergence of the underlying algorithm [36].

1.2. *New Contribution and Outline of Paper*

The motivation of this paper stems from the good practical performance of "build-up" interior-point algorithms that solve linear and more generally semidefinite programs in nonstandard form by dynamically selecting and adding inequalities only if they become active on the way to optimality. To the best of our knowledge and our review of the literature, these methods currently remain without supporting theoretical analysis and without proofs of convergence and worst-case complexity. Hence, the new contribution of this paper is the formal statement of a general algorithm of this type, together with a first proof of its convergence and a detailed complexity analysis. To keep the proof relatively short in length, we base our algorithm and analysis on the basic LP framework of a feasible short-step method but believe that a similar analysis can also be extended to computationally superior infeasible (long-step) methods such as the interior-point method (IPM) implemented by the authors for solving semidefinite relaxations of binary quadratic programs [7].

We emphasize that while the basic method may appear like a cutting-plane method, there are substantial differences. First, unlike cutting-plane algorithms that separate or generate cuts at infeasible, optimal solutions, this alternative framework attempts to predict active inequalities before they become violated and add them dynamically without restarting the algorithm. Matching the empirical evidence in computational experiments [5, 7, 14, 16, 26], such methods can be effective in practice if the number of inequalities to be added is not too large. Indeed, under this same condition, our analysis for the first time shows that such methods are also efficient in theory. Second, and without further assumptions on the inequalities in the problem, however, we do not expect the algorithm presented here to be significantly better in practice or have better worst-case complexity than a standard method: some extra work is inevitable to choose and properly integrate selected inequalities, and of course all inequalities may be needed in the worst case. In particular, and a third major difference from a cutting-plane method that shall also be apparent from our problem formulation in Section 2, this general framework is generally not applicable unless the candidate inequalities are known in advance.

Finally, while the dynamic inclusion of inequalities can be accomplished quite "brutal" in practice, without restart this may also cause the algorithm to jam especially when adding many constraints near the optimal solution. To analyze and either prevent or explain such behavior in theory, our analysis requires a delicate balance between predicted infeasibilities and measures of centrality and optimality which may analogously result in exponential convergence in the worst case. Hence, another contribution of this paper is that our analysis provides clear insight into the conditions under which an algorithm of the proposed kind is polynomial or may be exponential.

The remaining paper is structured as follows. Section 2 gives the problem formulation with our assumptions and a small example that outlines the key mechanism of the new algorithm to add inequalities dynamically. Section 3 collects preliminaries and prepares our original analysis. In Section 3.1, we review some known results and derive several generalizations for the different predictor and corrector steps that are used in our method. Section 3.2 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 4. Some concluding remarks and ideas for future work are given in Section 5.

2. Problem Formulation and Main Result

We consider LP problems in the following nonstandard 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. We use X and S as the usual notation for the diagonal square matrices built of the elements of x and s, and whenever convenient we write I and $e := (1, 1, \dots, 1)^T$ for the identity matrix and the vector of all ones of suitable dimension, respectively. 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 algorithm and analysis for the more practical aspects of detecting unboundedness or infeasibility, we can adopt similar indicators and termination criteria to those used by similar IPMs [22, 30, 40].

Assumption 2 *Problem* (1) *has a feasible solution* (x, y, z, s) *that satisfies the conditions* $(x, s) > 0$, $z = 0$, $\|Xs - \mu e\| \leq (1/4)\mu$ *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 point can be found [23, 24, 32, 33, among others]. However, we do not require that problem (1) has a strictly feasible solution because we assume that $z = 0$, [**4.**] **which is a natural condition as a dual variable should not be present in the problem when removing its primal inequality**. In particular, if there was some entry of z that was strictly positive for all feasible points, then its 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 big number $M < \infty$ such that the primal residual $r = Px - q$ at any feasible point is bounded from above by M.*

Without loss of generality, we can always 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 term $\mathcal{O}(\ell)$ for the numbers of inequalities in problem (1). We will use this assumption in the proof of Lemma 4.4 where it would also suffice that $Px \leq q + Me$ only for all x that are actually encountered by the algorithm.

2.1. *Outline of Algorithm and Main Result*

Based on Assumption 2, we initialize the algorithm with a feasible iterate at which all inequalities (1c) are inactive with a residual of at least $\rho = \mu/\tau$ where μ is the initial barrier parameter. After removing these inequalities, the algorithm starts like a standard IPM and alternates between predictor steps to reduce μ and corrector steps to recenter the iterates and keep them in sufficient 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 current threshold value $\rho = \mu/\tau$, indicating that they tend to become active at optimality. The dependency of this threshold on μ makes it an adaptive threshold that decreases in proportion to the barrier parameter so that only active inequalities would be added at optimality. Whenever a new inequality is added, the algorithm augments both problem and iterate in such a way that centrality, barrier parameter, and primal feasibility are preserved. It is well known, however, that it is generally not possible to also maintain feasibility in the dual so that the algorithm temporarily continues with a sequence of three corrector steps that work together to fully absorb the new dual infeasibility:

(1) The first corrector step is a pure centering step that restore the iterate's proximity to the central path; it does not change the barrier parameter or the primal and dual residuals.
(2) The second corrector step is a feasibility-restoring step that reduces the amount of dual infeasibility; it does not modify the primal iterate but generally changes both centrality and the barrier parameter.
(3) The third corrector step is a modified centering step that does not change centrality or residuals, but serves to restore the barrier parameter that may have been changed in the second corrector step.

Whenever we change the primal iterate in the first and third corrector steps, we continue to check the residuals of dropped inequalities and, if necessary, add new inequalities and restore all infeasibilities inequality-wise starting from the last inequality added. This recursive nature of the algorithm leads to a possibly exponential worst-case complexity, which is proved in detail in Section 4.

THEOREM 2.1 *Let problem* (1) *be given, and* (x, y, s) *be a strictly feasible point that satisfies* $x^T s \le (1/\epsilon)^\kappa$ ***[6.] with*** $\epsilon > 0$ *and* $\kappa > 0$*, and Assumption 1 with* $\tau > 0$*. If the problem satisfies Assumptions 2 and 3, then the new algorithm finds an* ϵ*-optimal solution in* $\mathcal{O}(((\kappa + \tau + 1)/\epsilon)l(n + l)^{1/2}e^{\theta/11})$ *iterations, where* $\theta = \mathcal{O}(l/\sqrt{n + l})$ *and* l *is the number of inequalities that are added to the problem.*

We note that the statement of Theorem 2.1 depends on the number l of inequalities added rather than the number ℓ of total inequalities in problem (1). In particular, if $l = \mathcal{O}(\sqrt{n})$ or $l \le \ell = \mathcal{O}(\sqrt{n})$, then $\theta = \mathcal{O}(1)$ and $e^{\theta/11} = \mathcal{O}(1)$ so that the iteration bound reduces to $\mathcal{O}(((\kappa+\tau+1)/\epsilon)l(n+l)^{1/2})$ and the complexity is polynomial **[8.] in the size of the problem**. Based on additional parameters defined in our algorithm and insight gained from its subsequent analysis, we establish much weaker conditions for polynomiality in Theorem 4.2 and Section 4.2 below.

