SIAM J. OPTIM.
Vol. 15, No. 2, pp. 319–347

A NEW ITERATION-COMPLEXITY BOUND FOR THE MTY PREDICTOR-CORRECTOR ALGORITHM*

RENATO D. C. MONTEIRO† AND TAKASHI TSUCHIYA‡

Abstract. In this paper we present a new iteration-complexity bound for the Mizuno–Todd–Ye predictor-corrector (MTY P-C) primal-dual interior-point algorithm for linear programming. The analysis of the paper is based on the important notion of crossover events introduced by Vavasis and Ye. For a standard form linear program $\min\{c^T x : Ax = b,\ x \geq 0\}$ with decision variable $x \in \Re^n$, we show that the MTY P-C algorithm, started from a well-centered interior-feasible solution with duality gap $n\mu_0$, finds an interior-feasible solution with duality gap less than $n\eta$ in $\mathcal{O}(T(\mu_0/\eta)+n^{3.5}\log(\bar{\chi}_A^*))$ iterations, where $T(t) \equiv \min\{n^2 \log(\log t), \log t\}$ for all $t > 0$ and $\bar{\chi}_A^*$ is a scaling invariant condition number associated with the matrix A. More specifically, $\bar{\chi}_A^*$ is the infimum of all the conditions numbers $\bar{\chi}_{AD}$, where D varies over the set of positive diagonal matrices. Under the setting of the Turing machine model, our analysis yields an $\mathcal{O}(n^{3.5}L_A + \min\{n^2 \log L, L\})$ iteration-complexity bound for the MTY P-C algorithm to find a primal-dual optimal solution, where L_A and L are the input sizes of the matrix A and the data (A, b, c), respectively. This contrasts well with the classical iteration-complexity bound for the MTY P-C algorithm, which depends linearly on L instead of $\log L$.

Key words. interior-point algorithms, primal-dual algorithms, path-following, central path, layered least squares steps, condition number, polynomial complexity, crossover events, scale-invariance, predictor-corrector, affine scaling, strongly polynomial, linear programming

AMS subject classifications. 65K05, 68Q25, 90C05, 90C51, 90C60

DOI. 10.1137/S1052623402416803

1. Introduction.

We consider the linear programming (LP) problem

$$\begin{array}{ll} \text{minimize}_x & c^T x \\ \text{subject to} & Ax = b, \quad x \geq 0, \end{array} \tag{1.1}$$

and its associated dual problem

$$\begin{array}{ll} \text{maximize}_{(y,s)} & b^T y \\ \text{subject to} & A^T y + s = c, \quad s \geq 0, \end{array} \tag{1.2}$$

where $A \in \Re^{m\times n}$, $c \in \Re^n$, and $b \in \Re^m$ are given, and the vectors $x, s \in \Re^n$ and $y \in \Re^m$ are the unknown variables.

Karmarkar in his seminal paper [6] proposed the first polynomially convergent interior-point method with an $\mathcal{O}(nL)$ iteration-complexity bound, where L is the size of the LP instance (1.1). The first path-following interior-point algorithm was proposed by Renegar in his breakthrough paper [20]. Renegar's method closely follows the primal central path and exhibits an $\mathcal{O}(\sqrt{n}L)$ iteration-complexity bound.

*Received by the editors October 31, 2002; accepted for publication (in revised form) July 2, 2004; published electronically December 30, 2004.

http://www.siam.org/journals/siopt/15-2/41680.html

†School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332 (monteiro@isye.gatech.edu). This author was supported in part by NSF grants CCR-9902010, CCR-0203113, and INT-9910084 and by ONR grant N00014-03-1-0401.

‡The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-Ku, Tokyo, 106-8569, Japan (tsuchiya@sun312.ism.ac.jp). This author was supported in part by Japan-U.S. Joint Research Projects of Japan Society for the Promotion of Science "Algorithms for linear programs over symmetric cones" and the Grants-in-Aid for Scientific Research (C) 08680478 and (C) 15510144 of the Ministry of Science, Technology, Education and Culture of Japan.

The first path-following algorithm which simultaneously generates iterates in both the primal and dual spaces was proposed by Kojima, Mizuno, and Yoshise [7] and Tanabe [22], based on ideas suggested by Megiddo [10]. In contrast to Renegar's algorithm, Kojima et al.'s algorithm has an $\mathcal{O}(nL)$ iteration-complexity bound. A primal-dual path-following algorithm with an $\mathcal{O}(\sqrt{n}L)$ iteration-complexity bound was subsequently obtained by Kojima, Mizuno, and Yoshise [8] and Monteiro and Adler [15, 16] independently. Following these developments, many other primal-dual interior-point algorithms for LP have been proposed.

An outstanding open problem in optimization is whether there exists a strongly polynomial algorithm for LP, that is, one whose complexity is bounded by a polynomial of m and n only. A major effort in this direction is due to Tardos [23], who developed a polynomial-time algorithm whose complexity is bounded by a polynomial of m, n, and L_A, where L_A denotes the size of the (integral) matrix A. Such an algorithm gives a strongly polynomial method for the important class of LP problems where the entries of A are either 1, -1, or 0, e.g., LP formulations of network flow problems. Tardos's algorithm consists of solving a sequence of LP problems of small sizes by a standard polynomially convergent LP method and using their solutions to obtain the solution of the original LP problem.

The development of a method which works entirely in the context of the original LP problem and whose complexity is also bounded by a polynomial of m, n, and L_A is due to Vavasis and Ye [30, 31]. Their method is a primal-dual path-following interior-point algorithm similar to the ones mentioned above except that from time to time it uses a crucial step, namely, the layered least squares (LLS) direction. They showed that their method has an $O(n^{3.5}\log(\bar{\chi}_A + n))$ iteration-complexity bound, where $\bar{\chi}_A$ is a condition number associated with A which has the property that $\log\bar{\chi}_A = O(L_A)$ whenever A is integral. The number $\bar{\chi}_A$ was first introduced implicitly by Dikin [1] in the study of primal affine scaling (AS) algorithms and was later studied by several researchers including Vanderbei and Lagarias [29], Todd [24], and Stewart [21]. Properties of $\bar{\chi}_A$ are studied in [4, 27, 28].

The complexity analysis of Vavasis and Ye's algorithm is based on the notion of a crossover event, a combinatorial event concerning the central path. Intuitively, a crossover event occurs between two variables when one of them is larger than the other at a point in the central path and then becomes smaller asymptotically as the optimal solution set is approached. Vavasis and Ye showed that there can be at most $n(n-1)/2$ crossover events and that a distinct crossover event occurs every $O(n^{1.5}\log(\bar{\chi}_A+n))$ iterations, from which they deduced the overall $O(n^{3.5}\log(\bar{\chi}_A+n))$ iteration-complexity bound. In [14], an LP instance is given in which the number of crossover events is $\Theta(n^2)$.

One disadvantage of Vavasis and Ye's method is that it requires explicit knowledge of $\bar{\chi}_A$ in order to determine a partition of the variables into layers used in the computation of the LLS step. This difficulty was remedied in a variant proposed by Megiddo, Mizuno, and Tsuchiya [11] which does not require explicit knowledge of the number $\bar{\chi}_A$. They observed that at most n types of partitions arise as $\bar{\chi}_A$ varies from 1 to ∞, and that one of these can be used to compute the LLS step. Based on this idea, they developed a variant which computes the LLS steps for all these partitions and picks the one that yields the greatest duality gap reduction at the current iteration. Moreover, using the argument that, once the first LLS step is computed, the other ones can be cheaply computed by performing rank-one updates, they show that the overall complexity of their algorithm is exactly the same as Vavasis and Ye's algorithm.

Another approach that also remedies the above difficulty was proposed by Monteiro and Tsuchiya [19], who developed a variant of Vavasis and Ye's algorithm which has the same complexity as theirs and computes only one LLS step per iteration without any explicit knowledge of $\bar{\chi}_A$. The method is a P-C type algorithm like the one described in [13] except that at the predictor stage it takes a step along either the primal-dual AS step or the LLS step. In contrast to the LLS step used in the algorithm of Vavasis and Ye, the partition of variables used in the algorithm of [19] for computing the LLS step is constructed from the information provided by the AS direction and hence does not require any knowledge and/or guess on $\bar{\chi}_A$.

In this paper we present a new iteration-complexity bound for the Mizuno–Todd–Ye predictor-corrector (MTY P-C) primal-dual interior-point algorithm for LP. Using the notion of crossover events and a few other nontrivial ideas, we show that the MTY P-C algorithm started from a well-centered interior-feasible solution with duality gap $n\mu_0$ finds an interior-feasible solution with duality gap less than $n\eta$ in $\mathcal{O}(T(\mu_0/\eta) + n^{3.5}\log(\bar{\chi}_A^* + n))$ iterations, where

$$T(t) \equiv \min\{n^2 \log(\log t), \log t\} \quad \forall t > 0, \tag{1.3}$$

and $\bar{\chi}_A^*$ is a scaling invariant condition number associated with the matrix A. More specifically, $\bar{\chi}_A^*$ is the infimum of all the conditions numbers $\bar{\chi}_{AD}$, where D varies over the set of positive diagonal matrices. Thus, the derived iteration-complexity bound is scaling-invariant.

The iteration-complexity bound $\mathcal{O}(T(\mu_0/\eta) + n^{3.5}\log(\bar{\chi}_A^* + n))$ has the following intuitive geometric interpretation. The term $n^{3.5}\log(\bar{\chi}_A^* + n)$ corresponds to the complexity of tracing "the curved part" of the central trajectory, that is, the part of the trajectory where the step-size along the AS direction can be as small as $\Theta(1/\sqrt{n})$. This complexity depends only on n and A and does not depend on the initial duality gap and the tolerance $\eta > 0$. The other term, $T(\mu_0/\eta) = \min\{\mathcal{O}(n^2\log(\log \mu_0/\eta)), \mathcal{O}(\log \mu_0/\eta)\}$, which corresponds to the complexity of tracing the "straight part" of the trajectory, is the minimum of two bounds. The first bound is obtained by showing that the duality gap reduces R-quadratically over $\mathcal{O}(n^2)$ disjoint sets of consecutive iterations of the algorithm. The second bound, which is independent of the dimension of the problem, is due to the fact that the duality gap along the straight part of the central trajectory is divided by a factor at least 2 at every iteration. To some extent, the second bound gives a plausible explanation to why the number of iterations performed by many interior-point algorithms grows very slowly with n.

Furthermore, using the above result, we show that the MTY algorithm endowed with a certain scaling-invariant finite termination procedure terminates in at most $\mathcal{O}(T(\mu_0/\eta^*) + n^{3.5}\log(\bar{\chi}_A^* + n))$ iterations, where η^* is a scaling invariant threshold number determined by the data (A, b, c). In particular, our results imply that the MTY P-C algorithm solves (1.1) and (1.2) in $\mathcal{O}(\min\{n^2 \log L, L\} + n^{3.5}L_A)$ iterations under the Turing machine model. This contrasts well with the classical iteration-complexity bound for the MTY P-C algorithm, namely, $\mathcal{O}(\sqrt{n}L)$, which depends linearly on the input size L of the data (A, b, c) instead of the logarithm of L.

The organization of the paper is as follows. Section 2 consists of four subsections. In subsection 2.1, we review the notion of the primal-dual central path and its associated 2-norm neighborhoods. Subsection 2.2 introduces the condition number $\bar{\chi}_A$ of a matrix A and describes the properties of $\bar{\chi}_A$ that will be useful in our analysis. Subsection 2.3 reviews the MTY P-C algorithm and states the main result of this paper, which establishes a new scaling-invariant iteration-complexity bound for the

MTY P-C algorithm to find a near primal-dual optimal solution of (1.1) and (1.2). Subsection 2.4 describes a scaling-invariant finite termination procedure and gives an alternative scaling-invariant iteration-complexity bound for the MTY P-C algorithm to find a primal-dual optimal solution of (1.1) and (1.2). Section 3, which consists of four subsections, introduces some basic tools which are used in our convergence analysis. Subsection 3.1 discusses the notion of a crossover event. Subsection 3.2 describes the notion of an LLS direction and states a proximity result that gives sufficient conditions under which the AS direction can be well approximated by an LLS direction. Subsection 3.3 reviews, from a different perspective, an important result from Vavasis and Ye [30], which basically provides sufficient conditions for the occurrence of crossover events. Subsection 3.4 describes two ordered partitions of the set of variables which are frequently used in our analysis. Section 4 is dedicated to the proof of the main result stated in subsection 2.3. Section 5 deals with a few implications of our main result under the Turing machine model. Section 6 provides some concluding remarks.

The following notation is used throughout our paper. We denote the vector of all ones by e. Its dimension is always clear from the context. The symbols $\Re^n$, $\Re^n_+$, and $\Re^n_{++}$ denote n-dimensional Euclidean space, the nonnegative orthant of $\Re^n$, and the positive orthant of $\Re^n$, respectively. The set of all $m \times n$ matrices with real entries is denoted by $\Re^{m\times n}$. If J is a finite index set, then $|J|$ denotes its cardinality, that is, the number of elements of J. For $J \subseteq \{1, \dots, n\}$ and $w \in \Re^n$, we let w_J denote the subvector $[w_i]_{i\in J}$; moreover, if E is an $m \times n$ matrix, then E_J denotes the $m \times |J|$ submatrix of E corresponding to J. For a vector $w \in \Re^n$, we let $\max(w)$ and $\min(w)$ denote the largest and the smallest component of w, respectively; $\mathrm{Diag}(w)$ denotes the diagonal matrix whose ith diagonal element is w_i for $i = 1, \dots, n$; and w^{-1} denotes the vector $[\mathrm{Diag}(w)]^{-1}e$ whenever it is well defined. For two vectors $u, v \in \Re^n$, uv denotes their Hadamard product, i.e., the vector in $\Re^n$ whose ith component is $u_i v_i$. The Euclidean norm, the 1-norm, and the ∞-norm are denoted by $\|\cdot\|$, $\|\cdot\|_1$, and $\|\cdot\|_\infty$, respectively. For a matrix E, $\mathrm{Im}(E)$ denotes the subspace generated by the columns of E, and $\mathrm{Ker}(E)$ denotes the subspace orthogonal to the rows of E. The superscript^T denotes transpose.

2. Problem and primal-dual P-C interior-point algorithms. In this section we review the MTY P-C algorithm [13] for solving the pair of LP problems (1.1) and (1.2). We also present our main convergence result, which establishes a new polynomial iteration-complexity bound for this algorithm, namely, $\mathcal{O}(T(\mu_0/\eta) + n^{3.5}\log(\bar{\chi}^*_A + n))$, where $T(\mu_0/\eta)$ is defined in (1.3) and $\bar{\chi}^*_A$ is a certain scaling-invariant condition number associated with the matrix A, $n\mu_0$ is the initial duality gap, and $n\eta$ is the required upper bound on the duality gap of the final iterate.

2.1. The problem, the central path, and its 2-norm neighborhood. In this subsection we state our assumptions and describe the primal-dual central path and its corresponding 2-norm neighborhoods.

Given $A \in \Re^{m\times n}$, $c \in \Re^n$, and $b \in \Re^m$, consider the pairs of linear programs (1.1) and (1.2), where $x \in \Re^n$ and $(y, s) \in \Re^m \times \Re^n$ are their respective variables. The set of strictly feasible solutions for these problems are

$$\mathcal{P}^{++} \equiv \{x \in \Re^n : Ax = b,\ x > 0\},$$
$$\mathcal{D}^{++} \equiv \{(y, s) \in \Re^{m\times n} : A^T y + s = c,\ s > 0\}.$$

Throughout the paper we make the following assumptions on the pair of problems

(1.1) and (1.2).

A.1. $\mathcal{P}^{++}$ and $\mathcal{D}^{++}$ are nonempty.

A.2. The rows of A are linearly independent.

Under the above assumptions, it is well known that for any $\nu > 0$ the system

$$xs = \nu e, \tag{2.1}$$

$$Ax = b, \quad x > 0, \tag{2.2}$$

$$A^T y + s = c, \quad s > 0, \tag{2.3}$$

has a unique solution (x, y, s), which we denote by $(x(\nu), y(\nu), s(\nu))$. The central path is the set consisting of all these solutions as ν varies in $(0, \infty)$. As ν converges to zero, the path $(x(\nu), y(\nu), s(\nu))$ converges to a primal-dual optimal solution (x^*, y^*, s^*) for problems (1.1) and (1.2). Given a point $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$, its duality gap and its normalized duality gap are defined as $x^T s$ and $\mu = \mu(x, s) \equiv x^T s/n$, respectively, and the point $(x(\mu), y(\mu), s(\mu))$ is said to be the central point associated with w. Note that $(x(\mu), y(\mu), s(\mu))$ also has normalized duality gap μ. We define the proximity measure of a point $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ with respect to the central path by

$$\phi(w) = \|xs/\mu - e\|.$$

Clearly, $\phi(w) = 0$ if and only if $w = (x(\mu), y(\mu), s(\mu))$, or equivalently, w coincides with its associated central point. The 2-norm neighborhood of the central path with opening $\beta > 0$ is defined as

$$\mathcal{N}(\beta) \equiv \{w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++} : \phi(w) \le \beta\}.$$

Finally, for any point $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$, we define

$$\delta(w) \equiv s^{1/2} x^{-1/2} \in \Re^n. \tag{2.4}$$

2.2. Condition number. In this subsection we define the condition number $\bar{\chi}_A$ associated with the constraint matrix A and state the properties of $\bar{\chi}_A$, which will play an important role in our analysis.

