Annals of Operations Research 47(1993)509–539

Global convergence of the affine scaling algorithm for primal degenerate strictly convex quadratic programming problems*

Takashi Tsuchiya
The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan

In this paper we deal with global convergence of the affine scaling algorithm for strictly convex QP problems satisfying a dual nondegeneracy condition. By means of the local Karmarkar potential function which was successfully applied to demonstrate global convergence of the affine scaling algorithm for LP, we show global convergence of the algorithm when the step-size 1/8 is adopted without requiring any primal nondegeneracy condition.

0. Introduction

Since Karmarkar [5] proposed the projective scaling algorithm for linear programming in 1984, a number of interior point algorithms have been proposed and implemented. The affine scaling algorithm, originated by Dikin [3] and rediscovered by several authors including Barnes [2], Vanderbei et al. [12], and Adler et al. [1] is one of the most popular interior point algorithms obtained by substituting the affine scaling transformation in place of the projective transformation in Karmarkar's algorithm.

One of the major problems in the theoretical analysis of the affine scaling algorithms is global convergence under the existence of degeneracy. Global analysis of the algorithm reduces to the analysis of the behavior near the boundary of the feasible polyhedron. Global convergence of the affine scaling algorithm for linear programming was shown in [10, 11], by introducing the local Karmarkar potential function as a tool to analyze the behavior of the algorithm near a degenerate boundary of the feasible region.

In this paper we deal with global convergence of the affine scaling algorithm for convex quadratic programming problems proposed by Dikin and Zolkaltsev [4]

*This paper was presented at the 14th International Symposium on Mathematical Programming, held August 5–9, 1991, in Amsterdam, The Netherlands.

and Ye [13]. We focus our analysis on strictly convex quadratic programming problems satisfying a certain dual nondegeneracy condition, and extend the analysis given in [10, 11] to demonstrate that the step-size 1/8 is sufficient in guaranteeing the global convergence of the algorithm.

In view of the nondegeneracy conditions, the best global convergence result so far is obtained by Sun [8] for general convex quadratic programming problems with the choice of the step-sizes as small as 2^{-L} (L is the input size of the problem) without requiring any nondegeneracy conditions, where he demonstrated the ergodic convergence of the dual estimate following the proof of the global convergence of the algorithm for linear programming by Tseng and Luo [9]. If we assume the primal nondegeneracy condition and restrict the problem to strictly convex quadratic programming, the global convergence with a step-size less than one follows by adding a little argument to the convergence results in [15]. Though our result is for strictly convex quadratic programming and requires the dual nondegeneracy condition, it may still be of interest because the analysis does not require any primal nondegeneracy condition, i.e., the feasible region may be primal degenerate, and because the step-size 1/8 is a considerable improvement compared with 2^{-L}. This is an intermediate result towards a proof of global convergence of the affine scaling algorithm with the step-size, say, 1/8, for convex quadratic programming without requiring nondegeneracy assumptions.

1. Problem and the main result

Let us consider the following strictly convex quadratic programming problem $\langle\mathbf{D}\rangle$ to minimize a strictly convex quadratic function $F(x)$ over a polyhedron $\mathscr{P} \in \mathbb{R}^n$:

$$\begin{aligned} &\text{minimize} \quad F(x), \\ &\text{subject to} \quad x \in \mathscr{P}, \\ &\mathscr{P} = \{x \in \mathbb{R}^n | A^{\mathrm{T}}x - b \geq 0\}, \\ &A = (a_1, \ldots, a_m) \in \mathbb{R}^{n\times m}, \quad x \in \mathbb{R}^n, \quad b \in \mathbb{R}^m, \end{aligned} \tag{1.1}$$

where we assume the following:

(A1) The feasible region $\mathscr{P}$ has an interior point and $Rank(A) = n$.

For a vector v, we denote by $[v]$ the diagonal matrix whose diagonal entries are elements of v. We denote the slack variables $A^{\mathrm{T}}x - b$ by $\xi(x)$, and define the "metric" matrix $G(x)$ for the affine scaling algorithm as follows:

$$G(x) = A[\xi(x)]^{-2}A^{\mathrm{T}}. \tag{1.2}$$

1 and I denote the vector of all ones and the identity matrix of proper dimension, respectively. We use $\|\cdot\|$ (without subscript) for the 2-norm. For the sequence $\{x^{(\nu)}\}$ ($\nu = 1,\ldots; x^{(\nu)} \in \mathbb{R}^n$), we abbreviate $\{f(x^{(\nu)})\}$, $\{g(x^{(\nu)})\}$ etc. as $\{f^{(\nu)}\}$, $\{g^{(\nu)}\}$ etc. We denote by x^+ the new point obtained by performing one iterative step at the point $x \in \mathbb{R}^n$, and use f^+, g^+, etc. to denote $f(x^+), g(x^+)$, etc. We do not indicate arguments of functions when they are obvious from the context.