2.2. *An Example*

To illustrate the algorithm's key idea of adding relevant inequalities dynamically, we use the following problem with the obvious optimal solution $x^* = 1$:

$$\min x \text{ s.t. } x \ge 1 \text{ and } x \ge 0. \tag{2}$$

This problem is simple enough to be solved without centering steps so that we can focus primarily on the augmentation mechanism to add the necessary inequality. We start with the reduced problem in primal-dual standard form:

$$\min x \text{ s.t. } x \ge 0, \qquad\qquad \max 0 \text{ s.t. } s = 1 \text{ and } s \ge 0. \tag{3}$$

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 $(x, s) = (x_0, 1)$ at which the dropped inequality is satisfied, say $x_0 = 4$ and $\mu = xs = x_0 = 4$ so that $x_0 - 1 = 3 \ge (1/\tau)4$ for any $\tau \ge 4/3$. 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 problem (3) in a single step. Because a full step into this direction violates the dropped inequality $x \ge 1$, we reduce the step length to some $\alpha < 1$ and add the primal constraint $x - r = 1$ with a strictly feasible primal slack $r = x - 1 > 0$. This yields the augmented (original) problem with $\hat{A} = [1, -1]$, $\hat{b} = 1$, and $\hat{c} = [1, 0]^T$ that now we write in primal-dual standard form:

$$\begin{aligned} &\min x && \max z && \text{(4a)}\\ &\text{s.t. } x - r = 1 && \text{s.t. } z + s = 1,\ -z + t = 0 && \text{(4b)}\\ &(x, r) \ge 0 && (s, t) \ge 0 && \text{(4c)} \end{aligned}$$

To preserve the barrier parameter $\mu(\alpha) = x(\alpha)^T s$ for the augmented iterate $(\hat{x}_r, \hat{s}_t) = \big((x(\alpha), r)^T, (s, t)^T\big)$ with $t = \tau = 2$, we compute the reduced step length $\alpha = 1/2$ as the largest value for which $r = x(\alpha) - 1 \ge \mu(\alpha)/\tau$. This ensures that the resulting iterate $(x, r) = (2, 1)$ remains primal feasible and perfectly centered with $\hat{\mu} = \mu(\alpha) = 2$. 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 negative residual $\zeta = 0 - (-z + t) = -\tau = -2$. Hence, we now continue with a feasibility-restoring corrector step into the dual direction computed from the following system for the augmented problem (4):

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

With $(x, r) = (2, 1)$, $(s, t) = (1, 2)$, and $\zeta = -2$, the solution to this system is

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

from which we only use the dual directions to update the dual iterate:

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

After this step, the new point is again feasible and has reduced the barrier parameter to $\mu = (xs + rt)/2 = (6/5 + 2/5)/2 = 4/5$. The next direction from

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

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 primal and dual steps with $\alpha_p = 7/10$ and $\alpha_d = 7/2$:

$$\begin{aligned} (x, r) &= (2, 1) + (7/10)(-10/7, -10/7) = (1, 0) \\ (z, s, t) &= (2/5, 3/5, 2/5) + (7/2)(6/35, -6/35, 6/35) = (1, 0, 1). \end{aligned}$$

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.

3. Preliminaries and New Results

In this section, we first collect some of the relevant preliminaries of primal-dual path-following interior-point algorithms and then characterize four different variants of the resulting search directions in predictor or corrector steps using several known properties and some new results that we prove in this paper. To begin, let us consider an LP in standard form with n primal variables:

$$\begin{aligned} &\min\ c^T x && \max\ b^T y && \text{(5a)}\\ &\text{s.t. } Ax = b && \text{s.t. } A^T y + s = c && \text{(5b)}\\ &x \geq 0, && s \geq 0. && \text{(5c)} \end{aligned}$$

The associated (primal) logarithmic barrier formulation for problem (5) is

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

with positive barrier parameter $\mu > 0$ and first-order optimality conditions

$$Ax = b \tag{6a}$$

$$A^T y + s = c \tag{6b}$$

$$XSe = \mu e. \tag{6c}$$

We refer to these conditions as primal feasibility, dual feasibility, and complementarity, respectively. It is well-known that if problem (5) has a strictly feasible solution, then the nonlinear system (6) has a unique solution for every $\mu > 0$. The set of these solutions is the so-called central path [40]. Starting from a (strictly) feasible initial point, primal-dual path-following methods in each iteration compute a Newton direction from (6) 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 (5) as μ is reduced to zero. Specifically, in this paper we will use the short-step neighborhood

$$\mathcal{N}_2^n(\gamma) := \left\{(x, s) \in \mathbb{R}^n \times \mathbb{R}^n : (x, s) > 0,\ \|Xs - \mu e\|_2 \leq \gamma\mu \text{ for } \mu = x^T s/n\right\}.$$

To simplify notation, we also write $\|\cdot\| := \|\cdot\|_2$ without subscript to denote the canonical 2-norms for both vectors or matrices, and we frequently drop the superscript n if the dimension of x and s is clear from the context. Unlike definitions in other papers, here it is important that $\mathcal{N}_2^n(\gamma)$ only depends on the problem dimension n and the centrality parameter γ but is independent of the actual central path or the set of feasible solutions. This is convenient to clearly distinguish between centrality and feasibility and 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 which in every iteration require a (strictly) feasible iterate and generally allow to use less notation.

3.1. *Predictor and Corrector Steps*

Now let (x, y, s) be a strictly feasible primal-dual iterate for problem (5) with $Ax = b$ and $A^T y + s = c$. Set $\beta \geq 0$, $\mu = x^T s/n$, 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 Newton directions of (6) and can be computed from the following linear system:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \xi_{\beta\mu} \end{bmatrix}. \tag{7}$$

Given a direction $(\Delta x, \Delta y, \Delta z)$ at an iterate (x, y, s) with step size α, we write

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

where $Ax(\alpha) = Ax = b$ and $A^T y(\alpha) + s(\alpha) = A^T y + s = c$ because $A\Delta x = 0$ and $A^T \Delta y + \Delta s = 0$ for all directions computed from (7). We also know that

$$x(\alpha)^T s(\alpha) = (1 - \alpha(1 - \beta))x^T s \tag{8a}$$

$$X(\alpha)s(\alpha) = (1 - \alpha)Xs + \alpha\beta(x^T s/n)e + \alpha^2 \Delta X \Delta s; \tag{8b}$$

see e.g. [31, 40]. In particular, equation (8a) implies that

$$\mu(\alpha) := x(\alpha)^T s(\alpha)/n = (1 - \alpha(1 - \beta))\mu = \begin{cases} (1-\alpha)\mu & \text{if } \beta = 0; \\ \mu & \text{if } \beta = 1. \end{cases} \tag{9}$$

This also shows that even though steps are in both $x(\alpha)$ and $s(\alpha)$, the barrier parameter $\mu(\alpha)$ is a linear function of α so that specific values that achieve or not exceed certain reductions can be found relatively easily. Furthermore, for $\alpha \in [0, 1]$ and general $\beta \geq 0$, we directly find $\mu(\alpha) = \beta\mu$ if $\alpha = 1$ so that

$$\left| 1 - \frac{\mu(1)}{\mu(\alpha)} \right| = \left| 1 - \frac{\beta}{1 - \alpha(1 - \beta)} \right| = \left| \frac{(1-\alpha)(1-\beta)}{1 - \alpha + \alpha\beta} \right| \leq |1 - \beta|. \tag{10}$$

Another result in [40] provides bounds on the term $\Delta X \Delta s$ in equation (8b) and is used in our proof of Lemma 3.6 in Section 3.1.4. Note that the parameters σ and θ in the original statement are replaced by β and γ in this paper.