Let $\mathcal{D}$ denote the set of all positive definite $n \times n$ diagonal matrices and define

(2.5)

$$\begin{aligned}\bar{\chi}_A &\equiv \sup\{\|A^T(A\tilde{D}A^T)^{-1}A\tilde{D}\| : \tilde{D} \in \mathcal{D}\} \\ &= \sup\left\{\frac{\|A^T y\|}{\|c\|} : y = \operatorname{argmin}_{\tilde{y}\in\Re^n} \|\tilde{D}^{1/2}(A^T\tilde{y} - c)\| \text{ for some } 0 \neq c \in \Re^n \text{ and } \tilde{D} \in \mathcal{D}\right\}.\end{aligned}$$

The parameter $\bar{\chi}_A$ plays a fundamental role in the complexity analysis of algorithms for LP and LLS problems (see [30] and references therein). Its finiteness was first established by Dikin [1]. Other authors have also given alternative derivations of the finiteness of $\bar{\chi}_A$ (see, for example, Stewart [21], Todd [24], and Vanderbei and Lagarias [29]).

We summarize in the next proposition a few important facts about the parameter $\bar{\chi}_A$.

PROPOSITION 2.1. *Let $A \in \Re^{m\times n}$ with full row rank be given. Then, the following statements hold:*

(a) $\bar{\chi}_{HA} = \bar{\chi}_A$ *for any nonsingular matrix* $H \in \Re^{m\times m}$*;*

(b) $\bar{\chi}_A = \max\{\|G^{-1}A\| : G \in \mathcal{G}\}$*, where* $\mathcal{G}$ *denotes the set of all* $m \times m$ *nonsingular submatrices of* A*;*

(c) *if the entries of* A *are all integers, then* $\bar{\chi}_A$ *is bounded by* $2^{\mathcal{O}(L_A)}$*, where* L_A *is the input bit length of* A*;*

(d) $\bar{\chi}_A = \bar{\chi}_F$ *for any* $F \in \Re^{(n-m)\times n}$ *such that* $\mathrm{Ker}(A) = \mathrm{Im}(F^T)$*.*

Proof. Statement (a) readily follows from the definition (2.5). The inequality $\bar{\chi}_A \geq \max\{\|G^{-1}A\| : G \in \mathcal{G}\}$ is established in Lemma 3 of [30] while the proof of the reverse inequality is given in [24] (see also Theorem 1 of [25]). Hence, (b) holds. The proof of (c) can be found in Lemma 24 of [30]. A proof of (d) can be found in [4]. □

We now state a Hoffman-type result for a system of linear equalities whose proof can be found in Lemma 2.3 of [19].

LEMMA 2.2. *Let* $A \in \Re^{m\times n}$ *with full row rank be given and let* $(\mathcal{K}, \mathcal{L})$ *be an arbitrary bipartition of the index set* $\{1, \ldots, n\}$*. Assume that* $\bar{w} \in \Re^{|\mathcal{L}|}$ *is an arbitrary vector such that the system* $A_{\mathcal{K}} u = A_{\mathcal{L}} \bar{w}$ *is feasible. Then, this system has a feasible solution* $\bar{u}$ *such that* $\|\bar{u}\| \leq \bar{\chi}_A \|\bar{w}\|$*.*

2.3. P-C step and its properties. In this subsection we review the well-known MTY P-C algorithm [13] and its main properties. We also state the main result of this paper, which establishes a new scaling-invariant iteration-complexity bound of the MTY P-C algorithm for finding a near primal-dual optimal solution of (1.1) and (1.2).

Each iteration of the MTY P-C algorithm consists of two steps, namely, the predictor (or AS) step and the corrector (or centrality) step. The search direction used by both steps at a given point in $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ is the unique solution of the following linear system of equations:

$$\begin{aligned} S\Delta x + X\Delta s &= \sigma\mu e - xs, \\ A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0, \end{aligned} \tag{2.6}$$

where $\mu = \mu(x, s)$ and $\sigma \in \Re$ is a prespecified parameter, commonly referred to as the centrality parameter. When $\sigma = 0$, we denote the solution of (2.6) by $(\Delta x^{\mathrm{a}}, \Delta y^{\mathrm{a}}, \Delta s^{\mathrm{a}})$ and refer to it as the (primal-dual) AS direction at w; it is the direction used in the predictor step. When $\sigma = 1$, we denote the solution of (2.6) by $(\Delta x^{\mathrm{c}}, \Delta y^{\mathrm{c}}, \Delta s^{\mathrm{c}})$ and refer to it as the centrality direction at w; it is the direction used in the corrector step.

To describe an entire iteration of the MTY P-C algorithm, suppose that a constant $\beta \in (0, 1/4]$ is given. Given a point $w = (x, y, s) \in \mathcal{N}(\beta)$, this algorithm generates the next point $w^+ = (x^+, y^+, s^+) \in \mathcal{N}(\beta)$ as follows. It first moves along the direction $(\Delta x^{\mathrm{a}}, \Delta y^{\mathrm{a}}, \Delta s^{\mathrm{a}})$ until it hits the boundary of the enlarged neighborhood $\mathcal{N}(2\beta)$. More specifically, it computes the point $w^{\mathrm{a}} = (x^{\mathrm{a}}, y^{\mathrm{a}}, s^{\mathrm{a}}) \equiv (x, y, s) + \alpha_{\mathrm{a}}(\Delta x^{\mathrm{a}}, \Delta y^{\mathrm{a}}, \Delta s^{\mathrm{a}})$, where

$$\alpha_{\mathrm{a}} \equiv \sup\{\alpha \in [0, 1] : (x, y, s) + \alpha'(\Delta x^{\mathrm{a}}, \Delta y^{\mathrm{a}}, \Delta s^{\mathrm{a}}) \in \mathcal{N}(2\beta)\ \forall \alpha' \in [0, \alpha]\}. \tag{2.7}$$

Next, the point $w^+ = (x^+, y^+, s^+)$ inside the smaller neighborhood $\mathcal{N}(\beta)$ is generated by taking a unit step along the centrality direction $(\Delta x^{\mathrm{c}}, \Delta y^{\mathrm{c}}, \Delta s^{\mathrm{c}})$ at the point w^{a}, that is, $(x^+, y^+, s^+) \equiv (x^{\mathrm{a}}, y^{\mathrm{a}}, s^{\mathrm{a}}) + (\Delta x^{\mathrm{c}}, \Delta y^{\mathrm{c}}, \Delta s^{\mathrm{c}}) \in \mathcal{N}(\beta)$. Starting from a point $w^0 \in \mathcal{N}(\beta)$ and successively performing iterations as described above, the MTY P-C algorithm generates a sequence of points $\{w^k\} \subseteq \mathcal{N}(\beta)$, which converges to the primal-dual optimal face of problems (1.1) and (1.2).

The convergence analysis of the sequence $\{w^k\}$, as well as the finite termination of the MTY P-C algorithm, has been the subject of study of several papers. In what follows, we review the properties of the P-C iteration, which yields the classical polynomial iteration-complexity bound for the MTY P-C algorithm. We also discuss alternative properties of the P-C iteration, which will be used in our analysis to derive a new polynomial iteration-complexity bound for the MTY P-C algorithm. A scaling-invariant finite termination procedure for the MTY P-C algorithm and its relationship with another well-known finite termination procedure will be discussed in subsection 2.4. A detailed proof of the next two propositions can be found, for example, in [34].

PROPOSITION 2.3 (predictor step). *Suppose that* $w = (x, y, s) \in \mathcal{N}(\beta)$ *for some constant* $\beta \in (0, 1/4]$. *Let* $\Delta w^{\mathrm{a}} = (\Delta x^{\mathrm{a}}, \Delta y^{\mathrm{a}}, \Delta s^{\mathrm{a}})$ *denote the AS direction at* w *and let* α_{a} *be the step-size computed according to* (2.7). *Then the following statements hold:*

(a) *the point* $w + \alpha\Delta w^{\mathrm{a}}$ *has normalized duality gap* $\mu(\alpha) = (1-\alpha)\mu$ *for all* $\alpha \in \Re$;
(b) $\alpha_{\mathrm{a}} \geq \max\{1 - \chi/\beta, \sqrt{\beta/n}\}$, *where* $\chi \equiv \|\Delta x^{\mathrm{a}} \Delta s^{\mathrm{a}}\|/\mu$.

Proof. It is well known (see, for example, section 4.5.1 of [34]) that (a) holds, $\chi \leq n/2$, and

$$\alpha^{\mathrm{a}} \geq \frac{2}{1 + \sqrt{1 + 4\chi/\beta}}.$$

Using these two inequalities we see, after a simple verification, that (b) holds. ∎

PROPOSITION 2.4 (corrector step). *Suppose that* $w = (x, y, s) \in \mathcal{N}(2\beta)$ *for some constant* $\beta \in (0, 1/4]$ *and let* $(\Delta x^{\mathrm{c}}, \Delta y^{\mathrm{c}}, \Delta s^{\mathrm{c}})$ *denote the corrector step at* w. *Then* $w + \Delta w^{\mathrm{c}} \in \mathcal{N}(\beta)$. *Moreover, the (normalized) duality gap of* $w + \Delta w^{\mathrm{c}}$ *is the same as that of* w.

For a search direction $\Delta w = (\Delta x, \Delta y, \Delta s)$ at a point $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$, the quantity

(2.8)
$$(Rx, Rs) \equiv \left(\frac{\delta(x + \Delta x)}{\sqrt{\mu}}, \frac{\delta^{-1}(s + \Delta s)}{\sqrt{\mu}} \right) = \left(\frac{x^{1/2}s^{1/2} + \delta\Delta x}{\sqrt{\mu}}, \frac{x^{1/2}s^{1/2} + \delta^{-1}\Delta s}{\sqrt{\mu}} \right),$$

where $\delta \equiv \delta(w)$, appears quite often in our analysis. We refer to it as the *residual* of Δw. Throughout this paper, we denote the residual of the AS direction $\Delta w^{\mathrm{a}} = (\Delta x^{\mathrm{a}}, \Delta y^{\mathrm{a}}, \Delta s^{\mathrm{a}})$ at w as $(Rx^{\mathrm{a}}(w), Rs^{\mathrm{a}}(w))$. Note that if $(Rx^{\mathrm{a}}, Rs^{\mathrm{a}}) \equiv (Rx^{\mathrm{a}}(w), Rs^{\mathrm{a}}(w))$, then

$$Rx^{\mathrm{a}} = -\frac{1}{\sqrt{\mu}}\delta^{-1}\Delta s^{\mathrm{a}}, \qquad Rs^{\mathrm{a}} = -\frac{1}{\sqrt{\mu}}\delta\Delta x^{\mathrm{a}}, \tag{2.9}$$

and

$$Rx^{\mathrm{a}} + Rs^{\mathrm{a}} = \frac{x^{1/2}s^{1/2}}{\sqrt{\mu}}, \tag{2.10}$$

due to the fact that $(\Delta x^{\mathrm{a}}, \Delta y^{\mathrm{a}}, \Delta s^{\mathrm{a}})$ satisfies the first equation in (2.6) with $\sigma = 0$. The following quantity plays an important role in our analysis:

$$\varepsilon_\infty^{\mathrm{a}}(w) \equiv \max_i \{\min\{|Rx_i^{\mathrm{a}}(w)|, |Rs_i^{\mathrm{a}}(w)|\}\}. \tag{2.11}$$

In terms of this quantity, we have the following bound on the reduction on the duality gap during an iteration of the MTY P-C algorithm.

LEMMA 2.5. *Suppose that $w \in \mathcal{N}(\beta)$ for some constant $\beta \in (0, 1/4]$ and let w^+ be the point obtained after a single iteration of the MTY P-C algorithm with base point w. Then, $w^+ \in \mathcal{N}(\beta)$ and*

$$\mu(w^+) \le \min\left\{1 - \sqrt{\frac{\beta}{n}}, \frac{\sqrt{n}\varepsilon_\infty^{\mathrm{a}}(w)}{\beta}\right\}\mu(w). \tag{2.12}$$

Proof. Using (2.6) and (2.9), we easily see that $(Rx^{\mathrm{a}})^T Rs^{\mathrm{a}} = 0$. Using this observation together with (2.10), we easily see that $\max\{\|Rx^{\mathrm{a}}\|, \|Rs^{\mathrm{a}}\|\} \le \sqrt{n}$. The result now immediately follows from this conclusion, relation (2.11), and Propositions 2.3 and 2.4. □

We end this section by stating the main result of this paper. This result establishes a new iteration-complexity bound for the MTY P-C algorithm.

THEOREM 2.6. *Given a termination tolerance $\eta > 0$ for the normalized duality gap and an initial point $w^0 \in \mathcal{N}(\beta)$ with $\beta \in (0, 1/4]$, the MTY P-C algorithm generates an iterate $w^k \in \mathcal{N}(\beta)$ satisfying $\mu(w^k) \le \eta$ in at most*

$$\mathcal{O}(\min\{\sqrt{n}\log(\mu_0/\eta), T(\mu_0/\eta) + n^{3.5}\log(\bar{\chi}_A^* + n)\}) \tag{2.13}$$

iterations, where $\mu_0 \equiv \mu(w^0)$, $T(\cdot)$ is defined in (1.3), and

$$\bar{\chi}_A^* \equiv \inf\{\bar{\chi}_{AD} : D \in \mathcal{D}\}.$$

A few observations are in order at this point. First, the first bound in (2.13) is the classical one derived in [13] (see also [8, 15, 16, 17]), and follows as an immediate consequence of the first bound on the duality gap reduction obtained in Lemma 2.5. The second bound in (2.13) is the one which will be established in this paper. Observe that, in contrast to the classical iteration-complexity bound which is proportional to $\log(\mu_0/\eta)$, the new bound depends linearly on $\log(\log(\mu_0/\eta))$. Second, note that the MTY P-C algorithm is scaling-invariant; i.e., if the change of variables $(x, y, s) = (D\tilde{x}, \tilde{y}, D^{-1}\tilde{s})$ for some $D \in \mathcal{D}$ is performed on the pair of problems (1.1) and (1.2) and the MTY P-C algorithm is applied to the new dual pair of scaled problems, then the sequence of iterates $\{\tilde{w}^k\}$ generated satisfies $(x^k, y^k, s^k) = (D\tilde{x}^k, \tilde{y}^k, D^{-1}\tilde{s}^k)$ for all $k \ge 1$ as long as the initial iterate $\tilde{w}^0 \in \mathcal{N}(\beta)$ in the $\tilde{w}$-space satisfies $(x^0, y^0, s^0) = (D\tilde{x}^0, \tilde{y}^0, D^{-1}\tilde{s}^0)$. For this reason, the MTY P-C algorithm should have an iteration-complexity bound which does not depend on the scaled space where the sequence of iterates is generated. Note that the iteration-complexity bound (2.13) has this property since μ_0, $\bar{\chi}_A^*$, and the inequality $\mu(w^k) \le \eta$ are all scaling-invariant. Third, to establish the iteration-complexity bound stated in Theorem 2.6, it is sufficient to establish that, in the scaled space, the iteration-complexity bound is

$$\mathcal{O}(\min\{\sqrt{n}\log(\mu_0/\eta), T(\mu_0/\eta) + n^{3.5}\log(\bar{\chi}_{AD} + n)\}). \tag{2.14}$$

Indeed, since the number of iterations of the MTY P-C algorithm does not depend on the underlying scaled space, it follows that this number is majorized by the infimum of (2.14) over all $D \in \mathcal{D}$, i.e., by the bound (2.13). Moreover, without loss of generality, we will consider the MTY P-C algorithm applied to (1.1) and (1.2) without any scaling and will establish the iteration-complexity bound (2.14) with $D = I$.

Vavasis and Ye [30, 31] developed a primal-dual path-following interior-point algorithm which solves the LP pair (1.1) and (1.2) in $O(n^{3.5}\log(\bar{\chi}_A + n))$ iterations. In contrast to the other standard primal-dual path-following interior-point methods developed in the literature, including the MTY P-C algorithm presented in this subsection, Vavasis and Ye's algorithm uses from time to time a crucial step, namely, the LLS step, which unfortunately is not scaling-invariant. Hence, the quantity $\bar{\chi}_A$ in its complexity bound cannot be replaced by $\bar{\chi}^*_A$. For this reason, a definitive comparison of the iteration-complexity bound corresponding to Vavasis and Ye's algorithm and the one stated in Theorem 2.6 is not apparent. While the one of Theorem 2.6 contains the extra term $T(\mu_0/\eta)$, which is not present in Vavasis and Ye's iteration-complexity bound, the iteration-complexity bound of Theorem 2.6 depends on $\bar{\chi}^*_A$ instead of the larger quantity $\bar{\chi}_A$. The following simple example shows that the difference between these two quantities can be substantial. It exhibits a family of matrices A such that the ratio $\bar{\chi}^*_A/\bar{\chi}_A$ converges to 0.