Let $x^{(\nu)}$ be an interior point of the polyhedron $\mathscr{P}$. In the affine scaling algorithm for $\langle\mathbf{D}\rangle$, we determine the next iterate $x^{(\nu+1)}$ as the optimal solution for the following minimization problem

$$\begin{aligned} &\text{minimize}_x \quad F(x), \\ &\text{subject to} \quad \{(x - x^{(\nu)})^t G(x^{(\nu)})(x - x^{(\nu)})\}^{1/2} \le \mu^{(\nu)}, \end{aligned} \tag{1.3}$$

where $\mu^{(\nu)}$ is a constant such that $0 < \mu_{\min} \le \mu^{(\nu)} \le \mu_{\max} < 1$.

It is well-known that $x^{(\nu+1)}$ also remains an interior point if $0 < \mu^{(\nu)} < 1$. If $x^{(\nu+1)} = x^{(\nu)}$, then we terminate the iteration. In this case, since $x^{(\nu)}$ is an interior point, $x^{(\nu)}$ is the global minimum point of F over $\mathbb{R}^n$. Note that the optimization problem of this type appears in the context of the trust region algorithms [6]. Recently, Ye [14] gave a remarkable result that the problem can be solved in $O(\log\log(1/\epsilon))$ iterations to the precision ϵ by a combination of Newton's method and a binary search, where each iteration solves a system of linear equations. Thus, (1.3) can be solved efficiently. In terms of the slack variables, we may write the problem (1.3) as

$$\begin{aligned} &\text{minimize}_x \quad F(x), \\ &\text{subject to} \quad \|[\xi(x^{(\nu)})]^{-1}(\xi(x) - \xi(x^{(\nu)}))\| \le \mu^{(\nu)}. \end{aligned} \tag{1.4}$$

The Karush–Kuhn–Tucker condition for (1.3) is

$$\begin{aligned} &\frac{\partial F}{\partial x}(x) - \lambda G(x^{(\nu)})(x - x^{(\nu)}) = 0, \\ &\lambda(\{(x - x^{(\nu)})^{\mathrm{T}} G(x^{(\nu)})(x - x^{(\nu)})\}^{1/2} - \mu^{(\nu)}) = 0, \\ &\lambda \ge 0, \end{aligned} \tag{1.5}$$

and it is not difficult to see that $x^{(\nu+1)}$ can be written as follows (except for the special case where $x^{(\nu+1)}$ happens to be exactly the optimal solution of $\langle\mathbf{D}\rangle$):

$$x^{(\nu+1)} = x^{(\nu)} - \mu^{(\nu)} \frac{G(x^{(\nu)})^{-1} g(x^{(\nu+1)})}{\{g(x^{(\nu+1)})^{\mathrm{T}} G(x^{(\nu)})^{-1} g(x^{(\nu+1)})\}^{1/2}}, \tag{1.6}$$

where $g(x) = \partial F(x)/\partial x$. Because of the convexity of $F(x)$, we have the following relation

$$\begin{aligned} F(x^{(\nu+1)}) - F(x^{(\nu)}) &\le g(x^{(\nu+1)})^{\mathrm{T}}(x^{(\nu+1)} - x^{(\nu)}) \\ &= -\mu^{(\nu)}\{g(x^{(\nu+1)})^{\mathrm{T}} G(x^{(\nu)})^{-1} g(x^{(\nu+1)})\}^{1/2}. \end{aligned} \tag{1.7}$$

To describe our main result, it is necessary to introduce a dual nondegeneracy condition for $\langle \mathbf{D} \rangle$. Let $\mathscr{X}$ be a face of $\mathscr{P}$. We define "the subproblem $\langle \mathbf{D}_{\mathscr{X}} \rangle$ associated with $\mathscr{X}$ as follows:

$$\text{minimize } F(x), \quad \text{subject to } x \in \mathscr{X}. \tag{1.8}$$

A point $x \in \mathscr{P}$ is said to be "a face-optimal point" if x is the optimal solution for the subproblem associated with a face of $\mathscr{P}$.

Given a face-optimal point x^*, let $\mathscr{X}$ be the face that contains x^* as its interior point, and let $g^* = \partial F(x^*)/\partial x$. It is easily seen that the linear function $g^{*\mathrm{T}}x$ is constant on $\mathscr{X}$. If there exists no other face of $\mathscr{P}$ than $\mathscr{X}$ containing $\mathscr{X}$ on which $g^{*\mathrm{T}}x$ is constant, we refer to x^* as "a dual nondegenerate face-optimal point". We require the following nondegeneracy condition concerning the face-optimal points.

(A2) Every face-optimal point of $\langle \mathbf{D} \rangle$ is dual nondegenerate.