LEMMA 3.1 (Lemma 5.4 in [40]) *Let $\beta \geq 0$, $\gamma \in (0, 1)$, $(x, s) \in \mathcal{N}_2(\gamma)$ with $\mu = x^T s/n$, and $(\Delta x, \Delta s)$ be obtained from system* (7). *Then*

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

3.1.1. The Predictor Step: Reducing the Barrier Parameter

If we set $\beta < 1$, then we call the direction obtained from system (7) 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 (9). 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 when $\gamma = 1/4$, and is well known and widely used in the IPM literature [31, 40, among others].

LEMMA 3.2 (Lemma 4 in [31]) *Let $\beta = 0$, $\gamma = 1/4$, $(x, s) \in \mathcal{N}_2(\gamma)$, and $(\Delta x, \Delta s)$ be obtained from system (7) 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*

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

From (9), we see that this predictor step reduces the barrier parameter $\mu(\bar{\alpha}) = (1 - \bar{\alpha}) x^T s/n$ proportionally to the step length. Moreover, Lemma 3.2 implies that the new iterate $(x(\bar{\alpha}), y(\bar{\alpha}), s(\bar{\alpha}))$ still belongs to the wider neighborhood $\mathcal{N}_2(2\gamma)$ so that a single centering corrector step can be taken to restore centrality of the new iterate in the original neighborhood $\mathcal{N}_2(\gamma)$.

3.1.2. *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 (7) 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 (8a).

LEMMA 3.3 (Lemma 3 in [31] / Lemma 4.2 in [30]) *Let $\beta = 1$, $\gamma = 1/4$, $(x, s) \in \mathcal{N}_2(2\gamma)$ with $\mu = x^T s/n$, and $(\Delta x, \Delta s)$ be obtained from (7) with $\xi_{\beta\mu} = \mu e - Xs$. Then $(\bar{x}, \bar{s}) = (x, s) + (\Delta x, \Delta s) \in \mathcal{N}_2(\gamma)$ with $\bar{\mu} = \bar{x}^T \bar{s}/n = \mu$.*

Hence, given an iterate in the wider neighborhood $\mathcal{N}_2(2\gamma)$, it suffices to take a single full step into **[9.] this direction** to recenter the iterate in the narrower neighborhood $\mathcal{N}_2(\gamma)$ with no change in its barrier parameter.

3.1.3. *The Second Corrector Step: Restoring Dual Feasibility*

Let us now consider the situation where we have a dual infeasible point (x, y, s) so that $Ax = b$ but $A^T y + s \neq c$. The following analysis can be repeated similarly for primal infeasibility, but we omit this discussion because our method is always primal feasible. Let $\mu = x^T s/n$ as before 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 modified Newton system

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ \xi_c \\ 0 \end{bmatrix} \tag{11}$$

a feasibility-restoring direction. Indeed, given a step size α, it follows that

$$\xi_c(\alpha) := c - A^T y(\alpha) - s(\alpha) = (1 - \alpha)\xi_c \tag{12}$$

so that a full dual step with step size $\alpha = 1$ completely restores dual feasibility. However, it is not guaranteed that a full step also maintains centrality and non-negativity unless the initial infeasibility is sufficiently small. The following lemma is especially important for our analysis and gives a slightly extended version of the original result in [13] by including 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 3.4 (Lemma 3.3 in [13]) *Let $(x,s) \in \mathcal{N}_2(\gamma)$ with $\mu = x^T s/n$, $\delta < \sqrt{n}$, and $(\Delta y, \Delta s)$ be the dual direction obtained from system* (11). *If*

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

then 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) $(1-\delta/\sqrt{n})\mu \leq \bar{\mu} \leq (1+\delta/\sqrt{n})\mu$;
(b) $\left\|\bar{X}\bar{s} - \bar{\mu}e\right\| \leq ((1+\gamma)\delta + \gamma + \delta)\mu$.

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

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

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 $\delta \leq 1/11$. These two observations form the basis of our new Lemma 3.5.

LEMMA 3.5 (Corollary to Lemma 3.4) *Let $\gamma = 1/4$, $(x,s) \in \mathcal{N}_2(\gamma)$ with $\mu = x^T s/n$, $\delta \leq 1/11$, $\lambda = (\delta/\sqrt{n})((1+\gamma)/(1-\gamma))^{1/2}$, $\sigma = \lambda/\left\|S^{-1}\xi_c\right\|$ and $(\Delta y, \Delta s)$ be the dual direction obtained from system* (11). *Let $\bar{\alpha} \leq 1$ be the largest step size such that*

$$(x, s(\alpha)) \in \mathcal{N}_2(2\gamma) \tag{13a}$$

$$\left(1 - \frac{\delta}{\sqrt{n}}\right) x^T s \leq x^T s(\alpha) \leq \left(1 + \frac{\delta}{\sqrt{n}}\right) x^T s \tag{13b}$$

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

(a) *If $\sigma \geq 1$, then $\bar{\alpha} = 1$ and $\bar{\xi}_c = 0$.*
(b) *If $\sigma \leq 1$, then $\bar{\alpha} \geq \sigma$ and $\left\|\bar{\xi}_c\right\| \leq (1-\sigma)\|\xi_c\|$.*

Proof. If $\sigma \geq 1$, or equivalently, if $\left\|S^{-1}\xi_c\right\| \leq \lambda$, then the above observation and Lemma 3.4 imply that the full dual step $(\Delta y, \Delta s)$ satisfies the conditions in (13) and absorbs the total dual infeasibility ξ_c, so $\bar{\alpha} = 1$ and $\bar{\xi}_c = 0$. On the other hand, if $\sigma \leq 1$, then let $(\Delta\tilde{y}, \Delta\tilde{s}) = \sigma(\Delta y, \Delta s)$ correspond to the dual solution of system (11) 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 (13) 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 (12). □

We highlight that the primal iterate $\bar{x} = x$ in Lemmata 3.4 and 3.5 does not change, which is important later because this implies that no new primal inequalities will become violated when restoring dual feasibility in the second corrector step. Furthermore, because $(\bar{x}, \bar{s})$ remains in 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 $\bar{\mu}$.

3.1.4. The Third Corrector Step: Restoring the Barrier Parameter

Similar to the first corrector step, this third 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, but generally can be either smaller or larger. The next lemma gives a new, general result that is based on Lemma 3.1.