Example. For $\varepsilon > 0$, define $A(\varepsilon) := (\varepsilon I \,|\, e_1)$, where I is the $m \times m$ identity matrix and e_1 is the first unit vector. It is easy to see that $\bar{\chi}_{A(\varepsilon)} = \Theta(\varepsilon^{-1})$ and that $\chi^*_{A(\varepsilon)} = O(1)$ for every $\varepsilon > 0$.

2.4. A scaling-invariant finite termination procedure. In this section, we describe a scaling-invariant finite termination procedure which, used in conjunction with the MTY P-C algorithm, allows one to find an exact primal-dual optimal solution of (1.1) and (1.2). We also derive an alternative scaling-invariant iteration-complexity bound for the MTY P-C algorithm to find an exact primal-dual optimal solution of (1.1) and (1.2).

The finite termination procedure described in this subsection is similar to the one described in Mehrotra and Ye [12] (see also [34]) except that ours uses a scaling-invariant scheme for guessing the optimal partition (B_*, N_*) associated with the pair of LP problems (1.1) and (1.2). (Recall that, by definition, $B_* \equiv \{i : x_i > 0 \text{ for some } x \in \text{opt}(1.1)\}$ and $N_* \equiv \{i : s_i > 0 \text{ for some } (y, s) \in \text{opt}(1.2)\}$, where $\text{opt}(\cdot)$ denotes the set of optimal solutions of problem $(\cdot)$.) Namely, given a point $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$, we define the AS-bipartition $(B(w), N(w))$ at w as

$$B(w) \equiv \{i : |Rs^{\mathrm{a}}_i(w)| \le |Rx^{\mathrm{a}}_i(w)|\}, \qquad N(w) \equiv \{i : |Rs^{\mathrm{a}}_i(w)| > |Rx^{\mathrm{a}}_i(w)|\}. \tag{2.15}$$

We now state the following finite termination procedure which uses the AS-bipartition for guessing the optimal partition. Note that this partition is a primal-dual symmetric variant of the indicators discussed in [2].

Finite termination (FT) procedure.

Given $\gamma \in (0,1)$ and $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ such that $xs \ge \gamma\mu(w)$.

(1) Find the AS-bipartition $(B, N) = (B(w), N(w))$ at w.

(2) Solve the following projection problems:

$$x^* \equiv \text{argmin}_{\tilde{x}}\{\|\delta(x - \tilde{x})\|^2 : A\tilde{x} = b,\ \tilde{x}_N = 0\}, \tag{2.16}$$

$$(y^*, s^*) \equiv \text{argmin}_{(\tilde{y},\tilde{s})}\{\|\delta^{-1}(s - \tilde{s})\|^2 : A^T\tilde{y} + \tilde{s} = c,\ \tilde{s}_B = 0\}, \tag{2.17}$$

where $\delta \equiv \delta(w)$.

(3) If $x^*_B > 0$ and $s^*_N > 0$, then $w^* = (x^*, y^*, s^*)$ is optimal; output w^* and declare success. Otherwise, exit the procedure and declare failure.

Our goal now will be to show that the FT procedure always finds a primal-dual optimal solution of (1.1) and (1.2) whenever $\mu(w)$ is less than a certain positive threshold constant defined in terms of additional condition measures associated with the pair of LP problems (1.1) and (1.2). We start by defining these condition measures. The first two condition measures are given by

$$\xi^P(A,b,c) \equiv \min_{i\in B_*} \max\{\bar{x}_i : \bar{x} \in \text{opt}(1.1)\},$$
$$\xi^D(A,b,c) \equiv \min_{i\in N_*} \max\{\bar{s}_i : (\bar{y},\bar{s}) \in \text{opt}(1.2)\}.$$

The third condition measure is defined as

$$\zeta(A,(B_*,N_*)) \equiv \max\{\zeta_1,\zeta_2\}, \tag{2.18}$$

where

$$\zeta_1 \equiv \max_{d_{N_*}\neq 0} \left\{ \min \left\{ \frac{\|d_{B_*}\|}{\|d_{N_*}\|} : d = (d_{B_*}, d_{N_*}) \in \text{Ker}(A) \right\} \right\},$$
$$\zeta_2 \equiv \max_{d_{B_*}\neq 0} \left\{ \min \left\{ \frac{\|d_{N_*}\|}{\|d_{B_*}\|} : d = (d_{B_*}, d_{N_*}) \in \text{Im}(A^T) \right\} \right\}.$$

Using Lemma 2.2 and Proposition 2.1(d), it is easy to see that $\zeta(A,(B_*,N_*)) \le \bar{\chi}_A$.

The following result states some well-known estimates on the size of the components of a point $w \in \mathcal{P}^{++}\times\mathcal{D}^{++}$.

LEMMA 2.7. *Let $w = (x,y,s) \in \mathcal{P}^{++}\times\mathcal{D}^{++}$ be given and define $\mu \equiv \mu(w)$, $\delta \equiv \delta(w)$, $\xi^P \equiv \xi^P(A,b,c)$, and $\xi^D \equiv \xi^D(A,b,c)$. Then*

$$\max(x_{N_*}) \le \frac{n\mu}{\xi^D}, \qquad \max(s_{B_*}) \le \frac{n\mu}{\xi^P}. \tag{2.19}$$

If we further assume that $xs \ge \gamma\mu e$ for some $\gamma \in (0,1)$, then

$$\min(x_{B_*}) \ge \frac{\gamma\xi^D}{n}, \qquad \min(s_{N_*}) \ge \frac{\gamma\xi^P}{n}, \tag{2.20}$$

$$\|\delta_{B_*}\|_\infty \|\delta_{N_*}^{-1}\|_\infty \le \frac{n^2\mu}{\gamma\xi^P\xi^D}. \tag{2.21}$$

Proof. The inequality (2.19) follows immediately from the definitions of $\xi^P(A,b,c)$ and $\xi^D(A,b,c)$ and the following identities: $x^Ts = n\mu$, $\bar{x}^T\bar{s} = 0$, and $(x-\bar{x})^T(s-\bar{s}) = 0$ for all $\bar{w} = (\bar{x},\bar{y},\bar{s}) \in \text{opt}(1.1)\times\text{opt}(1.2)$ (see Ye [33] for details). Moreover, using the assumption that $xs \ge \gamma\mu e$ and (2.19), we obtain $s_i \ge \gamma\mu/x_i \ge \gamma\xi^D/n$ for every $i \in N_*$ and $x_i \ge \gamma\mu/s_i \ge \gamma\xi^P/n$ for every $i \in B_*$, from which (2.20) follows. Inequality (2.21) follows immediately from (2.4), (2.19), and (2.20). ☐

The following result shows not only that the AS-bipartition $(B(w),N(w))$ is a correct guess for the optimal partition (B_*,N_*) but also that the FT procedure yields a primal-dual optimal solution of (1.1) and (1.2) whenever the duality gap $\mu(w)$ is less than a suitable threshold value defined in terms of the condition measures ξ and ζ introduced above, the dimension n, and the degree of centrality γ of w.

LEMMA 2.8. *Suppose that $\gamma \in (0,1)$ and $w = (x,y,s) \in \mathcal{P}^{++}\times\mathcal{D}^{++}$ are such that*

$$xs \ge \gamma\mu \quad \textit{and} \quad \mu < \frac{\gamma^{1.5}\xi^P\xi^D}{2\zeta n^{2.5}}, \tag{2.22}$$

where $\mu = \mu(w)$*,* $\zeta \equiv \zeta(A, (B_*, N_*))$*,* $\xi^P \equiv \xi^P(A, b, c)$*, and* $\xi^D \equiv \xi^D(A, b, c)$*. Then we have that*

(a) $(B(w), N(w)) = (B_*, N_*)$*;*

(b) *the FT procedure yields a strictly complementary primal-dual optimal solution, that is, a triple* $w^* = (x^*, y^*, s^*) \in \text{opt}(1.1) \times \text{opt}(1.2)$ *such that* $x^* + s^* > 0$*.*

Proof. Suppose $\gamma \in (0, 1)$ and $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ are such that (2.22) holds. We will first prove (a). Let $\Delta w^{\rm a} = (\Delta x^{\rm a}, \Delta y^{\rm a}, \Delta s^{\rm a})$ denote the AS direction at w. It is well known that

$$\max\{\|\delta \Delta x^{\rm a}\|, \|\delta^{-1} \Delta s^{\rm a}\|\} \le \sqrt{n\mu}, \tag{2.23}$$

where $\delta \equiv \delta(w)$. Moreover, it is easy to see that

$$\Delta x^{\rm a} = \text{argmin} \left\{ \tilde{s}^T \Delta x + \frac{1}{2} \|\delta \Delta x\|^2 : A \Delta x = 0 \right\}$$

for every $\tilde{s} \in c + \text{Im}(A^T)$. Now fix some $(\bar{y}, \bar{s}) \in \text{opt}(1.2)$ and let $\tilde{s} = \bar{s}$ in the above optimization problem. Splitting the variable $\Delta x^{\rm a}$ according to the partition (B_*, N_*), noting that $\bar{s}_{B_*} = 0$, and fixing the component Δx_{N_*} to be $\Delta x^{\rm a}_{N_*}$, we conclude that

$$\Delta x^{\rm a}_{B_*} = \text{argmin}_{\Delta x_{B_*}} \left\{ \frac{1}{2} \|\delta_{B_*} \Delta x_{B_*}\|^2 : A_{B_*} \Delta x_{B_*} = -A_{N_*} \Delta x^{\rm a}_{N_*} \right\}. \tag{2.24}$$

Now, let $\widetilde{\Delta x}_{B_*}$ denote the minimum norm solution of the system $A_{B_*} \Delta x_{B_*} = -A_{N_*} \Delta x^{\rm a}_{N_*}$. Using (2.21), (2.23), (2.24), and the definition of $\zeta = \zeta(A, (B_*, N_*))$, we obtain

$$\begin{aligned} \|\delta_{B_*} \Delta x^{\rm a}_{B_*}\| &\le \|\delta_{B_*} \widetilde{\Delta x}_{B_*}\| \le \|\delta_{B_*}\|_\infty \|\widetilde{\Delta x}_{B_*}\| \le \zeta \|\delta_{B_*}\|_\infty \|\Delta x^{\rm a}_{N_*}\| \\ &\le \zeta \|\delta_{B_*}\|_\infty \|\delta^{-1}_{N_*}\|_\infty \|\delta_{N_*} \Delta x^{\rm a}_{N_*}\| \le \sqrt{n\mu}\,\zeta \|\delta_{B_*}\|_\infty \|\delta^{-1}_{N_*}\|_\infty \\ &\le \frac{n^{2.5} \zeta \mu^{1.5}}{\gamma \xi^P \xi^D} < \frac{\sqrt{\gamma\mu}}{2}, \end{aligned}$$

where the last inequality is due to (2.22). The last inequality, together with (2.22), implies that for every $i \in B_*$, we have

$$\left| \frac{\delta_i^{-1} \Delta s^{\rm a}_i}{\sqrt{\mu}} \right| = \left| \frac{x_i^{1/2} s_i^{1/2}}{\sqrt{\mu}} + \frac{\delta_i \Delta x^{\rm a}_i}{\sqrt{\mu}} \right| \ge \frac{x_i^{1/2} s_i^{1/2}}{\sqrt{\mu}} - \left| \frac{\delta_i \Delta x^{\rm a}_i}{\sqrt{\mu}} \right| > \sqrt{\gamma} - \frac{\sqrt{\gamma}}{2} = \frac{\sqrt{\gamma}}{2} \ge \left| \frac{\delta_i \Delta x^{\rm a}_i}{\sqrt{\mu}} \right|,$$

or equivalently, $|Rx^{\rm a}_i(w)| > |Rs^{\rm a}_i(w)|$ for every $i \in B_*$. Hence, $B_* \subseteq B(w)$ in view of (2.15). Similarly, we can show that $N_* \subseteq N(w)$. Therefore, $(B(w), N(w)) = (B_*, N_*)$.

We now prove (b). Let (x^*, y^*, s^*) denote the point determined by (2.16) and (2.17). Observe that $x_{B_*} - x^*_{B_*}$ is a feasible solution of the system $A_{B_*} \Delta x_{B_*} = -A_{N_*} x_{N_*}$. Now let $\widehat{\Delta x}_{B_*}$ denote the minimum norm solution of this system. Using (2.21), (2.22), (2.16), and the definition of $\zeta = \zeta(A, (B_*, N_*))$, we obtain

$$\begin{aligned} &\|x^{-1}_{B_*}(x_{B_*} - x^*_{B_*})\| \\ &\quad \le \left\| x^{-1/2}_{B_*} s^{-1/2}_{B_*} \right\|_\infty \|\delta_{B_*}(x_{B_*} - x^*_{B_*})\| \le \frac{1}{\sqrt{\gamma\mu}} \|\delta_{B_*}(x_{B_*} - x^*_{B_*})\| \\ &\quad \le \frac{1}{\sqrt{\gamma\mu}} \|\delta_{B_*} \widehat{\Delta x}_{B_*}\| \le \frac{1}{\sqrt{\gamma\mu}} \|\delta_{B_*}\|_\infty \|\widehat{\Delta x}_{B_*}\| \le \frac{\zeta}{\sqrt{\gamma\mu}} \|\delta_{B_*}\|_\infty \|x_{N_*}\| \\ &\quad \le \frac{\zeta}{\sqrt{\gamma\mu}} \|\delta_{B_*}\|_\infty \|\delta^{-1}_{N_*}\|_\infty \|\delta_{N_*} x_{N_*}\| = \frac{\zeta}{\sqrt{\gamma\mu}} \|\delta_{B_*}\|_\infty \|\delta^{-1}_{N_*}\|_\infty \left\| x^{1/2}_{N_*} s^{1/2}_{N_*} \right\| \\ &\quad \le \frac{\zeta\sqrt{n}}{\sqrt{\gamma}} \|\delta_{B_*}\|_\infty \|\delta^{-1}_{N_*}\|_\infty \le \frac{n^{2.5} \zeta \mu}{\gamma^{1.5} \xi^P \xi^D} < \frac{1}{2}. \end{aligned}$$

The above inequality clearly implies that $x^*_{B_*} > 0$. In a similar way, we can prove that $s^*_{N_*} > 0$. Hence, (b) follows. $\square$

The next result gives an iteration-complexity bound for the MTY P-C algorithm, used in conjunction with the FT procedure to find a primal-dual optimal solution of (1.1) and (1.2).

THEOREM 2.9. *Suppose that $w^0 \in \mathcal{N}(\beta)$ with $\beta \in (0, 1/4]$ is given. Then, the version of the MTY P-C algorithm in which the FT procedure is invoked at every iterate w^k started from w^0 finds a primal-dual strictly complementary optimal solution w^* in at most*

$$\mathcal{O}(\min\{\sqrt{n}\log(n\mu_0/\eta_*), T(\mu_0/\eta^*) + n^{3.5}\log(\bar{\chi}^*_A + n)\}) \tag{2.25}$$

iterations, where $\mu_0 = \mu(w^0)$, the function $T(\cdot)$ is defined in (1.3)*, and*

$$\eta_* \equiv \sup\left\{\frac{\xi^P(AD, b, Dc)\xi^D(AD, b, Dc)}{\zeta(AD, (B_*, N_*))} : D \in \mathcal{D}\right\}.$$

Proof. This result follows immediately from Theorem 2.6, Lemma 2.8, and the scaling-invariance of the algorithm under consideration. $\square$

A few implications of Theorem 2.9 under the Turing machine model will be discussed in section 5. For now, we mention that the complexity bound (2.25) remains invariant not only under the transformation that changes the LP data from (A, c, b) to (AD, Dc, b) for some $D \in \mathcal{D}$, but also under the transformation that changes (A, c, b) to $(A, \alpha c, \beta b)$ for some fixed positive scalars α and β.

3. Basic tools. In this section we introduce the basic tools that will be used in the proof of Theorem 2.6. The analysis relies heavily on the notion of crossover events due to Vavasis and Ye [30]. In subsection 3.1, we give a definition of crossover event, which is slightly different than the one introduced in [30], and then discuss some of its properties. In subsection 3.2, we describe the notion of an LLS direction introduced in [30] and then state a proximity result that gives sufficient conditions under which the AS direction can be well approximated by an LLS direction. Subsection 3.3 reviews from a different perspective an important result from [30], namely, Lemma 17, that essentially guarantees the occurrence of crossover events. Since this result is stated in terms of the residual of an LLS step, the use of the proximity result of subsection 3.2 allows us to obtain a similar result stated in terms of the residual of the AS direction. In subsection 3.4, we introduce two ordered partitions of the set of variables, which play an important role in our analysis.

3.1. Crossover events. In this subsection we discuss the important notion of a crossover event developed by Vavasis and Ye [30].