When applied to the case of linear programming, **(A2)** is equivalent to the assumption of dual nondegeneracy required in [11]. Hence we refer to this assumption as "the assumption of dual nondegeneracy". Now, we are ready to describe the main theorem in this paper.

THEOREM 1.1

Let $\langle \mathbf{D} \rangle$ be a problem satisfying the assumptions **(A1)** and **(A2)**, and apply the affine scaling algorithm with $0 < \mu_{\min} \le \mu_{\max} \le 1/8$. Then, the algorithm either (I) terminates after a finite number of iterations yielding the global minimum point of $F(x)$ over $\mathbb{R}^n$, or (II) generates an infinite sequence that converges to the optimal solution of $\langle \mathbf{D} \rangle$.

We emphasize that this theorem does not require the primal nondegeneracy condition. It applies to the cases where the feasible region is primal degenerate, provided that the problem satisfies the dual nondegeneracy condition **(A2)**. Since case (I) is trivial, in the remaining part, we focus on case (II), assuming that the algorithm generates an infinite sequence.

2. Preliminaries

In this section we introduce some more notations and describe basic results obtained in [10, 11] which will be used in this analysis. We also make some preliminary observations.

(1) We use the letters $\mathscr{A}, \mathscr{B}, \ldots, \mathscr{X}$ to denote the faces of $\mathscr{P}$. We do not treat the empty set as a face. For a face $\mathscr{X}$ of $\mathscr{P}$, we denote by $E(\mathscr{X})$ the set of indices of the constraints which are always satisfied with equality on the face. We sometimes abbreviate $E(\mathscr{X})$ as E when the face $\mathscr{X}$ which is associated with the notation E is obvious from the context.

(2) Given a set $X \subseteq \{1, \ldots, m\}$ of indices, we denote by A_X^{T}, b_X the matrix and the vector composed of the corresponding coefficient vectors and constants. We use $\xi_X(x)$ for $A_X^{\mathrm{T}}x - b_X$. Analogously, for a vector v, we denote by v_X the vector which is composed of the part of v associated with X. A matrix with a pair of index sets as the lower indices, $C_{X_1X_2}$, say, represents a matrix whose rows and columns are associated with the first set X_1 and the second one X_2, respectively.

(3) A point x on a face $\mathscr{X}$ of $\mathscr{P}$ is referred to as an "interior point of $\mathscr{X}$" if $\xi_{E(\mathscr{X})}(x) = 0$ and $\xi_i(x) > 0$ $(i \notin E(\mathscr{X}))$. The interior point of a vertex is the vertex itself. The face $\mathscr{X}$ is characterized as the smallest face (as a set) among the faces which contain the point x as their element.

(4) For an index set X, we use $|X|$ to denote its cardinality. If X is a (proper) subset of another index set Y, we denote $X \subseteq (\subset) Y$. Then we denote by $Y - X$ the set consisting of the indices which belong to X but not to Y. The complement of X, which is defined as $\{1, \ldots, m\} - X$, is written by X^c.

See [7] for the basic theory of polyhedra. Given an index set X, we define

$$\Phi_X(x) = \frac{\|\xi_X(x)\|}{\min_{i \notin X} \xi_i(x)} = \frac{\|\xi_X(x)\|}{\min_{i \in X^c} \xi_i(x)}. \tag{2.1}$$

The following lemma relates the existence of a sequence of interior points to the existence of a face with the quantity (2.1).

LEMMA 2.1 (LEMMA 3.2 OF [10])

Let X be a nonempty index set of constraints, and let $\{x^{(\nu)}\}$ be a sequence of interior points of $\mathscr{P}$. If

$$\text{(i) } \xi_X^{(\nu)} \to 0 \quad \text{and} \quad \text{(ii) } \Phi_X(x^{(\nu)}) = \frac{\|\xi_X^{(\nu)}\|}{\min_{i \notin X} \xi_i^{(\nu)}} \quad \text{converges to zero,}$$

then there exists a face $\mathscr{X}$ such that $E(\mathscr{X}) = X$. □

Let Z be an index set of constraints. We can choose the index set $B \subseteq Z$ such that the columns of A_B form a basis for the range space of A_Z. Since $Rank(A) = n$, due to the elementary theory of linear algebra, we can choose the index set $\bar{B}$ from the complement of Z such that $A_{B \cup \bar{B}}$ is a nonsingular matrix. Then $\xi_{B \cup \bar{B}}$ is regarded as another coordinate system, where the coordinate transformation is given by

$$\xi_{B \cup \bar{B}}(x) = A^{\mathrm{T}}_{B \cup \bar{B}} x - b_{B \cup \bar{B}} \quad \text{and} \quad x(\xi_{B \cup \bar{B}}) = (A^{\mathrm{T}}_{B \cup \bar{B}})^{-1}(\xi_{B \cup \bar{B}} + b_{B \cup \bar{B}}). \tag{2.2}$$