LEMMA 3.6 *Let $\gamma = 1/4$, $\delta \le 2/5$, $(x, s) \in \mathcal{N}_2(2\gamma)$ with $\mu = x^T s/n$, and $(\Delta x, \Delta s)$ be the direction from system (7) with $\xi_{\beta\mu} = \beta\mu e - Xs$. If $|1 - \beta| \le \delta/\sqrt{n}$, then $(x(\alpha), s(\alpha)) = (x, s) + \alpha(\Delta x, \Delta s) \in \mathcal{N}_2(2\gamma)$ for all $\alpha \in [0, 1]$.*

Proof. From (8), Lemma 3.1, (9), and the stated assumptions we see that

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

With $\delta \le 2/5$, it follows that $\sqrt{2}(1/4+\delta^2) \le 41\sqrt{2}/100 < 3/5$, and $\beta \ge 1-\delta/\sqrt{n} \ge 1 - \delta \ge 3/5$. Because $\alpha \ge \alpha^2$ for all $\alpha \in [0, 1]$, the above yields

$$\begin{aligned}
\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| &\le \frac{1 - \alpha + \alpha^2\sqrt{2}(1/4 + \delta^2)}{1 - \alpha(1-\beta)} \cdot \frac{\mu(\alpha)}{2} \\
&\le \frac{1 - \alpha\left(1 - \sqrt{2}(1/4 + \delta^2)\right)}{1 - \alpha(1-\beta)} \cdot \frac{\mu(\alpha)}{2} \le \frac{\mu(\alpha)}{2}
\end{aligned}$$

and thus $X(\alpha)s(\alpha) \ge (1/2)\mu(\alpha)e > 0$. This also implies that $(x(\alpha), s(\alpha)) > 0$ by continuity and thus shows that $(x(\alpha), s(\alpha)) \in \mathcal{N}_2(2\gamma)$ for all $\alpha \in [0, 1]$. $\square$

The final result in this section follows as a corollary specifically for restoring the barrier parameter after it has changed from μ to $\bar{\mu}$ in the second corrector step of Lemma 3.5, by setting $\beta = \mu/\bar{\mu}$. The upper bound of $\delta \le 2/7$ compared to $\delta \le 2/5$ in Lemma 3.6 is intentional as will become clear from the proof.

LEMMA 3.7 (Corollary to Lemma 3.6) *Let $\beta = \mu/\bar{\mu}$, $\gamma = 1/4$, $\delta \le 2/7$, $(\bar{x}, \bar{s}) \in \mathcal{N}_2(2\gamma)$ with $\bar{x}^T\bar{s} = \bar{\mu}$, and $(\Delta x, \Delta s)$ be the direction obtained from system (7) at $(\bar{x}, \bar{s})$ with $\xi_{\beta\mu} = \mu e - \bar{X}\bar{s}$. If*

$$\left(1 - \delta/\sqrt{n}\right)\mu \le \bar{\mu} \le \left(1 + \delta/\sqrt{n}\right)\mu \tag{14}$$

then $(\bar{x}(\alpha), \bar{s}(\alpha)) = (\bar{x}, \bar{s}) + \alpha(\Delta x, \Delta s) \in \mathcal{N}_2(2\gamma)$ 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 the two inequalities in (14) can equivalently be written as $\beta_1 \leq \beta \leq \beta_2$. 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

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

With $\delta \leq 2/7$, it now follows that $|1-\beta| \leq 1/((7/2)\sqrt{n}-1) \leq 1/((5/2)\sqrt{n}) = (2/5)/\sqrt{n}$ and the result follows immediately from Lemma 3.6. □

Finally, from (9) and (10) we see that a full step with $\alpha = 1$ recovers the old barrier parameter $\mu = \beta\bar{\mu}$ whereas a general step with $\alpha \in [0,1]$ still achieves a new value $\bar{\mu}(\alpha)$ whose difference to μ is no more than that between $\bar{\mu}$ and μ.

3.2. *Polynomiality of a Feasible Predictor-Corrector Method*

After the first polynomiality proofs for LP using the projective-scaling method by Karmarkar [20] and the ellipsoid method by Khachiyan [21], a variety of conceptually and notationally much simpler IPMs was proposed and subsequently proved to converge in polynomial time by Kojima et al. [22] and Mizuno [29, 30], among many others. We base the formulation in this paper on the feasible predictor-corrector algorithm described by Mizuno et al. [31] that alternates between affine-scaling predictor steps within the wider neighborhood $\mathcal{N}_2(1/2)$ and pure-centering corrector steps that recenter each iterate within $\mathcal{N}_2(1/4)$. **[11/15.] Without changing the resulting worst-case complexity bound, we apply the termination condition in Step 2 onto the barrier parameter μ rather than its (larger) duality gap $x^T s = n\mu$ to facilitate our later comparison.**

Algorithm 1 (Feasible Predictor-Corrector IPM in [31]) *Let problem* (5) *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 $\mu^1 = (x^1)^T s^1/n$, and let $\kappa > 0$ such that $\mu^1 \leq (1/\epsilon)^\kappa$. Set $k = 1$.*

Step 2 (Termination): *If $\mu^k \leq \epsilon$ 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 $(\Delta x, \Delta y, \Delta s)$ from system* (7) *at (x, s) with $\xi_{\beta\mu} = -Xs$. Let $\bar{\alpha}$ be the largest step size such that $(x(\alpha), s(\alpha)) \in \mathcal{N}_2(2\gamma)$ 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 $(\Delta\bar{x}, \Delta\bar{y}, \Delta\bar{s})$ from system* (7) *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 3.3 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

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

according to Lemma 3.2 and reduces the barrier parameter by a factor of at least $1 - \alpha^*$. From (9) together with the result of Lemma 3.3 that the barrier parameter remains unchanged during the corrector step, it follows that

$$\mu^{k+1} \leq (1 - 8^{-1/4} n^{-1/2}) \mu^k \leq (1 - 8^{-1/4} n^{-1/2})^k \mu^1 \tag{15}$$

by induction, where $\mu^1 \leq (1/\epsilon)^\kappa$. Solving $(1 - 8^{-1/4} n^{-1/2})^k (1/\epsilon)^\kappa \leq \epsilon$ for $k \geq (\kappa + 1) \log(\epsilon) / \log(1 - 8^{-1/4} n^{-1/2})$ and using the relationship that

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

for $K > 0$, we find that $\log(1 - 8^{-1/4} n^{-1/2}) = -\mathcal{O}(n^{-1/2})$. Finally combining terms, the following result is shown and corresponds to Theorem 1 in [31].

THEOREM 3.8 (Theorem 1 in [31]) *Algorithm 1 finds an ϵ-optimal solution to problem* (5) *in $\mathcal{O}\left((\kappa + 1) \log(1/\epsilon) \sqrt{n}\right)$ iterations, if it exists.*

We remark that when rewriting the original problem (1) in primal-dual standard form, the result in Theorem 3.8 provides a worst-case iteration bound of $\mathcal{O}\left((\kappa + 1) \log(1/\epsilon) \sqrt{n + \ell}\right)$. **[13.] Hence, despite the difference in ϵ and missing constant** τ, this bound depends on the number ℓ of total inequalities and thus could be much worse than the bound in Theorem 2.1, which only depends on the number l of those inequalities that need to be added at optimality. Conditions for polynomiality of our bound are given in Theorem 4.2 and Section 4.2 below.

4. New Algorithm

In this section, we formulate and analyze the new algorithm for solving LP problems in the non-standard form (1). For clarity, we refer to the full problem 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, denoted by (A, b, c, P, q). In this notation, (A, b, c) represent the data that is currently used, whereas (P, q) are the inequality data currently removed from the problem. In particular, at the beginning all inequalities are removed so that P and q are of full dimension ℓ. As the algorithm selects inequalities for inclusion, these are used to augment (A, b) and correspondingly removed from (P, q) which therefore decrease in size (compare Step 4.5 and the `Augment` function in Algorithm 3 below). The description of the algorithm is given in two parts.

(1) All feasible predictor-corrector steps are taken in Algorithm 2 which is analogous to Algorithm 1. We only ensure to keep a sufficient slack for all inequalities that are currently dropped by shortening step sizes, if necessary. We use

the counter k for the number of iterations that guarantee the minimum step length of Lemma 3.2 and thus reduce μ according to (15).