DEFINITION. *For two indices $i, j \in \{1, \ldots, n\}$ and a constant $\mathcal{C} \geq 1$, a $\mathcal{C}$*-crossover event *for the pair (i, j) is said to occur on the interval $(\nu', \nu]$ if*

$$\begin{aligned} &\textit{there exists } \nu_0 \in (\nu', \nu] \textit{ such that } \frac{s_j(\nu_0)}{s_i(\nu_0)} \leq \mathcal{C}, \\ &\textit{and } \frac{s_j(\tilde{\nu})}{s_i(\tilde{\nu})} > \mathcal{C} \ \forall\ \tilde{\nu} \leq \nu'. \end{aligned} \tag{3.1}$$

Moreover, the interval $(\nu', \nu]$ is said to contain a $\mathcal{C}$-crossover event if (3.1) *holds for some pair (i, j).*

Hence, the notion of a crossover event is independent of any algorithm and is a property of the central path only. Note that in view of (2.1), condition (3.1) can be reformulated into an equivalent condition involving only the primal variable. For our purposes, we will use only (3.1).

We have the following simple but crucial result about crossover events.

PROPOSITION 3.1. *Let $\mathcal{C} \geq 1$ be a given constant. There can be at most $n(n-1)/2$ disjoint intervals of the form $(\nu', \nu]$ containing $\mathcal{C}$-crossover events.*

The notion of $\mathcal{C}$-crossover events can be used to define the notion of $\mathcal{C}$-crossover events between two iterates of the MTY P-C algorithm as follows. We say that a $\mathcal{C}$-crossover event occurs between two iterates w^k and w^l, $k < l$, generated by the MTY P-C algorithm if the interval $(\mu(w^l), \mu(w^k)]$ contains a $\mathcal{C}$-crossover event. Note that in view of Proposition 3.1, there can be at most $n(n-1)/2$ disjoint intervals of this type. We will show in the remainder of this paper that there exists a constant $\mathcal{C} \geq 1$ with the following property: for any index k, there exists an index $l > k$ with the property $l - k = \mathcal{O}(\log(\log(\mu(w^0)/\eta)) + n^{1.5}\log(\bar{\chi}_A + n))$ and, if the MTY P-C algorithm does not terminate before or at the lth iteration, then a $\mathcal{C}$-crossover event must occur between the iterates w^k and w^l. Proposition 3.1 and a simple argument then show that the MTY P-C algorithm must terminate within $\mathcal{O}(n^2\log(\log(\mu(w^0)/\eta)) + n^{3.5}\log(\bar{\chi}_A + n))$ iterations, yielding part of Theorem 2.6. The other part of Theorem 2.6 is obtained using slightly different reasoning.

3.2. LLS directions and their relationship with the AS direction. In this subsection we describe another type of direction, which plays an important role on a criterion that guarantees the occurrence of crossover events (see Lemma 3.3), namely, the LLS direction. We also state a proximity result, which describes how the AS direction can be well approximated by suitable LLS directions.

The LLS direction was first introduced by Vavasis and Ye in [30] and is one of two directions used in their algorithm. While the algorithm in this paper does not rely on this direction, its analysis relies heavily on it by means of the implications of Lemma 3.3.

Let $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ and a partition $(J_1, \ldots, J_p)$ of the index set $\{1, \ldots, n\}$ be given and define $\delta \equiv \delta(w)$. The primal LLS direction $\Delta x^{\mathrm{ll}} = (\Delta x^{\mathrm{ll}}_{J_1}, \ldots, \Delta x^{\mathrm{ll}}_{J_p})$ at w with respect to J is defined recursively according to the order $\Delta x^{\mathrm{ll}}_{J_p}, \ldots, \Delta x^{\mathrm{ll}}_{J_1}$ as follows. Assume that the components $\Delta x^{\mathrm{ll}}_{J_p}, \ldots, \Delta x^{\mathrm{ll}}_{J_{k+1}}$ have been determined. Let $\Pi_{J_k} : \Re^n \to \Re^{J_k}$ denote the projection map defined as $\Pi_{J_k}(u) = u_{J_k}$ for all $u \in \Re^n$. Then $\Delta x^{\mathrm{ll}}_{J_k} \equiv \Pi_{J_k}(L^x_k)$, where L^x_k is given by

$$
\begin{aligned}
(3.2) \qquad L^x_k &\equiv \mathrm{Argmin}_{p \in \Re^n} \{\|\delta_{J_k}(x_{J_k} + p_{J_k})\|^2 : p \in L^x_{k+1}\} \\
&= \mathrm{Argmin}_{p \in \Re^n} \{\|\delta_{J_k}(x_{J_k} + p_{J_k})\|^2 : p \in \mathrm{Ker}(A),\ p_{J_i} = \Delta x^{\mathrm{ll}}_{J_i} \\
&\qquad \forall i = k+1, \ldots, p\},
\end{aligned}
$$

with the convention that $L^x_{p+1} \equiv \mathrm{Ker}(A)$. The slack component $\Delta s^{\mathrm{ll}} = (\Delta s^{\mathrm{ll}}_{J_1}, \ldots, \Delta s^{\mathrm{ll}}_{J_p})$ of the dual LLS direction $(\Delta y^{\mathrm{ll}}, \Delta s^{\mathrm{ll}})$ at w with respect to J is defined recursively as follows. Assume that the components $\Delta s^{\mathrm{ll}}_{J_1}, \ldots, \Delta s^{\mathrm{ll}}_{J_{k-1}}$ have been determined. Then $\Delta s^{\mathrm{ll}}_{J_k} \equiv \Pi_{J_k}(L^s_k)$, where L^s_k is given by

$$
\begin{aligned}
(3.3) \qquad L^s_k &\equiv \mathrm{Argmin}_{q \in \Re^n} \{\|\delta^{-1}_{J_k}(s_{J_k} + q_{J_k})\|^2 : q \in L^s_{k-1}\} \\
&= \mathrm{Argmin}_{q \in \Re^n} \{\|\delta^{-1}_{J_k}(s_{J_k} + q_{J_k})\|^2 : q \in \mathrm{Im}(A^T),\ q_{J_i} = \Delta s^{\mathrm{ll}}_{J_i} \\
&\qquad \forall i = 1, \ldots, k-1\},
\end{aligned}
$$

with the convention that $L^s_0 \equiv \mathrm{Im}(A^T)$. Finally, once Δs^{ll} has been determined, the component Δy^{ll} is determined from the relation $A^T \Delta y^{\mathrm{ll}} + \Delta s^{\mathrm{ll}} = 0$.

It is easy to verify that the AS direction is a special LLS direction, namely, the one with respect to the only partition in which $p = 1$ (see section 5 of [30]). Clearly, the LLS direction at a given $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ depends on the partition $J = (J_1, \ldots, J_p)$ used.

A partition $J = (J_1, \ldots, J_p)$ is said to be *ordered* at a point $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ if $\max(\delta_{J_i}) \le \min(\delta_{J_{i+1}})$ for all $i = 1, \ldots, p-1$, where $\delta \equiv \delta(w)$. In this case, the gap of J, denoted by $\operatorname{gap}(J)$, is defined as

$$\operatorname{gap}(J) = \min_{1 \le i \le p-1} \left\{ \frac{\min(\delta_{J_{i+1}})}{\max(\delta_{J_i})} \right\} = \frac{1}{\max_{1 \le i \le p-1}(\|\delta_{J_i}\|_\infty \|\delta^{-1}_{J_{i+1}}\|_\infty)} \ge 1,$$

with the convention that $\operatorname{gap}(J) = \infty$ if $p = 1$.

In the remainder of this subsection, we describe how the AS direction at a given $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ can be well approximated by suitable LLS steps, a result that will be important in our convergence analysis. Another result along this direction has also been obtained by Vavasis and Ye [32]. However, our result is more general and better suited for the development of this paper.

The proximity result below can be proved using the projection decomposition techniques developed in [26]. Another proof, using instead the techniques developed in [18], has been given in [19]. The result essentially states that the larger the gap of J is, the closer the AS direction and the LLS direction with respect to J will be to each other.

Proposition 3.2. *Let $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ and an ordered partition $J = (J_1, \ldots, J_p)$ at w be given, and let $(Rx^{\mathrm{a}}, Rs^{\mathrm{a}})$ and $(Rx^{\mathrm{ll}}, Rs^{\mathrm{ll}})$, respectively, denote the residuals of the AS direction at w and of the LLS direction at w with respect to J. If $\operatorname{gap}(J) \ge 4p\bar{\chi}_A$, then*

$$\max\{\|Rx^{\mathrm{a}} - Rx^{\mathrm{ll}}\|_\infty, \|Rs^{\mathrm{a}} - Rs^{\mathrm{ll}}\|_\infty\} \le \frac{12\sqrt{n}\bar{\chi}_A}{\operatorname{gap}(J)}.$$

In view of the above result, the AS direction can be well approximated by the LLS directions with respect to ordered partitions J, which have large gaps. The LLS direction with $p = 1$, which is the AS direction, provides the perfect approximation to the AS direction itself. However, this kind of trivial approximation is not useful for us due to the need for keeping under control the "spread" of some layers J_k. For an ordered partition $J = (J_1, \ldots, J_p)$ at w, the spread of the layer J_k, denoted by $\operatorname{spr}(J_k)$, is defined as

$$\operatorname{spr}(J_k) \equiv \frac{\max(\delta_{J_k})}{\min(\delta_{J_k})} \quad \forall k = 1, \ldots, p.$$

3.3. Relation between crossover events, the AS step, and the LLS directions. In this subsection we develop some variants of Lemma 17 of Vavasis and Ye [30] which are particularly suitable for our analysis. Specifically, we develop two estimates on the number of iterations that the MTY P-C algorithm needs to perform for some crossover event to occur. While the first estimate essentially depends on the size of the residual of the LLS step and the step-size at the initial iterate, the second one, derived with the aid of Proposition 3.2, depends only on the size of the residual of the AS direction at the initial iterate.

We start by stating an immediate consequence of Lemma 17 of [30], whose proof can be found in Lemma 3.4 of Monteiro and Tsuchiya [19].

LEMMA 3.3. *Let $w = (x, y, s) \in \mathcal{N}(\beta)$ for some $\beta \in (0, 1)$ and an ordered partition $J = (J_1, \ldots, J_p)$ at w be given. Let $\delta \equiv \delta(w)$, $\mu = \mu(w)$, and $(Rx^{\mathrm{ll}}, Rs^{\mathrm{ll}})$ denote the residual of the LLS direction $(\Delta x^{\mathrm{ll}}, \Delta y^{\mathrm{ll}}, \Delta s^{\mathrm{ll}})$ at w with respect to J. Then, for any index $q \in \{1, \ldots, p\}$, any constant*

$$\mathcal{C}_q \geq (1+\beta)\,\mathrm{spr}(J_q)/(1-\beta)^2,$$

and any $\mu' \in (0, \mu)$ such that

$$\frac{\mu'}{\mu} \leq \frac{\|Rx^{\mathrm{ll}}_{J_q}\|_\infty \|Rs^{\mathrm{ll}}_{J_q}\|_\infty}{n^3 \mathcal{C}_q^2 \bar{\chi}_A^2},$$

the interval $(\mu', \mu]$ contains a $\mathcal{C}_q$-crossover event.

Using the above result, we can now derive the main result of this subsection, which provides two estimates on the number of iterations that the MTY P-C algorithm needs to perform until some crossover event occurs. This result is a slight variation of Lemma 3.5 of Monteiro and Tsuchiya [19], which is more suitable for our analysis.

LEMMA 3.4. *Suppose $\beta \in (0, 1/4]$ and that $J = (J_1, \ldots, J_p)$ is an ordered partition at w. Let $w = (x, y, s) \in \mathcal{N}(2\beta)$ be such that $\mu(w^{k+1}) \leq \mu(w) \leq \mu(w^k)$ for some iteration index k of the MTY P-C algorithm. Let $(Rx^{\mathrm{ll}}, Rs^{\mathrm{ll}})$ denote the residual of the LLS direction $(\Delta x^{\mathrm{ll}}, \Delta y^{\mathrm{ll}}, \Delta s^{\mathrm{ll}})$ at w with respect to J. Then, for every $q \in \{1, \ldots, p\}$ and every $\mathcal{C}_q \geq (1+2\beta)\,\mathrm{spr}(J_q)/(1-2\beta)^2$, the following statements hold:*

(a) *There exists an iteration index $l > k$ such that*

$$l - k = \mathcal{O}\left(\sqrt{n}\left(\log(\bar{\chi}_A + n) + \log \mathcal{C}_q + \log\left(\frac{\mu(w^{k+1})/\mu(w)}{\|Rx^{\mathrm{ll}}_{J_q}\|_\infty \|Rs^{\mathrm{ll}}_{J_q}\|_\infty}\right)\right)\right) \tag{3.4}$$

and with the property that either a $\mathcal{C}_q$-crossover event occurs between w^k and w^l or the algorithm terminates at or before the lth iteration.

(b) *If, in addition,*

$$\mathrm{gap}(J) \geq \max\left\{4n\bar{\chi}_A, \frac{24\sqrt{n}\bar{\chi}_A}{\varepsilon^{\mathrm{a}}_{J_q}}\right\}, \tag{3.5}$$

where $\varepsilon^{\mathrm{a}}_{J_q} \equiv \min\{\|Rx^{\mathrm{a}}_{J_q}\|_\infty, \|Rs^{\mathrm{a}}_{J_q}\|_\infty\}$, then the iteration index l above satisfies

$$l - k = \mathcal{O}(\sqrt{n}(\log(\bar{\chi}_A + n) + \log \mathcal{C}_q + \log[(\varepsilon^{\mathrm{a}}_{J_q})^{-1}])). \tag{3.6}$$

Proof. To prove (a), assume that the MTY P-C algorithm does not terminate at or before the lth iteration. Lemma 3.3 guarantees that the interval $(\mu(w^l), \mu(w)]$ contains a $\mathcal{C}_q$-crossover event, and hence that a $\mathcal{C}_q$-crossover event occurs between w^k and w^l whenever

$$\frac{\mu(w^l)}{\mu(w)} = \frac{\mu(w^l)}{\mu(w^{k+1})}\frac{\mu(w^{k+1})}{\mu(w)} \leq \frac{\|Rx^{\mathrm{ll}}_{J_q}\|_\infty \|Rs^{\mathrm{ll}}_{J_q}\|_\infty}{n^3 \mathcal{C}_q^2 \bar{\chi}_A^2}. \tag{3.7}$$

Since, by Lemma 2.5, $\mu(w^{j+1})/\mu(w^j) \leq 1 - \sqrt{\beta/n}$ for all $j \geq 0$, we conclude that (3.7) holds for any l satisfying

$$\log\left(\frac{\mu(w^{k+1})}{\mu(w)}\right) + (l-k-1)\log\left(1 - \sqrt{\frac{\beta}{n}}\right) \leq \log\left[\frac{\|Rx^{\mathrm{ll}}_{J_q}\|_\infty \|Rs^{\mathrm{ll}}_{J_q}\|_\infty}{n^3 \mathcal{C}_q^2 \bar{\chi}_A^2}\right].$$

Now, using the fact that $\log(1+x) < x$ for any $x > -1$, it is easy to see that (3.4) holds for the smallest l satisfying the above inequality. Hence, (a) follows.

To prove (b), it is sufficient to show that the bound in (3.4) is bounded above by the one in (3.6) when (3.5) holds. Indeed, by Proposition 3.2 and (3.5), it follows that

$$\max\{\|Rx^{\mathrm{a}} - Rx^{\mathrm{ll}}\|_\infty, \|Rs^{\mathrm{a}} - Rs^{\mathrm{ll}}\|_\infty\} \le \frac{12\sqrt{n}\bar{\chi}_A}{\mathrm{gap}(J)} \le \frac{\varepsilon^{\mathrm{a}}_{J_q}}{2}.$$

Hence, we have

$$\begin{aligned}
&\min\{\|Rx^{\mathrm{ll}}_{J_q}\|_\infty, \|Rs^{\mathrm{ll}}_{J_q}\|_\infty\} \\
&\quad \ge \min\{\|Rx^{\mathrm{a}}_{J_q}\|_\infty - \|Rx^{\mathrm{a}} - Rx^{\mathrm{ll}}\|_\infty, \|Rs^{\mathrm{a}}_{J_q}\|_\infty - \|Rs^{\mathrm{a}} - Rs^{\mathrm{ll}}\|_\infty\} \\
&\quad \ge \min\{\|Rx^{\mathrm{a}}_{J_q}\|_\infty, \|Rs^{\mathrm{a}}_{J_q}\|_\infty\} - \frac{\varepsilon^{\mathrm{a}}_{J_q}}{2} = \varepsilon^{\mathrm{a}}_{J_q} - \frac{\varepsilon^{\mathrm{a}}_{J_q}}{2} = \frac{\varepsilon^{\mathrm{a}}_{J_q}}{2}.
\end{aligned}$$

Using this estimate in (3.4) together with the fact that $\mu(w^{k+1})/\mu(w) \le 1$, we conclude that the right-hand side of (3.4) is bounded above by the right-hand side of (3.6). ◻

3.4. Two important ordered partitions. In this subsection we describe two ordered partitions which are crucial in the analysis of this paper.