We refer to the pair $(B, \bar{B})$ as a "pair of basis index sets associated with the index set Z". In this paper we use the letters B and $\bar{B}$ as the notation for such pairs of basis index sets. When we want to make clear that the pair is associated with the index set Z, we write them as $B(Z)$ and $\bar{B}(Z)$. We refer to $(\xi_{B(Z)}, \xi_{\bar{B}(Z)})$ as the "slack coordinate associated with the index set Z". We denote by $R(Z, B)$ the index set $Z - B$. Due to the definitions, there exists a matrix T_{BR} such that

$$A_R = A_B T_{BR}. \tag{2.3}$$

Thus, with the index set Z and its associated pair of basis index sets $(B, \bar{B})$ determined, we define the matrices $\bar{A}_{B(Z)}$ and $\bar{A}_{\bar{B}(Z)}$ as

$$\begin{pmatrix} \bar{A}_{B(Z)} \\ \bar{A}_{\bar{B}(Z)} \end{pmatrix} \equiv (A_{B(Z)} \quad A_{\bar{B}(Z)})^{-1} = (A_{B(Z) \cup \bar{B}(Z)})^{-1}. \tag{2.4}$$

Then, we have

$$\begin{aligned} &\begin{pmatrix} \bar{A}_{B(Z)} A_{B(Z)} & \bar{A}_{B(Z)} A_{\bar{B}(Z)} \\ \bar{A}_{\bar{B}(Z)} A_{B(Z)} & \bar{A}_{\bar{B}(Z)} A_{\bar{B}(Z)} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}, \\ &A_{B(Z)} \bar{A}_{B(Z)} + A_{\bar{B}(Z)} \bar{A}_{\bar{B}(Z)} = I. \end{aligned} \tag{2.5}$$

Note that $A_{B(Z)} \bar{A}_{B(Z)}$ and $A_{B(Z)} \bar{A}_{\bar{B}(Z)}$ are projection matrices. With these notations, the constraints can be categorized into four groups:

$$\xi(x) = A^{\mathrm{T}} x - b = \begin{pmatrix} A^{\mathrm{T}}_{R(Z,B)} x - b_{R(Z,B)} \\ A^{\mathrm{T}}_{B(Z)} x - b_{B(Z)} \\ A^{\mathrm{T}}_{\bar{B}(Z)} x - b_{\bar{B}(Z)} \\ A^{\mathrm{T}}_{N(Z,\bar{B})} x - b_{N(Z,\bar{B})} \end{pmatrix} = \begin{pmatrix} \xi_R(x) \\ \xi_B(x) \\ \xi_{\bar{B}}(x) \\ \xi_N(x) \end{pmatrix}, \tag{2.6}$$

where $N(Z, \bar{B}) = \{1, \ldots, m\} - Z - \bar{B} = \{1, \ldots, m\} - R - B \cup \bar{B}$. We use R and N also as global notations in this paper. We omit the arguments (Z, B) of R and $(Z, \bar{B})$ of N if they are obvious from the context.

We abuse notation by introducing the abbreviation $\nabla_{B'} = \partial/\partial\xi_{B'}$, where $B' \subseteq B(Z) \cup \bar{B}(Z)$. We define

$$\eta(x) \equiv (\eta_R(x), \eta_B(x), \eta_{\bar{B}}(x), \eta(x)) = (0, \nabla_{B\cup\bar{B}}F(x), 0). \tag{2.7}$$

When Z is an index set for the always-active constraints on a face $\mathscr{X}$, say, i.e., in the case of $Z = E(\mathscr{X})$, we use conventional notations $B(\mathscr{X})$ and $\bar{B}(\mathscr{X})$ for $B(E(\mathscr{X}))$ and $\bar{B}(E(\mathscr{X}))$, respectively.

Let x^* be a face-optimal point of $\langle\mathbf{D}\rangle$ which is an interior point of the face $\mathscr{X}$ of $\mathscr{P}$. The point x^* is the minimum point of F over the face $\mathscr{X}$. Choose a pair of the basis index set $(B(\mathscr{X}), \bar{B}(\mathscr{X}))$ associated with $E(\mathscr{X})$ to take the slack coordinate $(\xi_{B(\mathscr{X})}, \xi_{\bar{B}(\mathscr{X})})$. We have the following proposition.

PROPOSITION 2.2

The objective function F is represented as follows, in terms of the slack coordinate $(\xi_{B(\mathscr{X})}, \xi_{\bar{B}(\mathscr{X})})$:

$$F(x) = \tfrac{1}{2}(\xi_B^{\mathrm{T}} \quad (\xi_{\bar{B}} - \xi_{\bar{B}}^*)^{\mathrm{T}}) \begin{pmatrix} H_{BB} & H_{B\bar{B}} \\ H_{B\bar{B}}^{\mathrm{T}} & H_{\overline{BB}} \end{pmatrix} \begin{pmatrix} \xi_B \\ \xi_{\bar{B}} - \xi_{\bar{B}}^* \end{pmatrix} + \eta_B^{*\mathrm{T}}\xi_B + F^*. \tag{2.8}$$

Here H is the Hessian matrix of F with respect to $(\xi_B, \xi_{\bar{B}})$, and ξ^*, η^*, F^* are the slack variables, the gradient and the function value of F at x^*, respectively.