(2) To add new inequalities, we use a separate function `Augment` in Algorithm 3 that consists primarily of the three corrector steps discussed in Section 3.1. We use the counter l for the total number of calls to this function, or equivalently, for the total number of inequalities added to the problem.

In addition, we use a third counter h for the number of nested recursive calls to the function `Augment` after the third corrector step (from Step 4.5) in Algorithm 3. The counters k, l, and h are needed for our subsequent complexity discussion. To only gain a general understanding of how the algorithm works, these counters can be ignored.

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

Step 1 (Initialization): *Set* $\gamma = 1/4$, $\delta = 1/11$, *and* $\epsilon > 0$. *Let* $(x^1, y^1, s^1) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ *with* $(x^1, s^1) \in \mathcal{N}_2^n(\gamma)$ *be a strictly feasible iterate for the initial instance* $(A, b, c, P, q) \in \mathbb{R}^{m\times n} \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{\ell\times n} \times \mathbb{R}^\ell$ *in which all inequalities* $Px \geq q$ *are removed from the problem. Set* $\mu^1 = (x^1)^T s^1/n$, *and let* $\kappa > 0$ *and* $\tau > 0$ *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* $\mu^k \leq \epsilon$ *stop with*

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

Step 3 (Predictor Step): *Set* $(x, y, s, \mu) = (x^k, y^k, s^k, \mu^k)$, $\beta = 0$, $\mathrm{Pred} = \mathrm{True}$, *and compute* $(\Delta x, \Delta y, \Delta s)$ *from system* (7) *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*

$$Px(\alpha) - (1/\tau)\mu(\alpha)e \geq q \tag{16a}$$

$$\mu(\alpha) \geq \epsilon \tag{16b}$$

for all $\alpha \in [0, \bar{\alpha}]$, *and set*

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

If $\bar{\alpha}$ *was determined by one of the conditions in* (16), *set* $\mathrm{Pred} = \mathrm{False}$.

Step 3.5 (Augmentation): *If* $\bar{\alpha}$ *was decided by* (16a) *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* $(\Delta\bar{x}, \Delta\bar{y}, \Delta\bar{s})$ *from system* (7) *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*

$$P\bar{x}(\alpha) - (\bar{\mu}/\tau)e \geq q \tag{17}$$

for all $\alpha \in [0, \tilde{\alpha}]$, *and set*

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

Step 4.5 (Augmentation): *If $\tilde{\alpha} < 1$ was determined by* (17) *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 to Step 2 with*

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

We point out and explain the differences between Algorithms 1 and 2. First, the new condition (16b) implies that Algorithm 2 never reduces the barrier parameter μ below ϵ in the predictor step. We utilize this condition later in the proof of Lemma 4.4. Similarly, conditions (16a) and (17) 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, π) from (P, q) 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 refer to a specific inequality and its residual, and their relevance for our analysis will become clear later in this section.

Algorithm 3 (`Augment`$(A, b, c, P, q, x, y, s, l, h)$) *Let $(x, y, s) \in \mathbb{R}^{n+l} \times \mathbb{R}^{m+l} \times \mathbb{R}^{n+l}$ be the current iterate for a problem instance $(A, b, c, P, q) \in \mathbb{R}^{(m+l)\times(n+l)} \times \mathbb{R}^{m+l} \times \mathbb{R}^{n+l} \times \mathbb{R}^{(\ell-l)\times(n+l)} \times \mathbb{R}^{\ell-l}$ with l (slacked) 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 $(\check{P}, \check{q}) \in \mathbb{R}^{(\ell-l-1)\times(n+l)} \times \mathbb{R}^{\ell-l-1}$ be the subsystem of (P, q) without this inequality and augment instance and iterate by adding it as slacked equality constraint:*

$$\left(\hat{A}, \hat{b}, \hat{c}, \hat{P}, \hat{q}\right) := \left(\begin{bmatrix} A & 0 \\ p^T & -1 \end{bmatrix}, \begin{bmatrix} b \\ \pi \end{bmatrix}, \begin{bmatrix} c \\ 0 \end{bmatrix}, \left[\check{P}\ 0\right], \check{q}\right) \tag{18a}$$

$$(\hat{x}, \hat{y}, \hat{s}) = (\hat{x}_r, \hat{y}_z, \hat{s}_t) := \left(\begin{bmatrix} x \\ r \end{bmatrix}, \begin{bmatrix} y \\ z \end{bmatrix}, \begin{bmatrix} s \\ t \end{bmatrix}\right) = \left(\begin{bmatrix} x \\ \rho \end{bmatrix}, \begin{bmatrix} y \\ 0 \end{bmatrix}, \begin{bmatrix} s \\ \tau \end{bmatrix}\right). \tag{18b}$$

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

Step 2 (First Corrector/Recentering Step): *Set $\beta = 1$ and compute the solution $(\Delta\hat{x}, \Delta\hat{y}, \Delta\hat{s})$ of the augmented system* (7) *at $(\hat{x}, \hat{s})$ with $\xi_{\beta\mu} = \mu e - \hat{X}\hat{s}$. Let $\alpha' \le 1$ be the largest step size such that*

$$\hat{P}\hat{x}(\alpha) - (\mu/\tau)e \ge \hat{q} \tag{19}$$

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

$$\left(\hat{x}', \hat{y}', \hat{s}'\right) = \left(\hat{x}(\alpha'), \hat{y}(\alpha'), \hat{s}(\alpha')\right). \tag{20}$$

Step 2.5 (Augmentation): *If $\alpha' < 1$ 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 solution $(\Delta\hat{x}', \Delta\hat{y}', \Delta\hat{s}')$ of the augmented system (11) at $(\hat{x}', \hat{s}')$ with ξ_c as defined above. Let $\alpha'' \leq 1$ be the largest step size such that*

$$(\hat{x}', \hat{s}'(\alpha)) \in \mathcal{N}_2^{n+l}(2\gamma) \tag{21a}$$

$$\left(1 - \frac{\delta}{\sqrt{n+l}}\right)(\hat{x}')^T\hat{s}' \leq (\hat{x}')^T\hat{s}'(\alpha) \leq \left(1 + \frac{\delta}{\sqrt{n+l}}\right)(\hat{x}')^T\hat{s}' \tag{21b}$$

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

$$\left(\hat{x}'', \hat{y}'', \hat{s}''\right) = \left(\hat{x}', \hat{y}'(\alpha''), \hat{s}'(\alpha'')\right). \tag{22}$$

Step 4 (Third Corrector/Barrier Step): *Set $\hat{\mu} = (\hat{x}'')^T\hat{s}''/(n+l)$, $\beta = \mu/\hat{\mu}$, and compute the solution $(\Delta\hat{x}'', \Delta\hat{y}'', \Delta\hat{s}'')$ of the augmented system (7) with $\xi_{\beta\mu} = \mu e - \hat{X}''\hat{s}''$. Let $\alpha''' \leq 1$ be the largest step size such that*

$$\hat{P}\hat{x}''(\alpha) - (1/\tau)\hat{\mu}(\alpha)e \geq \hat{q} \tag{23}$$

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

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

Step 4.5 (Augmentation): *If $\alpha''' < 1$ increase h by 1 (nested call from Step 4.5), call Algorithm 3, decrease h by 1 (upon return), 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 and repeat; 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 $p^Tx \geq \pi$ 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 4.3 below, these choices guarantee that the augmented iterate $(\hat{x}_r, \hat{y}_z, \hat{s}_t)$ preserves centrality, the barrier parameter μ, primal feasibility, and the amounts of dual infeasibility in all existing dual constraints, if any. In particular, it follows that the only new infeasibility in the augmented residual $\hat{\xi}_c = [\xi_c, \zeta]^T$ is caused by the new dual constraint $-z + t = 0$ which has an initial residual of $\zeta = -\tau$ in the last component of $\hat{\xi}_c$ indexed by $\iota = l$. As already outlined in Section 2.1, the three subsequent corrector steps work together to successively reduce ζ to zero while maintaining centrality, the current barrier parameter, and all other residuals.