The first ordered partition is due to Vavasis and Ye [30]. Given a point $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ and a parameter $\bar{g} \ge 1$, this partition, which we refer to as the *VY $\bar{g}$-partition* at w, is defined as follows. Let $(i_1, \ldots, i_n)$ be an ordering of $\{1, \ldots, n\}$ such that $\delta_{i_1} \le \cdots \le \delta_{i_n}$, where $\delta = \delta(w)$. For $k = 2, \ldots, n$, let $r_k \equiv \delta_{i_k}/\delta_{i_{k-1}}$ and define $r_1 \equiv \infty$. Let $k_1 < \cdots < k_p$ be all the indices k such that $r_k > \bar{g}$ for all $j = 1, \ldots, p$. The VY $\bar{g}$-partition J is then defined as $J = (J_1, \ldots, J_p)$, where $J_q \equiv \{i_{k_q}, i_{k_q+1}, \ldots, i_{k_{q+1}-1}\}$ for all $q = 1, \ldots, p$. (Here, by convention, $k_{p+1} \equiv n+1$.) More generally, given a subset $I \subseteq \{1, \ldots, n\}$, we can similarly define the *VY $\bar{g}$-partition* of I at w by taking an ordering $(i_1, \ldots, i_m)$ of I satisfying $\delta_{i_1} \le \cdots \le \delta_{i_m}$ where $m = |I|$, defining the ratios $r_1, \ldots, r_m$ as above, and proceeding exactly as in the construction above to obtain an ordered partition $J = (J_1, \ldots, J_p)$ of I.

It is easy to see that the following result holds for the partition J described in the previous paragraph (see section 5 of [30]).

PROPOSITION 3.5. *Given a subset $I \subseteq \{1, \ldots, n\}$, a point $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$, and a constant $\bar{g} \ge 1$, the VY $\bar{g}$-partition $J = (J_1, \ldots, J_p)$ of I at w satisfies* $\mathrm{gap}(J) > \bar{g}$ *and* $\mathrm{spr}(J_q) \le \bar{g}^{|J_q|} \le \bar{g}^n$ *for all $q = 1, \ldots, p$.*

The second ordered partition, which is used heavily in our analysis, is obtained as follows. Given a point $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$, we first use (2.15) to compute the AS-bipartition $(B, N) = (B(w), N(w))$ at w. Next, an order $(i_1, \ldots, i_n)$ of the index variables is chosen such that $\delta_{i_1} \le \cdots \le \delta_{i_n}$. Then, the first block of consecutive indices in the n-tuple $(i_1, \ldots, i_n)$ lying in the same set B or N are placed in the first layer $\mathcal{J}_1$, the next block of consecutive indices lying in the other set is placed in $\mathcal{J}_2$, and so on. As an example assume that $(i_1, i_2, i_3, i_4, i_5, i_6, i_7) \in B \times B \times N \times B \times B \times N \times N$. In this case, we have $\mathcal{J}_1 = \{i_1, i_2\}$, $\mathcal{J}_2 = \{i_3\}$, $\mathcal{J}_3 = \{i_4, i_5\}$, and $\mathcal{J}_4 = \{i_6, i_7\}$. A partition obtained according to the above construction is clearly ordered at w. We refer to it as an *ordered AS-partition* and denote it by $\mathcal{J}(w)$.

Note that an ordered AS-partition is not uniquely determined since there can be more than one n-tuple $(i_1, \ldots, i_n)$ satisfying $\delta_{i_1} \le \cdots \le \delta_{i_n}$. This situation occurs exactly when there are two or more indices i with the same value for δ_i. It can be easily seen that there exists a unique ordered AS-partition at w if and only if there do not exist $i \in B(w)$ and $j \in N(w)$ such that $\delta_i = \delta_j$. Hence, if the AS-bipartition $(B(w), N(w))$ does not have the latter property, there can be multiple ordered AS-partitions at w. In spite of this ambiguity, our analysis in this paper is valid for any chosen ordered AS-partition. So there is no need to have the notion of ordered AS-partition uniquely defined although this can be easily accomplished.

4. Convergence analysis of the MTY P-C algorithm. In this section, we provide the proof of Theorem 2.6.

We first introduce some global constants which will be used in the convergence analysis of this section. Let

$$\mathcal{C} \equiv \frac{(1+2\beta)}{(1-2\beta)^2}(2\bar{g})^n \quad \text{and} \quad \bar{g} \equiv \frac{24\bar{\chi}_A n}{\tau}, \tag{4.1}$$

where

$$\tau = \tau(\beta) \equiv \frac{\beta(1-\beta)^3(1-2\beta)^2}{4(1+\beta)^2(1+2\beta)}. \tag{4.2}$$

Note that, for a point $w \in \mathcal{P}^{++} \times \mathcal{D}^{++}$, we have in view of (2.11) and (2.15) that

$$\varepsilon_\infty^{\mathrm{a}}(w) = \max\{\|Rx_N^{\mathrm{a}}(w)\|_\infty, \|Rs_B^{\mathrm{a}}(w)\|_\infty\}, \tag{4.3}$$

where $B \equiv B(w)$ and $N \equiv N(w)$. Clearly, $\varepsilon_\infty^{\mathrm{a}}(w)$ is an upper bound on the absolute value of the small components of the residual $(Rx^{\mathrm{a}}(w), Rs^{\mathrm{a}}(w))$. The next result gives a lower bound on the absolute value of the large components of the residual $(Rx^{\mathrm{a}}(w), Rs^{\mathrm{a}}(w))$. For a proof of this result, we refer the reader to Lemma 2.6 of Monteiro and Tsuchiya [19].

LEMMA 4.1. *Suppose that $w = (x, y, s) \in \mathcal{N}(\beta)$ for some $\beta \in (0, 1)$. Then, we have*

$$\max\{|Rx_i^{\mathrm{a}}(w)|, |Rs_i^{\mathrm{a}}(w)|\} \geq \frac{\sqrt{1-\beta}}{2}$$

for all $i = 1, \ldots, n$, or equivalently, in view of the definition of $(B, N) \equiv (B(w), N(w))$, we have

$$\min\left\{\min_{i\in B}|Rx_i^{\mathrm{a}}(w)|, \min_{i\in N}|Rs_i^{\mathrm{a}}(w)|\right\} \geq \frac{\sqrt{1-\beta}}{2}.$$

Lemma 3.4 gives a good idea of part of the effort that will be undertaken in this section, namely, to find conditions under which the bounds (3.4) or (3.6) obtained in this result can be majorized by the quantity $\mathcal{O}(n\log(\bar{\chi}_A + n) + \log(\log(\mu(w^0)/\eta)))$. We will break our analysis into the following three cases:

(i) $\mathrm{gap}(\mathcal{J}(w)) \leq 2\bar{g}$ (Lemma 4.2);

(ii) $\mathrm{gap}(\mathcal{J}(w)) \geq \bar{g}$ and $\varepsilon_\infty^{\mathrm{a}}(w) \geq \tau\bar{g}/(\sqrt{n}\,\mathrm{gap}(\mathcal{J}(w)))$ (Lemma 4.5);

(iii) neither (i) nor (ii) holds (Lemma 4.11).

We will now give an outline of the approaches used to tackle each of the above three cases. In case (i), it is easy to see that at least one variable in N and one variable in B are in the same layer J_q of a VY $2\bar{g}$-partition $J = (J_1, \ldots, J_p)$ at w. We call a layer of this type a "mixed VY-layer." The proof of Lemma 4.2 essentially shows that the existence of a mixed layer implies that the quantity $\varepsilon_{J_q}^{\mathrm{a}}$ in Lemma 3.4(b) is not too small and that a $\mathcal{C}$-crossover event must occur within $\mathcal{O}(n^{1.5}\log(\bar{\chi}_A + n))$ iterations of the MTY P-C algorithm.

In case (ii), the quantity $\text{gap}(\mathcal{J}(w))$ is sufficiently large to guarantee that the AS-direction can be well approximated by an LLS direction at w with respect to a suitable ordered partition at w obtained from an ordered AS-partition at w by breaking one of its layers into smaller layers. Using the close proximity of these two directions and the fact that a step along this LLS direction followed by at most $\mathcal{O}(n^{1.5}\log(\bar{\chi}_A + n))$ regular MTY P-C steps yields a $\mathcal{C}$-crossover event, we show that case (ii) also implies that a $\mathcal{C}$-crossover event must occur within $\mathcal{O}(n^{1.5}\log(\bar{\chi}_A + n))$ iterations of the MTY P-C algorithm.

In case (iii), the proximity between the AS direction and the LLS direction mentioned above is not small enough to guarantee that the reduction of the duality gap along these two directions are of the same order of magnitude. However, in this case we manage to show that within $\mathcal{O}(\log(\log(\mu(w^0)/\eta)))$ iterations of the MTY P-C algorithm we return to a situation in which either a mixed layer arises, or case (ii) above holds, or an index changes status (moves from B to N or vice versa) from one iteration to the next. The same kind of techniques used to handle case (i) can be used to show that the latter possibility also implies that a $\mathcal{C}$-crossover event must occur within $\mathcal{O}(n^{1.5}\log(\bar{\chi}_A + n))$ iterations of the MTY P-C algorithm. Overall, we conclude that case (iii) also implies that a $\mathcal{C}$-crossover event must occur within $\mathcal{O}(n^{1.5}\log(\bar{\chi}_A + n) + \log(\log(\mu(w^0)/\eta)))$ iterations of the MTY P-C algorithm.

The first result below is a slight variation of Lemma 4.2 of [19]. It considers case (i) above, which can be handled by applying Lemma 3.4(b) with the ordered partition $J = (J_1, \ldots, J_p)$ chosen to be the VY $2\bar{g}$-partition at w.

LEMMA 4.2. *Suppose $\beta \in (0, 1/4]$ and $w = (x, y, s) \in \mathcal{N}(2\beta)$ is such that $\mu(w^{k+1}) \le \mu(w) \le \mu(w^k)$ for some iteration index k of the MTY P-C algorithm. Assume that* $\text{gap}(\mathcal{J}) \le 2\bar{g}$, *where $\mathcal{J} = (\mathcal{J}_1, \ldots, \mathcal{J}_r)$ is an ordered AS-partition at w. Then there exists an iteration index $l > k$ such that $l - k = \mathcal{O}(n^{1.5}\log(\bar{\chi}_A + n))$ and with the property that either a $\mathcal{C}$-crossover event occurs between w^k and w^l or the algorithm terminates at or before the lth iteration.*

Proof. The proof that the required iteration index l exists is based on Lemma 3.4(b). Indeed, let $J = (J_1, \ldots, J_p)$ be a VY $2\bar{g}$-partition at w. The assumption $\text{gap}(\mathcal{J}) \le 2\bar{g}$ implies the existence of two indices i, j lying in some layer J_q of J, with one in $B(w)$ and the other in $N(w)$. Without loss of generality, assume that $i \in B(w)$ and $j \in N(w)$. By Lemma 4.1, we then have $|Rx_i^{\mathrm{a}}(w)| \ge \sqrt{1-2\beta}/2$ and $|Rs_j^{\mathrm{a}}(w)| \ge \sqrt{1-2\beta}/2$. Since $i, j \in J_q$, this implies that

$$\varepsilon_{J_q}^{\mathrm{a}} \equiv \min\{\|Rx_{J_q}^{\mathrm{a}}(w)\|_\infty, \|Rs_{J_q}^{\mathrm{a}}(w)\|_\infty\} \ge \frac{\sqrt{1-2\beta}}{2}. \tag{4.4}$$

Using the above inequality, the fact that $\text{gap}(J) \ge 2\bar{g}$, and relations (4.1) and (4.2), we easily see that (3.5) holds. Since by Proposition 3.5 the spread of every layer of a VY $2\bar{g}$-partition at w is bounded above by $(2\bar{g})^n$, we conclude that $\text{spr}(J_q) \le (2\bar{g})^n$. Hence, we may set $\mathcal{C}_q = \mathcal{C} \equiv (1+2\beta)(2\bar{g})^n/(1-2\beta)^2$ in Lemma 3.4, from which it follows that

$$\log(\mathcal{C}_q) = \mathcal{O}(n\log\bar{g}) = \mathcal{O}(n\log(\bar{\chi}_A + n)), \tag{4.5}$$

where the last equality is due to (4.1). The result now follows from Lemma 3.4(b) by noting that the bound in (3.6) is $\mathcal{O}(n^{1.5}\log(\bar{\chi}_A + n))$ in view of (4.4) and (4.5). ∎

From now on we consider cases (ii) and (iii), i.e., the situation in which $\mathrm{gap}(\mathcal{J}(w)) \geq \bar{g}$. For the sake of future reference, we note that the condition $\mathrm{gap}(\mathcal{J}(w)) \geq \bar{g}$ implies

$$\frac{\tau \bar{g}}{\sqrt{n}\,\mathrm{gap}(\mathcal{J}(w))} \leq \tau \leq \frac{\sqrt{1-2\beta}}{4}, \tag{4.6}$$

due to (4.2). The following lemma is a slight variation of Lemma 4.3 of [19]. For some suitably chosen ordered partition J at w, it provides an upper bound on the right-hand side of (3.4) in terms of the residual of the LLS step with respect to $\mathcal{J}(w)$. Note that for this result we assume only that $\mathrm{gap}(\mathcal{J}(w)) \geq \bar{g}$.

LEMMA 4.3. *Suppose $\beta \in (0, 1/4]$ and $w = (x, y, s) \in \mathcal{N}(2\beta)$ is such that $\mu(w^{k+1}) \leq \mu(w) \leq \mu(w^k)$ for some iteration index k of the MTY P-C algorithm. Assume that $\mathrm{gap}(\mathcal{J}) \geq \bar{g}$, where $\mathcal{J} = (\mathcal{J}_1, \ldots, \mathcal{J}_r)$ is an ordered AS-partition at w. Let $(Rx^{\mathrm{l}}(w), Rs^{\mathrm{l}}(w))$ denote the residual of the LLS direction at w with respect to $\mathcal{J}$ and define*

$$\varepsilon^{\mathrm{l}}_{\infty}(w) \equiv \max\{\|Rx^{\mathrm{l}}_N(w)\|_\infty, \|Rs^{\mathrm{l}}_B(w)\|_\infty\}. \tag{4.7}$$

Then, there exists an iteration index $l > k$ such that

$$l - k = \mathcal{O}\left(n^{1.5} \log(\bar{\chi}_A + n) + \sqrt{n} \log\left(\frac{\mu(w^{k+1})/\mu(w)}{\varepsilon^{\mathrm{l}}_\infty(w)}\right)\right) \tag{4.8}$$

and with the property that either a $\mathcal{C}$-crossover event occurs between w^k and w^l or the algorithm terminates at or before the lth iteration.