Proof

By Taylor's expansion, it is easy to see that $F(x)$ is written as

$$\begin{aligned} F(x) =& \tfrac{1}{2}(\xi_B^{\mathrm{T}} \quad (\xi_{\bar{B}} - \xi_{\bar{B}}^*)^{\mathrm{T}}) \begin{pmatrix} H_{BB} & H_{B\bar{B}} \\ H_{B\bar{B}}^{\mathrm{T}} & H_{\overline{BB}} \end{pmatrix} \begin{pmatrix} \xi_B \\ \xi_{\bar{B}} - \xi_{\bar{B}}^* \end{pmatrix} \\ &+ (\eta_B^{*\mathrm{T}} \quad \eta_{\bar{B}}^{*\mathrm{T}}) \begin{pmatrix} \xi_B \\ \xi_{\bar{B}} - \xi_{\bar{B}}^* \end{pmatrix} + F^*. \end{aligned} \tag{2.9}$$

It is enough to show $\eta_{\bar{B}}^* = 0$. Assume, by contradiction, $\eta_{\bar{B}}^* \neq 0$. Since $A_{B\cup\bar{B}}^{\mathrm{T}}$ is an invertible matrix, there exists a vector y such that

$$A_{B\cup\bar{B}}^{\mathrm{T}} y = -\begin{pmatrix} 0 \\ \eta_{\bar{B}}^* \end{pmatrix}. \tag{2.10}$$

Since $\xi^*_{E^c(\mathscr{X})} > 0$ while $\xi^*_{E(\mathscr{X})} = 0$, we have, for sufficiently small $\epsilon > 0$,

$$F(x^* + \epsilon y) < F(x^*), \quad (x^* + \epsilon y) \in \mathscr{X}, \tag{2.11}$$

which is a contradiction. Thus we have $\eta^*_{\bar{B}} = 0$. □

We introduce another coordinate here, which has its origin at the face-optimal point x^* and in terms of which we have a simpler form of $F(x)$:

$$\begin{pmatrix} \theta_B \\ \theta_{\bar{B}} \end{pmatrix} \equiv \begin{pmatrix} I & 0 \\ H_{\bar{B}\bar{B}}^{-1} H_{B\bar{B}}^{\mathrm{T}} & I \end{pmatrix} \begin{pmatrix} \xi_B \\ \xi_{\bar{B}} \end{pmatrix} - \begin{pmatrix} 0 \\ \xi^*_{\bar{B}} \end{pmatrix} = \begin{pmatrix} \xi_B \\ H_{\bar{B}\bar{B}}^{-1} \eta_{\bar{B}} \end{pmatrix}. \tag{2.12}$$

We refer to the coordinate $(\theta_B, \theta_{\bar{B}})$ as "the local coordinate associated with $(x^*, (\xi_B, \xi_{\bar{B}}))$". It is easy to verify the following lemma.

LEMMA 2.3

With the notations above, the objective function is written as follows:

$$F(x) = \tfrac{1}{2}\theta_B^{\mathrm{T}} \tilde{H}_{BB} \theta_B + \tfrac{1}{2}\theta_{\bar{B}}^{\mathrm{T}} H_{\bar{B}\bar{B}} \theta_{\bar{B}} + \eta_B^{*\mathrm{T}} \theta_B + F^*, \tag{2.13}$$

where

$$\tilde{H}_{BB} = H_{BB} - H_{B\bar{B}} H_{\bar{B}\bar{B}}^{-1} H_{B\bar{B}}^{\mathrm{T}}. \tag{2.14}$$

□

Let $\tilde{\eta}(x)$ be a vector in $\mathbb{R}^m$ such that $A\tilde{\eta}(x) = g(x) = \partial F(x)/\partial x$. The iteration (1.5) is written as follows in the space of slack variables $\xi(x)$:

$$\begin{aligned} \xi(x^+) = A^{\mathrm{T}} x^+ - b &= A^{\mathrm{T}} x - b - \mu A^{\mathrm{T}} \frac{G(x)^{-1} g(x^+)}{\{g(x^+)^{\mathrm{T}} G(x)^{-1} g(x^+)\}^{1/2}} \\ &= A^{\mathrm{T}} x - b - \mu A^{\mathrm{T}} \frac{G(x)^{-1} A\tilde{\eta}(x^+)}{\{\tilde{\eta}(x^+)^{\mathrm{T}} A^{\mathrm{T}} G(x)^{-1} A \tilde{\eta}(x^+)\}^{1/2}} \\ &= \xi(x) - \mu[\xi(x)] \frac{P(x)\alpha(x)}{\{\alpha(x)^{\mathrm{T}} P(x) \alpha(x)\}^{1/2}}, \end{aligned} \tag{2.15}$$

where

$$P(x) = [\xi(x)]^{-1} A^{\mathrm{T}} G(x)^{-1} A [\xi(x)]^{-1} \quad \text{and} \quad \alpha(x) = [\xi(x)] \tilde{\eta}(x^+). \tag{2.16}$$