(1) The first corrector step in Step 2 corresponds to the regular centering step from Section 3.1.2. Like Step 4 in Algorithm 2, this step does not change the barrier parameter or any residuals and only serves to guarantee sufficient centrality before taking the second corrector step.
(2) The second corrector step in Step 3 corresponds to the feasibility step from Section 3.1.3 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 (21b).

(3) The third corrector step in Step 4 corresponds to the modified centering step from Section 3.1.4 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 keeps 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, whenever taking a primal step in the first and third corrector step we continue to use conditions (19) and (23) to check whether we need to add new inequalities from $(\hat{P}, \hat{q})$. In this case, the algorithm temporarily "freezes" the remaining infeasibility ζ in the component $n+\iota$ of $\hat{\xi}_c$, calls itself recursively to absorb the new infeasibility of the last added inequality first ("last-in-first-feasible"), and continues to reduce ζ only after all other residuals in those components indexed by $n+\iota+1, \ldots, n+l$ are successfully reduced to zero.

4.1. *Main Results*

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

$$(\hat{A}, \hat{b}, \hat{c}, \hat{P}, \hat{q}, \hat{x}, \hat{y}, \hat{s}, l, H) = \texttt{Augment}(A, b, c, P, q, x, y, s, j, h). \qquad (24)$$

Here $(A, b, c) \in \mathbb{R}^{(m+j)\times(n+j)} \times \mathbb{R}^{m+j} \times \mathbb{R}^{n+j}$ is the input instance with j inequalities that are already added as slacked equality constraints, $(P, q) \in \mathbb{R}^{(\ell-j)\times(n+j)} \times \mathbb{R}^{\ell-j}$ is the system of inequalities that remain dropped for the input instance (the j extra columns are all zero and correspond to the new slacks), and $(x, y, s) \in \mathbb{R}^{n+j} \times \mathbb{R}^{m+j} \times \mathbb{R}^{n+j}$ is the current iterate. Similarly, $(\hat{A}, \hat{b}, \hat{c}) \in \mathbb{R}^{(m+l)\times(n+l)} \times \mathbb{R}^{m+l} \times \mathbb{R}^{n+l}$ is the output instance with $l > j$ inequalities, $(\hat{P}, \hat{q}) \in \mathbb{R}^{(\ell-l)\times(n+l)} \times \mathbb{R}^{\ell-l}$ is the system of inequalities that remain dropped, and **[12.]** $(\hat{x}, \hat{y}, \hat{s}) \in \mathbb{R}^{n+l} \times \mathbb{R}^{m+l} \times \mathbb{R}^{n+l}$ is the new, augmented iterate. If a single inequality is added, then $l = j + 1$, but due to the possibility of recursive calls $l \geq j + 1$ in general. **[13.] Finally, the additional counter h is only changed in Step 4.5 of Algorithm 3 to keep track of the number of recursive calls after the third corrector step.** Theorem 4.1 and its proof in Subsections 4.1.1 and 4.1.2 are the main contribution of this paper.

THEOREM 4.1 *Let $\gamma = 1/4$, $\delta = 1/11$, and ϵ and τ be as in Algorithms 2 and 3. Let (x, y, s) be the an iterate for 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* **`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 call (24) in $\mathcal{O}((\tau/\epsilon)(n + l)^{1/2} e^{\delta\theta})$ iterations and $l - j - 1$ recursive calls to itself, where $\theta = h/\sqrt{n + l}$, [14.] h* ***is the maximum number of nested recursive calls to Algorithm 3 after the third corrector step****, $(\hat{x}, \hat{s}) \in \mathcal{N}_2^{n+l}(2\gamma)$, $\hat{\mu} = \hat{x}^T \hat{s}/(n + l) = \mu$, $\hat{A}\hat{x} = \hat{b}$, $\hat{P}\hat{x} - (1/\tau)\hat{\mu} \geq \hat{q}$, 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 4.1 implies that the integration of new inequalities in Algorithm 3 does 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 determined by (16a) so that we cannot guarantee the minimum step length from Lemma 3.2 and the corresponding reduction of the barrier parameter in (15). Hence, we do not increase the iteration counter k if an inequality is added after the predictor step, so that the estimate in (15) is still valid, 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 determined by (17) and although we know from (9) and Lemma 3.3 that we have not changed the barrier parameter, we cannot guarantee that the new iterate is sufficiently recentered. Because Lemma 3.7 with $\delta = 0$ implies that the new iterate 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 is a corollary of Theorems 3.8 and 4.1.

THEOREM 4.2 *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* $\mathcal{O}(((\kappa+\tau+1)/\epsilon)l(n+l)^{1/2}e^{\delta\theta})$ *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* $\mathcal{O}((\kappa+\tau+1)/\epsilon)l(n+l)^{1/2})$ *iterations.*

Note that the above statements about the combined algorithm use the weak estimate $\log(1/\epsilon) = \mathcal{O}(1/\epsilon)$ to combine the complexities from Algorithms 2 and 3. Also note that in the worst-case, h may be as large as l in which case Theorem 4.2 reduces to Theorem 2.1 as stated in Section 2.1. Nevertheless, Theorem 4.2 implies that the algorithm remains 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 4.2.

4.1.1. Proof of Theorem 4.1 when Adding a Single Inequality

The proof of Theorem 4.1 is split into two parts and starts with the analysis of Algorithm 3 when called from Step 3.5 or 4.5 of Algorithm 2 such that 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 (24), which therefore can be written as

$$(\hat{A}, \hat{b}, \hat{c}, \hat{P}, \hat{q}, \hat{x}, \hat{y}, \hat{s}, l, 0) = \texttt{Augment}(A, b, c, P, q, x, y, s, l-1, 0). \qquad (25)$$

The first result addresses the initial augmentation step.

LEMMA 4.3 *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* (18) *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* $\xi_c = c - A^Ty - s$, *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* $\xi_c = 0$ *and* $\hat{\xi}_c = [0, -\tau]^T$.

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

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

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

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

so that $(\hat{x},\hat{s}) \in \mathcal{N}_2^{n+l}(2\gamma)$. Third, using (18) we can explicitly write down the new system of primal constraints and substitute $r = \rho = p^Tx - \pi$ to find that

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

which shows that the augmented iterate remains primal feasible for the augmented instance. Fourth, for the new system of dropped inequalities $(\hat{P},\hat{q}) = \left(\begin{bmatrix}\check{P} \; 0\end{bmatrix}, \check{q}\right) \in \mathbb{R}^{(\ell-l)\times(n+l)} \times \mathbb{R}^{\ell-l}$ defined in Step 1 of Algorithm 3, we have

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

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

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

□

Lemma 4.3 shows that for an initial iterate (x,y,s) that satisfies the given assumptions, the augmented iterate $(\hat{x},\hat{y},\hat{s})$ preserves centrality, barrier parameter, primal feasibility, dual residuals of all former constraints, if any, and sufficient slack $\hat{P}\hat{x} - \hat{q} \geq \rho e$ for all those inequalities that remain dropped from the problem. This also means that $\alpha' = 0$ satisfies the step size conditions in the first corrector step so that the algorithm will continue. Furthermore, by our current assumption that we can take a full first corrector step, Lemma 3.3 implies that the next iterate $(\hat{x}',\hat{y}',\hat{s}') = (\hat{x}(1),\hat{y}(1),\hat{s}(1))$ defined in (20) 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 (19). Again, this means that $\alpha'' = 0$ satisfies the step size conditions in the subsequent second corrector step.