Proof. To simplify notation, let $(Rx^{\mathrm{l}}, Rs^{\mathrm{l}}) \equiv (Rx^{\mathrm{l}}(w), Rs^{\mathrm{l}}(w))$ and $\varepsilon^{\mathrm{l}}_\infty \equiv \varepsilon^{\mathrm{l}}_\infty(w)$. Assume without loss of generality that $\varepsilon^{\mathrm{l}}_\infty = \|Rx^{\mathrm{l}}_N\|_\infty$; the case in which $\varepsilon^{\mathrm{l}}_\infty = \|Rs^{\mathrm{l}}_B\|_\infty$ can be proved similarly. Then, $\varepsilon^{\mathrm{l}}_\infty = |Rx^{\mathrm{l}}_i|$ for some $i \in N$. Let $\mathcal{J}_t$ be the layer of $\mathcal{J}$ containing the index i and note that

$$\varepsilon^{\mathrm{l}}_\infty = |Rx^{\mathrm{l}}_i| = \|Rx^{\mathrm{l}}_{\mathcal{J}_t}\|_\infty \leq \|Rx^{\mathrm{l}}_{\mathcal{J}_t}\|. \tag{4.9}$$

Now, let $\mathcal{I} = (\mathcal{I}_1, \ldots, \mathcal{I}_p)$ be the VY $\bar{g}$-partition of $\mathcal{J}_t$ at w and consider the ordered partition $\mathcal{J}'$ defined as

$$\mathcal{J}' \equiv (\mathcal{J}_1, \ldots, \mathcal{J}_{t-1}, \mathcal{I}_1, \ldots, \mathcal{I}_p, \mathcal{J}_{t+1}, \ldots, \mathcal{J}_r).$$

Let $(Rx^{\mathrm{ll}}, Rs^{\mathrm{ll}})$ denote the residual of the LLS direction at w with respect to $\mathcal{J}'$. Using the definition of the LLS step, it is easy to see that $Rx^{\mathrm{l}}_{\mathcal{J}_j} = Rx^{\mathrm{ll}}_{\mathcal{J}_j}$ for all $j = t+1, \ldots, r$. Moreover, we have $\|Rx^{\mathrm{l}}_{\mathcal{J}_t}\| \leq \|Rx^{\mathrm{ll}}_{\mathcal{J}_t}\|$ since $\|Rx^{\mathrm{l}}_{\mathcal{J}_t}\|$ is the optimal value of the least squares problem which determines the $\Delta x^{\mathrm{l}}_{\mathcal{J}_t}$-component of the LLS step with respect to $\mathcal{J}$, whereas $\|Rx^{\mathrm{ll}}_{\mathcal{J}_t}\|$ is the objective value at a certain feasible solution for the same least squares problem. Hence, for some $q \in \{1, \ldots, p\}$, we have

$$\|Rx^{\mathrm{ll}}_{\mathcal{I}_q}\|_\infty = \|Rx^{\mathrm{ll}}_{\mathcal{J}_t}\|_\infty \geq \frac{1}{\sqrt{|\mathcal{J}_t|}} \|Rx^{\mathrm{ll}}_{\mathcal{J}_t}\| \geq \frac{1}{\sqrt{n}} \|Rx^{\mathrm{ll}}_{\mathcal{J}_t}\| \geq \frac{1}{\sqrt{n}} \|Rx^{\mathrm{l}}_{\mathcal{J}_t}\|. \tag{4.10}$$

Combining (4.9) and (4.10), we then obtain

$$\|Rx^{\mathrm{ll}}_{\mathcal{I}_q}\|_\infty \geq \frac{1}{\sqrt{n}} \varepsilon^{\mathrm{l}}_\infty. \tag{4.11}$$

Let us now bound the quantity $\|Rs^{\rm ll}_{\mathcal{I}_q}\|_\infty$ from below. Using the triangle inequality for norms, Lemma 4.1 together with the fact that $\mathcal{I}_q \subseteq N$, and Proposition 3.2 together with the fact that $\mathrm{gap}(\mathcal{J}') \geq \bar{g} = 24\bar{\chi}_A n/\tau \geq 96\bar{\chi}_A\sqrt{n}$, where the last inequality is due to (4.2), we obtain

$$\|Rs^{\rm ll}_{\mathcal{I}_q}\|_\infty \geq \|Rs^{\rm a}_{\mathcal{I}_q}\|_\infty - \|Rs^{\rm ll}_{\mathcal{I}_q} - Rs^{\rm a}_{\mathcal{I}_q}\|_\infty \geq \frac{1}{4} - \frac{12\sqrt{n}\bar{\chi}_A}{\mathrm{gap}(\mathcal{J}')} \geq \frac{1}{4} - \frac{1}{8} = \frac{1}{8}, \tag{4.12}$$

where $Rs^{\rm a} \equiv Rs^{\rm a}(w)$. Also, note that by (4.1) we have

$$\log \mathcal{C} = \mathcal{O}(n \log(\bar{\chi}_A + n)). \tag{4.13}$$

The result now follows from Lemma 3.4(a) with $J = \mathcal{J}'$ and $\mathcal{C}_q = \mathcal{C}$, the observation that Proposition 3.5 and (4.1) imply that $\mathcal{C} \geq (1+2\beta)\bar{g}^n/(1-2\beta)^2 \geq (1+2\beta)\,\mathrm{spr}(\mathcal{I}_q)/(1-2\beta)^2$, and the fact that the estimates (4.11)–(4.13) imply that the bound in (3.4) is majorized by the one in (4.8). $\square$

Our goal now will be to estimate the second logarithm that appears in the iteration-complexity bound (4.8). The next result gives a condition under which $\varepsilon^{\rm a}_\infty(w) = \mathcal{O}(\varepsilon^{\rm l}_\infty(w))$.

LEMMA 4.4. *Let $w = (x, y, s) \in \mathcal{P}^{++} \times \mathcal{D}^{++}$ be given and let $\mathcal{J} = (\mathcal{J}_1, \ldots, \mathcal{J}_r)$ denote an ordered AS-partition at w. If*

$$\mathrm{gap}(\mathcal{J}) \geq \max\left\{4n\bar{\chi}_A, \frac{24\sqrt{n}\bar{\chi}_A}{\varepsilon^{\rm a}_\infty(w)}\right\}, \tag{4.14}$$

then $\varepsilon^{\rm l}_\infty(w) \geq \varepsilon^{\rm a}_\infty(w)/2$, where $\varepsilon^{\rm l}_\infty(w)$ is defined in (4.7).

Proof. Let $(Rx^{\rm a}, Rs^{\rm a})$ and $(Rx^{\rm l}, Rs^{\rm l})$, respectively, denote the residuals of the AS direction at w and the LLS direction at w with respect to $\mathcal{J}(w)$. By Proposition 3.2 and condition (4.14), we have

$$\max\{\|Rx^{\rm a} - Rx^{\rm l}\|_\infty, \|Rs^{\rm a} - Rs^{\rm l}\|_\infty\} \leq \frac{12\sqrt{n}\bar{\chi}_A}{\mathrm{gap}(\mathcal{J})} \leq \frac{\varepsilon^{\rm a}_\infty(w)}{2}.$$

Hence, we have

$$\begin{aligned}
\varepsilon^{\rm l}_\infty(w) &\equiv \max\{\|Rx^{\rm l}_N\|_\infty, \|Rs^{\rm l}_B\|_\infty\} \\
&\geq \max\{\|Rx^{\rm a}_N\|_\infty - \|Rx^{\rm a}_N - Rx^{\rm l}_N\|_\infty, \|Rs^{\rm a}_B\|_\infty - \|Rs^{\rm a}_B - Rs^{\rm l}_B\|_\infty\} \\
&\geq \max\{\|Rx^{\rm a}_N\|_\infty, \|Rs^{\rm a}_B\|_\infty\} - \frac{\varepsilon^{\rm a}_\infty(w)}{2} = \varepsilon^{\rm a}_\infty(w) - \frac{\varepsilon^{\rm a}_\infty(w)}{2} = \frac{\varepsilon^{\rm a}_\infty(w)}{2}. \qquad \square
\end{aligned}$$

We are now ready to state and prove the result which takes care of case (ii).

LEMMA 4.5. *Suppose $\beta \in (0, 1/4]$ and $w = (x, y, s) \in \mathcal{N}(2\beta)$ is such that $\mu(w^{k+1}) \leq \mu(w) \leq \mu(w^k)$ for some iteration index k of the MTY P-C algorithm. Assume that $\mathrm{gap}(\mathcal{J}) \geq \bar{g}$ and $\varepsilon^{\rm a}_\infty(w) \geq \tau\bar{g}/(\sqrt{n}\,\mathrm{gap}(\mathcal{J}))$, where $\mathcal{J} = (\mathcal{J}_1, \ldots, \mathcal{J}_r)$ is an ordered AS-partition at w. Then, there exists an iteration index $l > k$ such that $l - k = \mathcal{O}(n^{1.5}\log(\bar{\chi}_A + n) + \sqrt{n}\log(\varepsilon^{\rm a}_\infty(w))^{-1})$ and with the property that either a $\mathcal{C}$-crossover event occurs between w^k and w^l or the algorithm terminates at or before the lth iteration. Furthermore, if $w = w^k$, then $l - k = \mathcal{O}(n^{1.5}\log(\bar{\chi}_A + n))$.*

Proof. The proof is based on Lemma 4.3 and essentially consists of showing that the term inside the second logarithm that appears in the bound (4.8) can always be bounded by $\mathcal{O}((\varepsilon^{\rm a}_\infty(w))^{-1})$ and that, when $w = w^k$, this term can also be bounded by $\mathcal{O}(\sqrt{n})$. Indeed, first note that the conditions $\mathrm{gap}(\mathcal{J}) \geq \bar{g}$ and $\varepsilon^{\mathrm{a}}_\infty(w) \geq \tau\bar{g}/\sqrt{n}\,\mathrm{gap}(\mathcal{J})$ and relations (4.1) and (4.2) clearly imply (4.14). Hence, by Lemma 4.4, it follows that $\varepsilon^{\mathrm{l}}_\infty(w) \geq \varepsilon^{\mathrm{a}}_\infty(w)/2$. Thus, we have

$$\frac{\mu(w^{k+1})/\mu(w)}{\varepsilon^{\mathrm{l}}_\infty(w)} \leq \frac{2\mu(w^{k+1})/\mu(w)}{\varepsilon^{\mathrm{a}}_\infty(w)} \leq \frac{2}{\varepsilon^{\mathrm{a}}_\infty(w)} = \mathcal{O}((\varepsilon^{\mathrm{a}}_\infty(w))^{-1}),$$

where in the second inequality we used the fact that $\mu(w) \geq \mu(w^{k+1})$. When $w = w^k$, it follows from the first inequality above and Lemma 2.5 that

$$\frac{\mu(w^{k+1})/\mu(w)}{\varepsilon^{\mathrm{l}}_\infty(w)} \leq \frac{2\mu(w^{k+1})/\mu(w)}{\varepsilon^{\mathrm{a}}_\infty(w)} \leq \frac{2\sqrt{n}}{\beta} = \mathcal{O}(\sqrt{n}).$$

The result now follows from Lemma 4.3 and the bounds obtained above. $\square$

From now on, we will consider the case in which an iterate $w = w^k$ of the MTY P-C algorithm satisfies $\mathrm{gap}(\mathcal{J}(w)) \geq \bar{g}$ and $\varepsilon^{\mathrm{a}}_\infty(w) \leq \tau\bar{g}/(\sqrt{n}\,\mathrm{gap}(\mathcal{J}(w)))$. Before tackling this case, we first need to establish two technical lemmas, namely, Lemmas 4.6 and 4.9, which give other sufficient conditions for the occurrence of $\mathcal{C}$-crossover events. Lemma 4.7 is used only in the proof of Lemma 4.8, which in turn is used only in the proof of Lemma 4.9.

LEMMA 4.6. *Suppose that $\beta \in (0, 1/4]$ and that $w = w^k \in \mathcal{N}(\beta)$ is an iterate of the MTY P-C algorithm and let $w^+ \equiv w^{k+1}$. Assume that either one of the following conditions holds:*

(a) *either $B(w) \cap N(w^+) \neq \emptyset$ or $N(w) \cap B(w^+) \neq \emptyset$;*

(b) *there exist indices i and j, one lying in $B(w)$ and the other in $N(w)$, such that $\delta_i(w)/\delta_j(w) \geq \bar{g}$ and $\delta_i(w^+)/\delta_j(w^+) \leq \bar{g}$.*

Then, there exists an iteration index $l > k$ such that $l - k = \mathcal{O}(n^{1.5}\log(\bar{\chi}_A + n))$ and with the property that either a $\mathcal{C}$-crossover event occurs between w^k and w^l or the algorithm terminates at or before the lth iteration.

Proof. Let $\mathcal{J} \equiv \mathcal{J}(w)$. If either one of the conditions $\mathrm{gap}(\mathcal{J}) \geq \bar{g}$ or $\varepsilon^{\mathrm{a}}_\infty(w) \leq \tau\bar{g}/(\sqrt{n}\,\mathrm{gap}(\mathcal{J}))$ does not hold, then the conclusion of the lemma follows immediately from Lemmas 4.2 and 4.5, regardless of whether one of conditions (a) or (b) holds. Assume then that $\mathrm{gap}(\mathcal{J}) \geq \bar{g}$ and $\varepsilon^{\mathrm{a}}_\infty(w) \leq \tau\bar{g}/(\sqrt{n}\,\mathrm{gap}(\mathcal{J}))$. This together with (4.6) imply that $\varepsilon^{\mathrm{a}}_\infty(w) \leq \sqrt{1-2\beta}/4$. We will first show that the conclusion of the lemma holds under condition (a). Let $w : [0, 1] \to \mathcal{N}(2\beta)$ be a continuous path such that $w(0) = w$, $w(1) = w^+$, and $\mu(w^+) \leq \mu(w(t)) \leq \mu(w)$ for all $t \in [0, 1]$; e.g., consider the path that traces the line segment from w to $w + \alpha_a\Delta w^{\mathrm{a}}$ and then the line segment from $w + \alpha_a\Delta w^{\mathrm{a}}$ to w^+. (It is straightforward to verify that these two segments lie in $\mathcal{N}(2\beta)$.) We will show more generally that if there exists $0 < t \leq 1$ such that either $B(w)\cap N(w(t)) \neq \emptyset$ or $N(w)\cap B(w(t)) \neq \emptyset$ (for $t = 1$ this is condition (a)), then the conclusion of the lemma follows. Indeed, assume that for some $0 < t \leq 1$ there exists an index j in the set $B(w) \cap N(w(t))$. (The proof is similar for the case in which $j \in N(w) \cap B(w(t))$.) Since $j \in B(w)$ and $\varepsilon^{\mathrm{a}}_\infty(w) \leq \sqrt{1-2\beta}/4$, we have $|Rs^{\mathrm{a}}_j(w(0))| = |Rs^{\mathrm{a}}_j(w)| \leq \varepsilon^{\mathrm{a}}_\infty(w) \leq \sqrt{1-2\beta}/4$. Moreover, since $j \in N(w(t))$, we have $|Rs^{\mathrm{a}}_j(w(t))| \geq \sqrt{1-2\beta}/2$ in view of Lemma 4.1. The intermediate value theorem applied to the continuous function $t \to |Rs^{\mathrm{a}}_j(w(t))|$ implies the existence of some $\bar{t} \in [0, t]$ such that $|Rs^{\mathrm{a}}_j(w(\bar{t}))| = \sqrt{1-2\beta}/4$. Letting $\bar{w} \equiv w(\bar{t})$, we have $\bar{w} \in \mathcal{N}(2\beta)$ and $|Rs^{\mathrm{a}}_j(\bar{w})| = \sqrt{1-2\beta}/4$. By Lemma 4.1 with β replaced by 2β we have that $\max\{|Rx^{\mathrm{a}}_j(\bar{w})|, |Rs^{\mathrm{a}}_j(\bar{w})|\} \geq \sqrt{1-2\beta}/2$. Since $|Rs^{\mathrm{a}}_j(\bar{w})| = \sqrt{1-2\beta}/4 < \sqrt{1-2\beta}/2$, we must have $|Rx^{\mathrm{a}}_j(\bar{w})| \geq \sqrt{1-2\beta}/2$. We thus proved that $\varepsilon^{\mathrm{a}}_\infty(\bar{w}) \geq$

$\min\{|Rx^{\rm a}_j(\bar{w})|, |Rs^{\rm a}_j(\bar{w})|\} \geq \sqrt{1-2\beta}/4$. If $\operatorname{gap}(\mathcal{J}(\bar{w})) \geq \bar{g}$, then by (4.6) and the fact that $\varepsilon^{\rm a}_\infty(\bar{w}) \geq \sqrt{1-2\beta}/4$, the conclusion of the lemma follows from Lemma 4.5 with $w = \bar{w}$. If, on the other hand, $\operatorname{gap}(\mathcal{J}(\bar{w})) < \bar{g}$, then the conclusion of the lemma follows from Lemma 4.2 with $w = \bar{w}$.

We will now show that the conclusion of the lemma holds under condition (b). Without loss of generality, assume that $i \in B(w)$ and $j \in N(w)$ is such that $\delta_i(w)/\delta_j(w) \geq \bar{g}$ and $\delta_i(w^+)/\delta_j(w^+) \leq \bar{g}$. The intermediate value theorem applied to the continuous function $t \to \delta_i(w(t))/\delta_j(w(t))$ implies the existence of some $\bar{t} \in [0,1]$ such that $\delta_i(w(\bar{t}))/\delta_j(w(\bar{t})) = \bar{g}$. If either $i \in N(w(\bar{t}))$ or $j \in B(w(\bar{t}))$, then the conclusion of the lemma holds in view of what we have already shown in the previous paragraph. Consider now the case in which $i \in B(w(\bar{t}))$ and $j \in N(w(\bar{t}))$. Since $\delta_i(w(\bar{t}))/\delta_j(w(\bar{t})) = \bar{g}$, this case implies that $\operatorname{gap}(\mathcal{J}(w(\bar{t}))) \leq \bar{g}$. Thus, the conclusion of the lemma follows from Lemma 4.2 with $w = w(\bar{t})$. ∎

It is worth noting that Lemma 4.6(a) gives a simple and intuitive scaling-invariant sufficient condition for a crossover event to occur within a reasonable number of iterations, namely, that an index changes its status from the iterate w to the next iterate w^+.

We now make an important observation that will be used in the proof of Lemma 4.11 below. Namely, if $\operatorname{gap} \mathcal{J}(w) \geq \bar{g}$ and neither condition (a) nor (b) of Lemma 4.6 holds, then we must have $(B(w^+), N(w^+)) = (B(w), N(w))$ and $\mathcal{J}(w) = \mathcal{J}(w^+)$; i.e., the AS-bipartition and the ordered AS-partition do not change during the iteration from w to w^+. Indeed, assume that $i, j \in \{1, \ldots, n\}$ are indices such that $\delta_i(w) > \delta_j(w)$ and one of them lies in $B(w)$ while the other lies in $N(w)$. Since $\operatorname{gap} \mathcal{J}(w) \geq \bar{g}$, this implies that $\delta_i(w)/\delta_j(w) \geq \bar{g}$. Moreover, since we are assuming that (b) of Lemma 4.6 does not hold, we must have $\delta_i(w^+)/\delta_j(w^+) \geq \bar{g}$, and hence that $\delta_i(w^+) > \delta_j(w^+)$. We have thus shown that the order of the δ_i's for indices of different types are preserved while moving from w to w^+. Moreover, since we are assuming that (a) of Lemma 4.6 does not hold, it follows that indices do not change status while moving from w to w^+. Using these two conclusions together, we easily see that $(B(w^+), N(w^+)) = (B(w), N(w))$ and $\mathcal{J}(w) = \mathcal{J}(w^+)$.