Note that $P(x)$ is a projection matrix. Multiplying both sides of (2.15) by $[\xi]^{-1}$, we have

$$[\xi]^{-1}\xi^{+} = \mathbf{1} - \mu\frac{P\alpha}{(\alpha^{\mathrm{T}}P\alpha)^{1/2}}. \tag{2.17}$$

From this formula, one easily verifies the following proposition.

PROPOSITION 2.4

For each component of the slack variables in the iteration (2.15), we have

$$(1-\mu^{(\nu)})\xi_i^{(\nu)} \le \xi_i^{(\nu+1)} \le (1+\mu^{(\nu)})\xi_i^{(\nu)}. \tag{2.18}$$

We analyze the asymptotic properties of $\{x^{(\nu)}\}$ on the basis of (2.15). Hence it is necessary to obtain an asymptotic formula of P when $\{x^{(\nu)}\}$ approaches to a face. This subject was studied extensively in [11]. From lemmas 4.1, 4.2, 4.3 of [11] we obtain the following lemma.

LEMMA 2.5

Let $\mathscr{X}$ be a face, and choose a pair of basis index sets $(B(\mathscr{X}), \bar{B}(\mathscr{X}))$ associated with $E(\mathscr{X})$ to take a slack coordinate $(\xi_{B(\mathscr{X})}, \xi_{\bar{B}(\mathscr{X})})$. Let x be an interior point of $\mathscr{P}$ and let the slack variables $\xi(x)$ be put in order as $\xi(x) = (\xi_{E(\mathscr{X})}(x), \xi_{\bar{B}(\mathscr{X})}(x), \xi_{N(\mathscr{X})}(x)) = (\xi_{R(\mathscr{X})}(x), \xi_{B(\mathscr{X})}(x), \xi_{\bar{B}(\mathscr{X})}(x), \xi_{N(\mathscr{X})}(x))$. Then the matrix $P(x)$ is written as follows:

$$P(x) = \begin{array}{c} \\ E(\mathscr{X}) \\ E^c(\mathscr{X}) \end{array}\!\!\begin{array}{c} \begin{array}{cc} E(\mathscr{X}) & E^c(\mathscr{X}) \end{array} \\ \begin{pmatrix} \hat{P}_{EE} & 0 \\ 0 & \hat{P}_{E^cE^c} \end{pmatrix} \end{array} + \Delta P, \tag{2.19}$$

where

$$\hat{P}_{EE} = \begin{pmatrix} S_{BR}^{\mathrm{T}} \\ I \end{pmatrix}(I + S_{BR}S_{BR}^{\mathrm{T}})^{-1}(S_{BR} \quad I),$$

$$\hat{P}_{E^cE^c} = \begin{pmatrix} I \\ S_{\bar{B}N}^{\mathrm{T}} \end{pmatrix}(I + S_{\bar{B}N}S_{\bar{B}N}^{\mathrm{T}})^{-1}(I \quad S_{\bar{B}N}), \tag{2.20}$$

$$S_{BR} = [\xi_B]\bar{A}_B A_R[\xi_R]^{-1},$$

$$S_{\bar{B}N} = [\xi_{\bar{B}}]\bar{A}_{\bar{B}}A_N[\xi_N]^{-1},$$

Here $\hat{P}_{EE}$ and $\hat{P}_{E^cE^c}$ are projection matrices. $\|\Delta P(x)\| = O(\Phi_E(x))$ as $\Phi_E(x) \to 0$.

Further, we have

$$\hat{P}_{EE}(x)\mathbf{1}_E = \mathbf{1}_E. \tag{2.21}$$

□

3. Convergence of the sequence

In this section we demonstrate that the sequence $\{x^{(\nu)}\}$ generated by the algorithm converges to a face-optimal point x^* of $\langle \mathbf{D} \rangle$, and observe properties of the sequence in the limit. As noted at the end of section 1, we assume that $\{x^{(\nu)}\}$ is an infinite sequence. The assumption **(A2)** is used only for lemma 3.6 and lemma 3.7. To make clear the role of **(A2)**, we state explicitly which assumptions among **(A1)** and **(A2)** are necessary for each result to hold.

LEMMA 3.1

Let $\{x^{(\nu)}\}$ be a sequence generated by the affine scaling algorithm for $\langle \mathbf{D} \rangle$ under the assumption **(A1)**. The sequence converges to a face-optimal point of $\langle \mathbf{D} \rangle$.