The next lemma gives an improved lower bound on this step size and the amount of infeasibility that can be absorbed, which depend on the termination criterion $\epsilon > 0$ in Algorithm 2 and the upper bound M from Assumption 3. Currently dealing with the addition of a single inequality only, we have written $\xi_c = [0,\zeta]^T$

with the only nonzero entry in the last component; this choice is without loss of generality, however, and can be modified to the more general ξ_c in the second corrector step of Algorithm 3 upon sorting inequalities. In that case, recall that despite the possibility of remaining nonzero residuals in other components, those are temporarily set to zero to restore feasibility for a single inequality at a time. This mechanism is addressed further in Subsection 4.1.2.

LEMMA 4.4 (Corollary to Lemma 3.5) *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 $\xi_c = [0, \zeta]^T$ 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 $\alpha'' \geq (1-\gamma^2)^{1/2}\delta\epsilon/\big(M(n+\ell)^{1/2}\tau\big)$ and $|\zeta|$ is reduced by at least $(1-\gamma^2)^{1/2}\delta\epsilon/\big(M(n+\ell)^{1/2}\big)$.*

Proof. Part (a) is clear from (12). For part (b), we first collect several bounds and then apply Lemma 3.5. Writing $(\hat{x}', \hat{y}', \hat{s}') = (\hat{x}'_r, \hat{y}'_z, \hat{s}'_t)$ for clarity, from $(\hat{x}'_r, \hat{s}'_t) \in \mathcal{N}_2^{n+l}(\gamma)$ with $\mu = \big((\hat{x}'_r)^T \hat{s}'_t\big)/(n+l) = \big((x')^T s' + r't'\big)/(n+l)$ we know that

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

where $r' \leq M$ by Assumption 3 and $\mu \geq \epsilon$ by (16b). Together, these imply

$$t' \geq (1-\gamma)\epsilon/M, \tag{26}$$

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

$$\left\|(\hat{S}'_t)^{-1}\xi_c\right\| = \left\| \begin{bmatrix} S' & 0 \\ 0 & t' \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \zeta \end{bmatrix} \right\| = \frac{|\zeta|}{t'} \leq \frac{M\tau}{(1-\gamma)\epsilon}. \tag{27}$$

Hence, using Lemma 3.5 the first statement of part (b) follows from

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

Combining the above with (26), (27), and Lemma 3.5, the second statement is shown similarly by writing

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

□

After the second corrector step, we know from (21) that the new iterate $(\hat{x}'', \hat{y}'', \hat{s}'')$ defined in (22) satisfies $(\hat{x}'', \hat{s}'') \in \mathcal{N}_2^{n+l}(2\gamma)$ and the bounds in (14) 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 $\alpha''' = 0$ satisfies the step size conditions for the third corrector step. In particular, under our assumption that a full Newton step is feasible, now equation (9) 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 3.7 and again satisfies all properties of Lemma 4.3(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 4.4. Because the initial residual is $|\zeta| = \tau$, it follows that the number of steps required to fully restore feasibility is bounded from above by

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

This completes the analysis of a single instance of Algorithm 3 and proves the following lemma as a special case of Theorem 4.1.

Lemma 4.5 *If (x, s, y) satisfies the assumptions in Lemma 4.3, then Algorithm 3 returns the output of the function call* (25) *with $j = l - 1$ and $H = h = 0$ in $\mathcal{O}((\tau/\epsilon)(n+l)^{1/2})$ iterations and with a new iterate $(\hat{x}, \hat{y}, \hat{s})$ that satisfies the properties in Lemma 4.3(a) and has a dual residual $\hat{\xi}_c = [\xi_c, 0]^T$.*

4.1.2. Proof of Theorem 4.1 when Adding Multiple Inequalities

Next, we analyze the case where we also detect and add new inequalities within Algorithm 3 after taking a primal step in either the first or third corrector step, when the step size α is determined by (19) or (23). Because both of these steps maintain iterates in $\mathcal{N}_2(2\gamma)$ and also satisfy all other assumptions of Lemma 4.3, 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 one or more remaining, only partially absorbed residuals from previously added inequalities.

In this case, Algorithm 3 restores feasibility of all such inequalities in a last-in-first-feasible fashion: whenever a new inequality is added ("last-in"), we "freeze" all remaining previous residuals by setting the corresponding entries in ξ_c to zero and only work towards feasibility of this new inequality ("first-feasible"). Similarly, whenever the `Augment` function returns to a previously added inequality with a remaining residual $|\zeta| \le \tau$, we can set $\xi_c = [0, \zeta, 0]^T$ where the trailing zero-vector corresponds to those inequalities that had been added later and whose residuals are therefore already fully restored. By sorting inequalities so that ζ occurs in the last component, like before we can use Lemma 4.4 to predict the amount by which $|\zeta|$ will be reduced. Because there is only a finite number of inequalities, this also implies that the process will eventually terminate and recursively return to every inequality that has been added to restore full dual feasibility and continue in Algorithm 2.

We now look at this basic mechanism 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 (20) 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. Nonetheless, in this case we can still repeat this step upon return from the recursive call of Algorithm 3 because equation (10) implies that the assumptions of Lemma 3.7 remain valid for the new choice of $\beta = \mu/\bar{\mu}(\alpha)$.

However, our analysis of Algorithm 3 now differs from that in Section 4.1.1 because the change in the barrier parameter may also change our estimate in Lemma 4.4. In particular, when 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 $\mu \geq \epsilon$ in Lemma 4.4 remains valid, the decrease is more critical because this estimate now must be replaced by the bound $\mu \geq (1-\delta/\sqrt{n+l})^h\epsilon$ to account for the worst case, in which it may vanish exponentially.

Lemma 4.6 (Corollary to Lemma 4.5) *Let $\theta = h/\sqrt{n+l}$. If (x, y, s) satisfies the assumptions in Lemma 4.3, then Algorithm 3 returns the output of the function call* (24) *in $\mathcal{O}((\tau/\epsilon)(n+l)^{1/2}e^{\delta\theta})$ 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 4.3(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 4.5 from Lemmata 4.3 and 4.4 if the bound $\mu \geq \epsilon$ in (26), or equivalently, the bound $1/\mu \leq 1/\epsilon$ in the rest of the proof is replaced by

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

□

To establish the remaining parts of Theorems 4.1 and 4.2, 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 $\mathcal{O}((1/\epsilon)(n+l)^{1/2})$ which is polynomial in the problem dimension.