The next lemma describes how the ratio between any pair of dual slacks varies as we move from one primal-dual feasible point to another one.

LEMMA 4.7. *For some $\beta, \beta' \in (0,1)$, let $w = (x,s,y) \in \mathcal{N}(\beta)$ and $w' = (x', s', y') \in \mathcal{N}(\beta')$ be points such that $\mu(w) = \mu(w')$. Then, for every $i, j \in \{1, \ldots, n\}$, we have*

$$\frac{s_i}{s_j} \leq \frac{(1+\beta)(1+\beta')}{(1-\beta)^2(1-\beta')^2} \frac{s'_i}{s'_j}. \tag{4.15}$$

Proof. Let $\mu \equiv \mu(w) = \mu(w')$. It is well known that

$$\frac{1-\beta}{1+\beta} s \leq s(\mu) \leq \frac{1}{1-\beta} s, \qquad \frac{1-\beta'}{1+\beta'} s' \leq s(\mu) \leq \frac{1}{1-\beta'} s'$$

(see, for example, Lemma 2.4(ii) of Gonzaga [3] and Proposition 2.1 of Monteiro and Tsuchiya [19]). For $i, j \in \{1, \ldots, n\}$, these relations then imply that

$$s_i \leq \frac{1+\beta}{1-\beta} s_i(\mu) \leq \frac{1+\beta}{(1-\beta)(1-\beta')} s'_i,$$

$$s_j \geq (1-\beta) s_j(\mu) \geq \frac{(1-\beta)(1-\beta')}{1+\beta'} s'_j,$$

from which (4.15) immediately follows. ∎

The following technical lemma describes how the ratio function $\delta_i(\cdot)/\delta_j(\cdot)$, for a fixed pair $(i,j) \in B(w) \times N(w)$, varies from $w = w^k$ to $w^+ = w^{k+1}$.

LEMMA 4.8. *Suppose $\beta \in (0, 1/4]$ and assume that $w = (x, s, y)$ and $w^+ = (x^+, s^+, y^+)$ are two consecutive iterates of the MTY P-C algorithm. Assume also that $\varepsilon^{\mathrm{a}}_\infty(w) \le \sqrt{1-\beta}/2$. Then, for every $i \in B(w)$ and $j \in N(w)$, we have*

$$\frac{\delta_i(w^+)}{\delta_j(w^+)} \le \frac{\sqrt{n}}{\tau} \frac{\delta_i(w)}{\delta_j(w)} \varepsilon^{\mathrm{a}}_\infty(w), \tag{4.16}$$

where $\tau = \tau(\beta)$ is the constant defined in (4.2).

Proof. To simplify notation, let $\varepsilon^{\mathrm{a}}_\infty \equiv \varepsilon^{\mathrm{a}}_\infty(w)$, $\delta^+ \equiv \delta_i(w^+)$, and $\delta \equiv \delta(w)$. Also, let $\Delta w^{\mathrm{a}} \equiv (\Delta x^{\mathrm{a}}, \Delta s^{\mathrm{a}}, \Delta y^{\mathrm{a}})$ denote the AS direction at w and define $\tilde{w} = (\tilde{x}, \tilde{s}, \tilde{y}) \equiv w + \alpha_a \Delta w^{\mathrm{a}}$. Using (2.4), the fact that $w^+ \in \mathcal{N}(\beta)$, and Lemma 4.7 with $w = w^+$ and $w' = \tilde{w}$, we obtain

$$\frac{\delta_i^+}{\delta_j^+} = \frac{(x_j^+ s_j^+)^{1/2}}{(x_i^+ s_i^+)^{1/2}} \frac{s_i^+}{s_j^+} \le \frac{\sqrt{1+\beta}}{\sqrt{1-\beta}} \frac{s_i^+}{s_j^+} \le \frac{(1+\beta)^{3/2}(1+2\beta)}{(1-\beta)^{5/2}(1-2\beta)^2} \frac{\tilde{s}_i}{\tilde{s}_j}. \tag{4.17}$$

Now, letting $(Rx^{\mathrm{a}}, Rs^{\mathrm{a}}) \equiv (Rx^{\mathrm{a}}(w), Rs^{\mathrm{a}}(w))$ and using Lemma 2.5, (2.11), and the fact that $i \in B(w)$, we obtain

$$\begin{aligned}
\tilde{s}_i &= s_i + \alpha_a \Delta s_i^{\mathrm{a}} = (1-\alpha_a) s_i + \alpha_a (s_i + \Delta s_i^{\mathrm{a}}) \\
&= \sqrt{\mu}\delta_i \left[(1-\alpha_a) \frac{x_i^{1/2} s_i^{1/2}}{\sqrt{\mu}} + \alpha_a R s_i^{\mathrm{a}} \right] \\
&\le \sqrt{\mu}\delta_i \left[\frac{\sqrt{n}\varepsilon^{\mathrm{a}}_\infty}{\beta} \sqrt{1+\beta} + \varepsilon^{\mathrm{a}}_\infty \right] \le \sqrt{\mu}\delta_i \varepsilon^{\mathrm{a}}_\infty \frac{2\sqrt{(1+\beta)n}}{\beta}.
\end{aligned} \tag{4.18}$$

Now, using relations (2.4), (2.9), and (4.3), the fact that $j \in N(w)$ and $\alpha_a \le 1$, and the assumption that $\varepsilon^{\mathrm{a}}_\infty \le \sqrt{1-\beta}/2$, we obtain

$$\begin{aligned}
\tilde{s}_j &= s_j + \alpha_a \Delta s_j^{\mathrm{a}} = \sqrt{\mu}\delta_j \left[\frac{x_j^{1/2} s_j^{1/2}}{\sqrt{\mu}} + \alpha_a \frac{\delta_j^{-1} \Delta s_j^{\mathrm{a}}}{\sqrt{\mu}} \right] \\
&= \sqrt{\mu}\delta_j \left[\frac{x_j^{1/2} s_j^{1/2}}{\sqrt{\mu}} - \alpha_a R x_j^{\mathrm{a}} \right] \ge \sqrt{\mu}\delta_j [\sqrt{1-\beta} - \varepsilon^{\mathrm{a}}_\infty] \ge \sqrt{\mu}\delta_j \frac{\sqrt{1-\beta}}{2}.
\end{aligned} \tag{4.19}$$

Merging (4.18) and (4.19) into (4.17) and noting (4.2), we obtain

$$\frac{\delta_i^+}{\delta_j^+} = \frac{(1+\beta)^{3/2}(1+2\beta)}{(1-\beta)^{5/2}(1-2\beta)^2} \frac{\tilde{s}_i}{\tilde{s}_j} \le \sqrt{n} \frac{4(1+\beta)^2(1+2\beta)}{\beta(1-\beta)^3(1-2\beta)^2} \frac{\delta_i}{\delta_j} \varepsilon^{\mathrm{a}}_\infty = \frac{\sqrt{n}}{\tau} \frac{\delta_i}{\delta_j} \varepsilon^{\mathrm{a}}_\infty. \qquad ∎$$

The next result gives another alternative sufficient condition for the occurrence of $\mathcal{C}$-crossover events. By the definition of the quantity $\operatorname{gap}(\mathcal{J}(w))$, there exists two indices i and j, one lying in $B(w)$ and the other in $N(w)$, such that $\operatorname{gap}(\mathcal{J}(w)) = \delta_i(w)/\delta_j(w)$. The lemma below considers the situation where $\operatorname{gap}(\mathcal{J}(w)) = \delta_i(w)/\delta_j(w)$ for some $i \in B(w)$ and $j \in N(w)$.

LEMMA 4.9. *Suppose $\beta \in (0, 1/4]$ and $w = w^k \in \mathcal{N}(\beta)$ is an iterate of the MTY P-C algorithm. Assume that* $\operatorname{gap}(\mathcal{J}(w)) = \delta_i(w)/\delta_j(w)$ *for some $i \in B(w)$ and $j \in$*

$N(w)$. Then there exists an iteration index $l > k$ such that $l-k = \mathcal{O}(n^{1.5}\log(\bar{\chi}_A+n))$ and with the property that either a $\mathcal{C}$-crossover event occurs between w^k and w^l or the algorithm terminates at or before the lth iteration.

Proof. Let $\mathcal{J} \equiv \mathcal{J}(w)$. If either condition $\mathrm{gap}(\mathcal{J}) \geq \bar{g}$ or $\varepsilon^{\mathrm{a}}_\infty(w) \leq \tau\bar{g}/(\sqrt{n}\,\mathrm{gap}(\mathcal{J}))$ does not hold, then the conclusion of the lemma follows immediately from Lemmas 4.2 and 4.5. Assume then that $\mathrm{gap}(\mathcal{J}) \geq \bar{g}$ and $\varepsilon^{\mathrm{a}}_\infty(w) \leq \tau\bar{g}/(\sqrt{n}\,\mathrm{gap}(\mathcal{J}))$. By assumption, there exist $i \in B(w)$ and $j \in N(w)$ such that $\delta_i(w)/\delta_j(w) = \mathrm{gap}(\mathcal{J}) \geq \bar{g}$. The assumptions on $\mathrm{gap}(\mathcal{J})$ and $\varepsilon^{\mathrm{a}}_\infty(w)$ together with (4.6) imply that $\varepsilon^{\mathrm{a}}_\infty(w) \leq \sqrt{1-2\beta}/4 \leq \sqrt{1-\beta}/2$. Hence, by Lemma 4.8 and the assumption on $\varepsilon^{\mathrm{a}}_\infty(w)$, we have

$$\frac{\delta_i(w^+)}{\delta_j(w^+)} \leq \frac{\sqrt{n}}{\tau}\frac{\delta_i(w)}{\delta_j(w)}\varepsilon^{\mathrm{a}}_\infty(w) = \frac{\sqrt{n}}{\tau}\,\mathrm{gap}(\mathcal{J})\varepsilon^{\mathrm{a}}_\infty(w) \leq \bar{g}, \tag{4.20}$$

where $w^+ \equiv w^{k+1}$. The conclusion of the lemma now follows from Lemma 4.6(b). $\square$

In order to prove our main lemma, namely Lemma 4.11, followed by the proof of Theorem 2.6, we introduce the set $\mathcal{K}$ consisting of those indices k such that the iterates w^k and w^{k+1} of the MTY P-C algorithm satisfy the following set of conditions, where $\mathcal{J}_l \equiv \mathcal{J}(w^l)$ for all $l \geq 0$:

(C1) $\mathrm{gap}(\mathcal{J}_k) > 2\bar{g}$ and $\varepsilon^{\mathrm{a}}_\infty(w^k) < \tau\bar{g}/(\sqrt{n}\ \mathrm{gap}(\mathcal{J}_k))$. (Note that if (C1) does not hold, then $w = w^k$ satisfies the hypotheses of either Lemma 4.2 or 4.5.)

(C2) $\mathrm{gap}(\mathcal{J}_{k+1}) = \delta_i(w^{k+1})/\delta_j(w^{k+1})$ for some $i \in N(w^{k+1})$ and $j \in B(w^{k+1})$. (Note that if (C2) does not hold, then $w = w^{k+1}$ satisfies the hypotheses of Lemma 4.9.)

(C3) $(B(w^{k+1}), N(w^{k+1})) = (B(w^k), N(w^k))$ and $\mathcal{J}_{k+1} = \mathcal{J}_k$. (Note that if (C3) does not hold, then the observation made in the second paragraph after Lemma 4.6 implies that $w = w^k$ satisfies the hypotheses of either Lemma 4.2 or 4.6.)

Note that if $k \notin \mathcal{K}$, then a $\mathcal{C}$-crossover event must occur after $\mathcal{O}(\log(\bar{\chi}_A + n))$ iterations, since $w = w^k$ satisfies the hypotheses of either Lemma 4.2, 4.5, or 4.6, or w^{k+1} satisfies the hypotheses of Lemma 4.9. The following result establishes a number of interesting properties regarding the above set $\mathcal{K}$.

LEMMA 4.10. *For every $k \in \mathcal{K}$, we have*

$$\frac{\mathrm{gap}(\mathcal{J}_{k+1})}{\bar{g}} \geq \left(\frac{\mathrm{gap}(\mathcal{J}_k)}{\bar{g}}\right)^2, \qquad \frac{\mu(w^k)}{\mu(w^{k+1})} \geq \left(\frac{\mathrm{gap}(\mathcal{J}_k)}{\bar{g}}\right) \geq 2. \tag{4.21}$$

Moreover, for some given constant $\eta > 0$, if $l, l+1, \ldots, \hat{l}-1, \hat{l}$ are consecutive indices in the set $\mathcal{K}(\eta) \equiv \{k \in \mathcal{K} : \mu(w^{k+1}) > \eta\}$, then $\hat{l} - l = \mathcal{O}(\log(\log(\mu(w^0)/\eta)))$.

Proof. Fix some $k \in \mathcal{K}$. Then, conditions (C1), (C2), and (C3) above hold for this k. In particular, by conditions (C2) and (C3), we know that there exist $i \in N(w^k)$ and $j \in B(w^k)$ such that

$$\mathrm{gap}(\mathcal{J}_{k+1}) = \frac{\delta_i(w^{k+1})}{\delta_j(w^{k+1})}.$$

Using (4.6) and condition (C1), we easily see that w^k satisfies the hypothesis of Lemma 4.8. Hence, it follows from the above identity, Lemma 4.8, and condition (C1) that

$$\mathrm{gap}(\mathcal{J}_{k+1}) \geq \frac{\tau}{\sqrt{n}\varepsilon^{\mathrm{a}}_\infty(w^k)}\frac{\delta_i(w^k)}{\delta_j(w^k)} \geq \frac{\tau}{\sqrt{n}\varepsilon^{\mathrm{a}}_\infty(w^k)}\,\mathrm{gap}(\mathcal{J}_k) \geq \frac{1}{\bar{g}}[\mathrm{gap}(\mathcal{J}_k)]^2,$$

where the second inequality uses the fact that $\mathcal{J}_k = \mathcal{J}_{k+1}$, which holds in view of (C3). The first inequality in (4.21) follows by dividing the above inequality by $\bar{g}$. Using Lemma 2.5, condition (C1), and the fact that $\tau \le \beta$, we obtain

$$\frac{\mu(w^k)}{\mu(w^{k+1})} \ge \frac{\beta}{\sqrt{n}\varepsilon_\infty^{\mathrm{a}}(w^k)} \ge \frac{\beta\,\mathrm{gap}(\mathcal{J}_k)}{\tau\bar{g}} \ge \frac{\mathrm{gap}(\mathcal{J}_k)}{\bar{g}} > 2 \quad \forall k \in \mathcal{K}, \tag{4.22}$$

showing that the second relation in (4.21) also holds. To show the last statement of the lemma, assume now that $l, l+1, \ldots, \hat{l}$ are consecutive indices in $\mathcal{K}(\eta)$. In view of the first relation in (4.21), we have

$$\log\left(\frac{\mathrm{gap}(\mathcal{J}_{k+1})}{\bar{g}}\right) \ge 2\log\left(\frac{\mathrm{gap}(\mathcal{J}_k)}{\bar{g}}\right) \quad \forall k = l, l+1, \ldots, \hat{l},$$

from where it follows that

$$\log\left(\frac{\mathrm{gap}(\mathcal{J}_{\hat{l}})}{\bar{g}}\right) \ge 2^{\hat{l}-l}\log\left(\frac{\mathrm{gap}(\mathcal{J}_l)}{\bar{g}}\right) \ge 2^{\hat{l}-l}\log 2.$$

Using the fact that $\hat{l} \in \mathcal{K}(\eta)$, relation (4.22) with $k = \hat{l}$, and the previous inequality, we obtain

$$\log\frac{\mu(w^0)}{\eta} \ge \log\frac{\mu(w^{\hat{l}})}{\mu(w^{\hat{l}+1})} \ge \log\left(\frac{\mathrm{gap}(\mathcal{J}_{\hat{l}})}{\bar{g}}\right) \ge 2^{\hat{l}-l}\log 2,$$

from which the last statement of the lemma immediately follows. ∎

We are now in a position to establish the main lemma of this section. Even though the main goal of this lemma is to cover the case where the hypotheses of neither Lemma 4.2 nor Lemma 4.5 hold, this result imposes no condition on the iterate w of the MTY P-C algorithm. Moreover, in contrast to the bounds of Lemmas 4.2 and 4.5, the bound derived in this lemma on the number of iterations needed to guarantee the occurrence of a $\mathcal{C}$-crossover event depends not only on the quantity $n^{1.5}\log(\bar{\chi}_A + n)$ but also on the term $\log(\log(\mu(w^0)/\eta))$.