Proof

Since the level set of $F(x)$ is compact and $\{F(x^{(\nu)})\}$ is a monotone decreasing sequence bounded below, the sequence $\{x^{(\nu)}\}$ has an accumulation point, say, x^*. We observe that x^* is a face-optimal point. Denote by $\mathscr{X}$ the face that contains x^* in its interior, and take a slack coordinate $(\xi_{B(\mathscr{X})}, \xi_{\bar{B}(\mathscr{X})})$ associated with the index set $E(\mathscr{X})$. Let $\{x^{(\nu_\tau)}\}$ a subsequence converging to x^*. Then, we have

$$\lim_{\tau\to\infty} g^{(\nu_\tau+1)\mathrm{T}} G(x^{(\nu_\tau)})^{-1} g^{(\nu_\tau+1)} = \lim_{\tau\to\infty} \alpha^{(\nu_\tau)\mathrm{T}} P(x^{(\nu_\tau)}) \alpha^{(\nu_\tau)} = 0, \tag{3.1}$$

where

$$\alpha^{(\nu)} = [\xi^{(\nu)}]\eta^{(\nu+1)}. \tag{3.2}$$

On the other hand, from lemma 2.5, we have

$$\begin{aligned} g^{(\nu_\tau+1)\mathrm{T}} G(x^{(\nu_\tau)})^{-1} g^{(\nu_\tau+1)} &= \alpha^{(\nu_\tau)\mathrm{T}} P^{(\nu_\tau)} \alpha^{(\nu_\tau)} \\ &= \alpha_E^{(\nu_\tau)\mathrm{T}} (\hat{P}_{EE}^{(\nu_\tau)} + \Delta P_{EE}^{(\nu_\tau)}) \alpha_E^{(\nu_\tau)\mathrm{T}} + 2\alpha_E^{(\nu_\tau)\mathrm{T}} \Delta P_{E\bar{B}}^{(\nu_\tau)} \alpha_{\bar{B}}^{(\nu_\tau)} \\ &\quad + \alpha_{\bar{B}}^{(\nu_\tau)\mathrm{T}} (I + S_{\bar{B}N}^{(\nu_\tau)} S_{\bar{B}N}^{(\nu_\tau)\mathrm{T}})^{-1} \alpha_{\bar{B}}^{(\nu_\tau)} + \alpha_{\bar{B}}^{(\nu_\tau)\mathrm{T}} \Delta P_{\overline{BB}}^{(\nu_\tau)} \alpha_{\bar{B}}^{(\nu_\tau)}. \end{aligned} \tag{3.3}$$

Since $\Phi_{E(\mathscr{X})}(x^{(\nu_\tau)}) \to 0$, $\|\alpha_{E(\mathscr{X})}^{(\nu_\tau)}\| \to 0$ and $\|\alpha^{(\nu_\tau)}\|$ is bounded, we see, by using lemma

2.5, that the first, the second and the fourth term on the rightmost hand side converges to zero, and hence, the third term, which is bounded below by

$$\alpha_{\bar{B}}^{(\nu_\tau)\mathrm{T}}(I+S_{\bar{B}N}^{(\nu_\tau)}S_{\bar{B}N}^{(\nu_\tau)\mathrm{T}})^{-1}\alpha_{\bar{B}}^{(\nu_\tau)} = \eta_{\bar{B}}^{(\nu_\tau+1)\mathrm{T}}([\xi_{\bar{B}}^{(\nu_\tau)}]^{-2}+\bar{A}_{\bar{B}}A_N[\xi_N^{(\nu_\tau)}]^{-2}A_N^t\bar{A}_{\bar{B}}^{\mathrm{T}})^{-1}\eta_{\bar{B}}^{(\nu_\tau+1)}$$

$$\geq M\left\{\min_{i\notin E(\mathscr{X})}\xi_i^{(\nu_\tau)}\right\}^2\|\eta_{\bar{B}(\mathscr{X})}^{(\nu_\tau+1)}\|^2, \tag{3.4}$$

where M is an appropriate positive constant, also converges to zero. Since $\min_{i\notin E(\mathscr{X})}\xi_i^{(\nu_\tau)}$ converges to a positive number, we see that $\eta_{\bar{B}(\mathscr{X})}^{(\nu_\tau+1)}$ tends to 0 as $\tau\to\infty$. From proposition 2.4, we see that each component of $\xi_{E(\mathscr{X})}^{(\nu_\tau)}$ is multiplied at most by the factor of $1+\mu^{(\nu_\tau)}$ (<2), and hence $\xi_{E(\mathscr{X})}^{(\nu_\tau)}\to 0$ ($\tau\to\infty$) immediately implies $\xi_{E(\mathscr{X})}^{(\nu_\tau+1)}\to 0$, then, every accumulation point $\tilde{x}^*$ of $\{x^{(\nu_\tau+1)}\}$ satisfies the conditions

$$\eta_{\bar{B}(\mathscr{X})}(\tilde{x}^*) = \nabla_{\bar{B}(\mathscr{X})}F(\tilde{x}^*) = 0, \quad \xi_{\bar{B}(\mathscr{X})}(\tilde{x}^*)\geq 0, \quad \xi_{B(\mathscr{X})}(\tilde{x}^*) = 0. \tag{3.5}$$