4.2. *Polynomiality Conditions and Discussion*

We conclude the analysis of Algorithms 2 and 3 with some further observations about their possible worst-case performance. First, we are currently not able to confirm general polynomial time complexity if large numbers of inequalities must be added at iterates that are already very close to optimality. In this case, we find that barrier parameter and primal slacks may vanish relatively faster than the (dual) residual of newly added inequalities and prematurely fall below the termination tolerance ϵ. This confirms the well-known, similar observation when using IPMs for warmstarts 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 be very small and IPMs tend to jam [5]. It is interesting to note, however, that our primal iterates are always feasible so that the algorithm **[20.] may** always be stopped (in polynomial time) with a nearly-optimal solution.

Second, our analysis specifically finds 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. Without such level of detail, we decided to count *all* recursive calls using the parameter h to shorten the proofs and explanations of Theorems 4.1 and 4.2, but the deeper theoretical insight implies that the worst-case iteration complexity of the full algorithm will be polynomial under the following decreasingly restrictive assumptions:

(i) if the total number of inequalities or the total number of inequalities added is of order $\mathcal{O}(\sqrt{n})$ (stated after Theorem 2.1);
(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 4.2);
(iii) if (ii) holds only for those inequalities for which the preceding second corrector step reduces the barrier parameter toward its lower possible bound;
(iv) if (iii) holds only for those inequalities for which this lower 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 is one of our ongoing research questions.

5. Concluding Remarks

We have presented the first convergence proof and complexity analysis for an interior-point framework that solves LP problems by dynamically selecting and adding inequalities as an integral part of the algorithm. Such algorithms have been implemented and used in practice by several groups of authors; they are motivated by applications with a large number of known inequalities for which only a relatively small, yet a priori unknown, subset of the inequalities is active at optimality. In particular, this situation occurs frequently for linear (or more generally conic) relaxations of discrete optimization problems in which the inequalities are large classes of cutting planes that are available in advance.

To show convergence and analyze the complexity of this framework, we formulated a new primal-dual interior-point algorithm for solving linear programs in nonstandard form with equality and inequality constraints. The algorithm uses a primal-dual path-following predictor-corrector short-step IPM that differs from both standard IPMs and cutting-plane methods in how it handles the inequality constraints. Unlike most standard IPMs for which the primal problem is written in standard form by adding nonnegative 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 which violated inequalities to add by a cutting-plane method, which has no best answer and is typically based on some heuristic. In particular, the algorithm maintains feasibility with respect to all primal constraints throughout, and can be terminated prematurely to find feasible nearly-optimal solutions.

Our analysis proves convergence of the new algorithm to an optimal solution at which all inequalities are satisfied regardless of whether they are added to the problem or not. It thus provides a theoretical foundation for similar schemes used in practice. We also establish conditions under which the complexity of the algorithm is polynomial in the problem dimension.

While the implementation of similar algorithms has already shown encouraging results in practice [5, 7, 14, 16, 26, among others], it now becomes 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 algorithm or theoretical analysis to establish an improved complexity, either with or without 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 need for recursive corrector steps and potentially lead to a new infeasible method with a significant impact also on actual implementations in practice.

References

[1] D. S. Atkinson and P. M. Vaidya. A cutting plane algorithm for convex programming that uses analytic centers. *Math. Programming*, 69(1, Ser. B):1–43, 1995. Nondifferentiable and large-scale optimization (Geneva, 1992).

[2] G. B. Dantzig and Y. Ye. A build-up interior method for linear programming: affine scaling form. Technical report, Stanford University/The University of Iowa, USA, 1991.

[3] D. den Hertog, C. Roos, and T. Terlaky. Adding and deleting constraints in the logarithmic barrier method for LP. In *Advances in optimization and approximation*, volume 1 of *Nonconvex Optim. Appl.*, pages 166–185. Kluwer Acad. Publ., Dordrecht, 1994.

[4] A. S. El-Bakry, R. A. Tapia, and Y. Zhang. A study of indicators for identifying zero variables in interior-point methods. *SIAM Rev.*, 36(1):45–72, 1994.

[5] A. Engau. Recent progress in interior-point methods: Cutting plane methods and warm starts. In M. F. Anjos and J. B. Lasserre, editors, *Handbook of Semidefinite, Conic, and Polynomial Optimization*, volume 166 of *International Series in Operations Research & Management Science*, chapter 17, pages 471–498. Springer, 2012.

[6] A. Engau, M. F. Anjos, and A. Vannelli. On interior-point warmstarts for linear and combinatorial optimization. *SIAM J. Optim.*, 20(4):1828–1861, 2010.

[7] A. Engau, M. F. Anjos, and A. Vannelli. On handling cutting-planes in interior-point methods for solving semidefinite relaxations of binary quadratic optimization problems. *Optim. Meth. Soft.*, 27(3):539–559, 2012.

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

[9] F. Facchinei, A. Fischer, and C. Kanzow. On the identification of zero variables in an interior-point framework. *SIAM J. Optim.*, 10(4):1058–1078, 2000.

[10] J.-L. Goffin, Z.-Q. Luo, and Y. Ye. Complexity analysis of an interior cutting plane method for convex feasibility problems. *SIAM J. Optim.*, 6(3):638–652, 1996.

[11] J.-L. Goffin and J.-P. Vial. Multiple cuts in the analytic center cutting plane method. *SIAM J. Optim.*, 11(1):266–288, 2000.

[12] J. Gondzio, P. González-Brevis, and P. Munari. New developments in the primal-dual column generation technique. *European J. Oper. Res.*, 224(1):41–51, 2013.

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

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

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

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

[17] C. Helmberg and F. Rendl. A spectral bundle method for semidefinite programming. *SIAM J. Optim.*, 10(3):673–696, 2000.

[18] J. Jung, D. O'Leary, and A. Tits. Adaptive constraint reduction for convex quadratic programming. *Computational Optimization and Applications*, 51:125–157, 2012.

[19] J. A. Kaliski and Y. Ye. A short-cut potential reduction algorithm for linear programming. *Management Science*, 39(6):757–776, 1993.

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

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

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

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

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

[25] Z.-Q. Luo, C. Roos, and T. Terlaky. Complexity analysis of logarithmic barrier decomposition methods for semi-infinite linear programming. In *Proceedings of the Stieltjes Workshop on High Performance Optimization Techniques (HPOPT '96) (Delft)*, volume 29, pages 379–394, 1999.

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

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

[28] J. E. Mitchell. Cutting plane methods and subgradient methods. In M. Oskoorouchi, editor, *Tutorials in Operations Research*, chapter 2, pages 34–61. INFORMS, 2009.

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

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

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

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

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

[34] C. Oberlin and S. J. Wright. Active set identification in nonlinear programming. *SIAM J. Optim.*, 17(2):577–605, 2006.

[35] S. Park. *Matrix reduction in numerical optimization.* PhD thesis, University of Maryland, 2011.

[36] F. A. Potra and R. Sheng. A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. *SIAM J. Optim.*, 8(4):1007–1028, 1998.

[37] A. L. Tits, P.-A. Absil, and W. P. Woessner. Constraint reduction for linear programs with many inequality constraints. *SIAM J. Optim.*, 17(1):119–146, 2006.

[38] K. Tone. An active-set strategy in an interior point method for linear programming. *Math. Programming*, 59(3, Ser. A):345–360, 1993.

[39] L. B. Winternitz, S. O. Nicholls, A L. Tits, and D. P. O'Leary. A constraint-reduced variant of Mehrotra's predictor-corrector algorithm. *Comput. Optim. Appl.*, 51(3):1001–1036, 2012.

[40] S. J. Wright. *Primal-dual interior-point methods.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.