LEMMA 4.11. *Suppose $\beta \in (0, 1/4]$ and $w = w^k \in \mathcal{N}(\beta)$ is an iterate of the MTY P-C algorithm. Then, there exists an iteration index $l > k$ such that*

$$l - k = \mathcal{O}(n^{1.5}\log(\bar{\chi}_A + n) + \log(\log(\mu(w^0)/\eta))) \tag{4.23}$$

and with the property that either a $\mathcal{C}$-crossover event occurs between w^k and w^l or the algorithm terminates at or before the lth iteration. Furthermore, we have

$$l - k = \mathcal{O}(n^{1.5}\log(\bar{\chi}_A + n) + \log(\mu(w^k)/\mu(w^l))). \tag{4.24}$$

Proof. Consider the set $\mathcal{K}(\eta)$ defined in Lemma 4.10 and let $\hat{k}$ be the first index greater than or equal to k such that $k \notin \mathcal{K}(\eta)$. Note that by Lemma 4.10, such an index $\hat{k}$ exists and satisfies $\hat{k} - k = \mathcal{O}(\log(\log(\mu(w^0)/\eta)))$. The condition $\hat{k} \notin \mathcal{K}(\eta)$ means that either $\mu(w^{\hat{k}+1}) \le \eta$ or $\hat{k} \notin \mathcal{K}$. If the first condition holds, then the conclusion of the lemma obviously holds with $l = \hat{k} + 1$. (The algorithm terminates at the lth iteration in this case.) If the latter condition holds, it follows that $w = w^{\hat{k}}$ satisfies the hypotheses of either Lemma 4.2, 4.5, or 4.6, or that $w^{\hat{k}+1}$ satisfies the hypotheses of Lemma 4.9. In any of these cases, we know that there exists an iteration index $l > \hat{k}$

such that $l - \hat{k} = \mathcal{O}(n^{1.5} \log(\bar{\chi}_A + n))$ and with the property that either a $\mathcal{C}$-crossover event occurs between $w^{\hat{k}}$ and w^l or the algorithm terminates at or before the lth iteration. Clearly, such an iteration index l satisfies the first conclusion of the lemma. To show that (4.24) also holds, it suffices to show that $\hat{k} - k = \mathcal{O}(\log(\mu(w^k)/\mu(w^l)))$. Indeed, since this conclusion obviously holds when $\hat{k} = k$, we may assume that $\hat{k} > k$. This implies that the indices $k, k+1, \ldots, \hat{k}-1$ are in $\mathcal{K}$, and hence, by the second relation in (4.22), we conclude that

$$\log \frac{\mu(w^k)}{\mu(w^l)} \geq \log \frac{\mu(w^k)}{\mu(w^{\hat{k}})} = \sum_{j=k}^{\hat{k}-1} \log \frac{\mu(w^j)}{\mu(w^{j+1})} \geq (\hat{k} - k) \log 2,$$

which clearly implies our claim above. ∎

We are now ready to give the proof of Theorem 2.6.

Proof of Theorem 2.6. Let k^* be the first index such that $\mu(w^{k^*}) \leq \eta$. Using Lemma 4.11 inductively starting from w^0, we can partition the set $\{0, 1, \ldots, k^*\}$ into disjoint subset of consecutive indices $K_1 := \{k_0, k_0+1, \ldots, k_1 - 1\}$, $K_2 := \{k_1, k_1 + 1, \ldots, k_2 - 1\}, \ldots, K_r := \{k_{r-1}, k_{r-1}+1, \ldots, k_r - 1\}$ such that $k_0 = 0$, $k_r = k^*$, and satisfying the following: (i) a $\mathcal{C}$-crossover event occurs between $w^{k_{i-1}}$ and $w^{k_i - 1}$ for $i = 1, \ldots, r-1$; and (ii) the cardinality of K_i satisfies

$$|K_i| = \mathcal{O}\left(\min\left\{\log\left(\log \frac{\mu^0}{\eta}\right), \log\left(\frac{\mu^{k_{i-1}}}{\mu^{k_i}}\right)\right\} + n^{1.5}\log(\bar{\chi}_A + n)\right) \quad \forall i = 1, \ldots, r.$$

Since the number of crossover event is $\mathcal{O}(n^2)$, it follows that $r = \mathcal{O}(n^2)$. These observations, (1.3), and the fact that $k^* = \sum_{i=1}^{r} |K_i|$ clearly imply that $k^* = \mathcal{O}(T(\mu_0/\eta) + n^{3.5}\log(\bar{\chi}_A + n))$. ∎

5. Implications of the main result under the Turing machine model. In section 2, we have presented two iteration-complexity results for the MTY P-C algorithm under the real number computational model. In this section, we discuss a few implications of these results under the Turing machine model. In this model, the entries of the data (A, b, c) are rational numbers of finite bit length and the arithmetic operations on these types of numbers are carried out approximately, i.e., in finite precision. For the sake of simplicity, here we just focus on the number of arithmetic operations and do not deal with the issue of finite precision. However, we observe that the issue of finite precision in the context of the MTY interior-point algorithm can be dealt with using standard and well-known techniques developed in the context of other interior-point methods (e.g., see [20]).

Let L and L_A be the input size of (A, b, c) and A, respectively. It is well known that $\bar{\chi}_A^* \leq \bar{\chi}_A \leq 2^{\mathcal{O}(L_A)}$ and $\xi(A, b, c) \geq 2^{-\mathcal{O}(L)}$. As was mentioned before, we have $\zeta(A, (B_*, N_*)) \leq \bar{\chi}_A$. Therefore, we have the following corollary which immediately follows from Theorem 2.9.

COROLLARY 5.1. *Assume that the data (A, b, c) is integral, and let L and L_A be defined as in the paragraph above. For some $\beta \in (0, 1/4]$, suppose that a point $w^0 \in \mathcal{N}(\beta)$ such that $\mu(w^0) = 2^{\mathcal{O}(L)}$ is given. Then, the version of the MTY P-C algorithm, in which the FT procedure is invoked at every iterate w^k, started from w^0 finds a primal-dual strictly complementary optimal solution w^* in at most*

$$\mathcal{O}(\min\{\sqrt{n}L, \min\{L, n^2 \log L\} + n^{3.5}(L_A + \log n)\}) \tag{5.1}$$

iterations.

Corollary 5.1 assumes that the initial iterate of the MTY P-C algorithm is a well-centered strictly feasible point whose duality gap is not too large. For a general dual pair of linear programs, even if such a point exists, computing this point is as hard as solving the pair of LP problems. In such a case, an auxiliary pair of dual LP problems is constructed whose optimal solution yields that of the original pair of LP problems. Using the big M idea, Vavasis and Ye [30] construct an auxiliary pair of LP problems, which we refer to as the *VY auxiliary LP pair*, associated with the pair of problems (1.1) and (1.2) in order to resolve the initialization issue for their LLS step algorithm. The VY auxiliary LP pair has the following properties: (i) the input size of its coefficient matrix is bounded by $\mathcal{O}(L_A)$; (ii) the sizes of its cost and right-hand coefficients are bounded by $\mathcal{O}(L)$; (iii) it admits a readily available well-centered initial point whose duality gap is $n2^{\mathcal{O}(L)}$; and (iv) if (1.1) and (1.2) have a primal-dual optimal solution, then this solution can be easily obtained from an optimal solution of the VY auxiliary LP pair. Therefore, we obtain the following theorem for solving a general pair of LP problems under the Turing machine model.

THEOREM 5.2. *Assume that the data* (A, b, c) *is integral, and suppose that* (1.1) *and* (1.2) *have a primal-dual optimal solution. Then, the MTY P-C algorithm, with the FT procedure invoked at every iteration, applied to the VY auxiliary LP pair, finds a strict complementary primal-dual optimal solution of* (1.1) *and* (1.2) *in a number of iterations bounded by* (5.1).

6. Concluding remarks. In this paper, we have developed a new iteration-complexity bound for the MTY P-C algorithm using the notion of crossover events due to Vavasis and Ye [30]. In contrast to the iteration-complexity bound developed in [30], ours is scaling-invariant and has an extra but relatively small term, namely, the term $T(\mu_0/\eta)$, where $T(\cdot)$ is defined in (1.3).

Note that the second bound in (2.12), i.e., the inequality $\mu(w^+)/\mu(w) \le \varepsilon_\infty^{\mathrm{a}}(w)\sqrt{n}/\beta$, plays an important role in our analysis. This is also the inequality which plays an important role in establishing that the sequence $\{\mu(w^k)\}$ generated by the MTY P-C algorithm is quadratically convergent. Observe also that if the sequence $\{\mu(w^k)\}$ generated by a fictitious algorithm satisfied $\mu(w^{k+1}) \le C_1\mu(w^k)^2$ for all k, where $C_1 = \mathcal{O}(1)$, then its iteration-complexity bound would be $\mathcal{O}(\log(\log(\mu(w^0)/\eta)))$. The term $n^2\log(\log(\mu(w^0)/\eta))$ which appears in our iteration-complexity bound is due to the fact that this type of quadratic convergence happens over $\mathcal{O}(n^2)$ disjoint finite sets of consecutive iteration indices. A natural conjecture is that any primal-dual interior-point algorithm which achieves the duality gap reduction given by (2.12) at every iteration has the same iteration-complexity bound as the one obtained in this paper. More generally, we conjecture that all interior-point algorithms whose corresponding sequence $\{\mu(w^k)\}$ converges superlinearly or quadratically are suitable for the same type of analysis performed in this paper.

Note that the iteration-complexity bound obtained in this paper under the real-computation model is with respect to a pair of dual LP problems satisfying Assumptions A.1 and A.2. For a pair of dual LP problems which does not satisfy A.1 or A.2, a natural open question is whether one can develop the same type of scaling-invariant iteration-complexity bound under the real-computation model obtained in this paper. Observe that one of the difficulties in this context is the proper choice of the big M constant in the VY auxiliary LP pair.

Our work was strongly motivated by the work of Vavasis and Ye [30]. Therefore, we wonder whether the MTY P-C algorithm with the FT procedure has the iteration-complexity bound $\mathcal{O}(n^{3.5}\log(\bar{\chi}_A^* + n))$, i.e., the iteration-complexity bound of Vavasis and Ye's algorithm but with $\bar{\chi}_A$ replaced by $\bar{\chi}_A^*$. The term involving the two log operators in our iteration-complexity bound is due to the possibility that the assumption of Lemma 4.9 is violated. When this assumption is violated, we have shown that the gap reduces quadratically until some suitable conditions are met in $\mathcal{O}(\log(\log(\mu(w^0)/\eta)))$ iterations which guarantee the occurrence of a $\mathcal{C}$-crossover event. A very challenging open question is whether these conditions are met in $\mathcal{O}(p(n, \log \bar{\chi}_A))$ iterations, where $p(x_1, x_2)$ is a polynomial in $\Re^2$. Note that this would provide a scaling-invariant iteration-complexity bound for the MTY P-C algorithm depending on n and $\bar{\chi}_A^*$ only.

Finally, another interesting topic for future research is whether the results obtained in this paper can be extended, possibly under some extra assumptions, to more general classes of problems such as convex quadratic programming or semidefinite programming.

Acknowledgments. The authors would like to thank the two anonymous referees and the associate editor for their constructive comments which have helped to improve the presentation of the paper. In particular, one of the referees found a few minor errors in the original proofs of Lemmas 4.6 and 4.8, which have been corrected in this present version.

This work was done while the first author was visiting Professor Cornelis Roos at the Department of Information Systems and Algorithms of Delft University of Technology during the summer of 2001; while visiting the second author at the Institute of Statistical Mathematics in August of 2002 and October of 2003; and while the second author was visiting the first author at the School of Industrial and Systems Engineering in June and September of 2002. They are grateful to their host institutions for the financial support and congenial scientific atmosphere provided during these visits.

REFERENCES

[1] I. I. Dikin, *On the speed of an iterative process,* Upr. Sist., 12 (1974), pp. 54–60 (in Russian).

[2] 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 (1994), pp. 45–72.

[3] C. C. Gonzaga, *Path-following methods for linear programming*, SIAM Rev., 34 (1992), pp. 167–224.

[4] C. C. Gonzaga and H. J. Lara, *A note on properties of condition numbers*, Linear Algebra Appl., 261 (1997), pp. 269–273.

[5] O. Güler and Y. Ye, *Convergence behavior of interior-point algorithms*, Math. Program., 60 (1993), pp. 215–228.

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

[7] M. Kojima, S. Mizuno, and A. Yoshise, *A primal-dual interior point algorithm for linear programming*, in Progress in Mathematical Programming, Interior-Point and Related Methods, N. Megiddo, ed., Springer-Verlag, New York, 1989, pp. 29–47.

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

[9] L. McLinden, *The complementarity problem for maximal monotone multifunction*, in Variational Inequalities and Complementarity Problems, R. W. Cottle, F. Giannessi, and J. L. Lions, eds., John Wiley, New York, 1980, pp. 251–270.

[10] N. Megiddo, *Pathways to the optimal set in linear programming*, in Progress in Mathematical Programming (Pacific Grove, CA, 1987), Springer, New York, Berlin, 1989, pp. 131–158.

[11] N. Megiddo, S. Mizuno, and T. Tsuchiya, *A modified layered-step interior-point algorithm for linear programming*, Math. Program., 82 (1998), pp. 339–355.

[12] S. Mehrotra and Y. Ye, *Finding an interior point in the optimal face of linear programs*, Math. Program., 62 (1993), pp. 497–515.

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

[14] S. Mizuno, N. Megiddo, and T. Tsuchiya, *A linear programming instance with many crossover events*, J. Complexity, 12 (1996), pp. 474–479.

[15] R. D. C. Monteiro and I. Adler, *Interior path following primal-dual algorithms. Part* I*: Linear programming*, Math. Program., 44 (1989), pp. 27–41.
[16] R. D. C. Monteiro and I. Adler, *Interior path-following primal-dual algorithms. Part* II*: Convex quadratic programming*, Math. Program., 44 (1989), pp. 43–66.
[17] R. D. C. Monteiro, I. Adler, and M. G. C. Resende, *A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension*, Math. Oper. Res., 15 (1990), pp. 191–214.
[18] R. D. C. Monteiro and T. Tsuchiya, *Global convergence of the affine scaling algorithm for convex quadratic programming*, SIAM J. Optim., 8 (1998), pp. 26–58.
[19] R. D. C. Monteiro and T. Tsuchiya, *A variant of the Vavasis–Ye layered-step interior-point algorithm for linear programming*, SIAM J. Optim., 13 (2003), pp. 1054–1079.
[20] J. Renegar, *A polynomial-time algorithm, based on Newton's method, for linear programming*, Math. Program., 40 (1988), pp. 59–93.
[21] G. W. Stewart, *On scaled projections and pseudoinverses*, Linear Algebra Appl., 112 (1989), pp. 189–193.
[22] K. Tanabe, *Centered Newton method for mathematical programming*, in System Modeling and Optimization, M. Iri and K. Yajima, eds., Springer-Verlag, Berlin, 1988, pp. 197–206.
[23] É. Tardos, *A strongly polynomial algorithm to solve combinatorial linear programs*, Oper. Res., 34 (1986), pp. 250–256.
[24] M. J. Todd, *A Dantzig-Wolfe-like variant of Karmarkar's interior-point linear programming algorithm*, Oper. Res., 38 (1990), pp. 1006–1018.
[25] M. J. Todd, L. Tunçel, and Y. Ye, *Characterizations, bounds, and probabilistic analysis of two complexity measures for linear programming problems*, Math. Program., 90 (2001), pp. 59–70.
[26] T. Tsuchiya, *Global convergence property of the affine scaling method for primal degenerate linear programming problems*, Math. Oper. Res., 17 (1992), pp. 527–557.
[27] L. Tunçel, *Approximating the complexity measure of Vavasis-Ye algorithm is NP-hard*, Math. Program., 86 (1999), pp. 219–223.
[28] L. Tunçel, *On the condition numbers for polyhedra in Karmarkar's form*, Oper. Res. Lett., 24 (1999), pp. 149–155.
[29] R. J. Vanderbei and J. C. Lagarias, *I. I. Dikin's convergence result for the affine-scaling algorithm*, in Mathematical Developments Arising from Linear Programming, Contemp. Math. 114, AMS, Providence, RI, 1990, pp. 109–119.
[30] S. Vavasis and Y. Ye, *A primal-dual accelerated interior point method whose running time depends only on A*, Math. Program., 74 (1996), pp. 79–120.
[31] S. Vavasis and Y. Ye, *A simplification to "a primal-dual interior point method whose running time depends only on the constraint matrix,"* in High Performance Optimization, Appl. Optim. 33, Kluwer, Dordrecht, 2000, pp. 233–243.
[32] S. Vavasis and Y. Ye, *On the relationship between layered least squares and affine scaling steps*, in The Mathematics of Numerical Analysis, Lectures in Appl. Math. 32, AMS, Providence, RI, 1996, pp. 857–865.
[33] Y. Ye, *On the finite convergence of interior-point algorithms for linear programming*, Math. Program., 57 (1992), pp. 325–335.
[34] Y. Ye, *Interior-Point Algorithms: Theory and Analysis*, John Wiley and Sons, New York, 1997.