Since (3.5) implies that $\tilde{x}^*$ is the unique face-optimal point where F is minimized over $\mathscr{X}$, the accumulation point of $x^{(\nu_\tau+1)}$ is unique, thus $x^{(\nu_\tau+1)}$ converges to $\tilde{x}^*$. We show $x^*=\tilde{x}^*$. If not, since both x^* and $\tilde{x}^*$ are on $\mathscr{X}$, we have $F(x^*)>F(\tilde{x}^*)$, then $F(x^*)$ cannot be an accumulation point. Hence, we have $x^*=\tilde{x}^*$.

From the discussion above, we see that, for any subsequence $\{x^{(\nu_\tau)}\}$ convergent to x^*, $\|x^{(\nu_\tau+1)}-x^{(\nu_\tau)}\|$ converges to zero as $\tau\to\infty$. This means that, for any given $\delta>0$, we can choose $\epsilon>0$ such that

$$\|x^{(\nu)}-x^*\|\leq\epsilon\Rightarrow\|x^{(\nu+1)}-x^{(\nu)}\|\leq\delta. \tag{3.6}$$

Thus every accumulation point is an optimal solution for some subproblem where the displacement vector converges to zero as the iterate approaches the accumulation point. Since the number of subproblems is finite, the number of accumulation points also is finite.

Now, we assume that there are two accumulation points x_1^* and x_2^*. Taking note of $F(x_1^*)=F(x_2^*)$ and the fact that the displacement vector converges to zero as the sequence approaches x_1^* and x_2^*, we see that there exists an accumulation point that is not a face-optimal point. However, this contradicts the fact that every accumulation point is a face-optimal point. Thus the sequence has a unique accumulation point. □

In the remaining part of this paper we denote the limiting point of $\{x^{(\nu)}\}$ by x^* and $F(x^*)$ by F^*. We investigate the properties of $\{x^{(\nu)}\}$ when it converges to x^*.

LEMMA 3.2

Let $\{x^{(\nu)}\}$ be the sequence generated by the affine scaling algorithm under the assumption (**A1**), and let x^* be the limiting point of the sequence, which is an interior point of the face $\mathscr{X}$. Choose a slack coordinate $(\xi_{B(\mathscr{X})}, \xi_{\bar{B}(\mathscr{X})})$ associated with $E(\mathscr{X})$. Let $(\theta_B, \theta_{\bar{B}})$ be the local coordinate associated with $(x^*, (\xi_B, \xi_{\bar{B}}))$. We have

$$\frac{\theta_{\bar{B}}^{(\nu+1)\mathrm{T}} H_{\overline{BB}} \theta_{\bar{B}}^{(\nu+1)}}{\|\xi_{E(\mathscr{X})}^{(\nu)}\|} \le M \tag{3.7}$$

for sufficiently large ν, where $M > 0$ is a constant.

Proof

In terms of the local coordinate, the objective function is represented as

$$F(x^{(\nu)}) = \tfrac{1}{2}\theta_B(x^{(\nu)})\tilde{H}_{BB}\theta_B(x^{(\nu)}) + \tfrac{1}{2}\theta_{\bar{B}}(x^{(\nu)})^{\mathrm{T}} H_{\overline{BB}}\theta_{\bar{B}}(x^{(\nu)}) + \eta_B^{*\mathrm{T}}\theta_B(x^{(\nu)}) + F^*. \tag{3.8}$$

Consider the point $y^{(\nu)}$ such that

$$(\theta_B(y^{(\nu)}), \theta_{\bar{B}}(y^{(\nu)})) = (\theta_B(x^{(\nu)}), 0) = (\xi_B(x^{(\nu)}), 0), \tag{3.9}$$

and let $\Omega^{(\nu)} = \{x \mid \|[\xi^{(\nu)}]^{-1}(\xi(x) - \xi^{(\nu)})\| \le \mu^{(\nu)}\}$.

For the time being, we assume that $y^{(\nu)} \in \Omega^{(\nu)}$ for sufficiently large ν, and observe that this implies the lemma. Since $\theta_{\bar{B}}(y^{(\nu)}) = 0$ and $\theta_B(y^{(\nu)}) = \theta_B(x^{(\nu)})$, from lemma 2.3, we have

$$F(y^{(\nu)}) = \tfrac{1}{2}\theta_B(x^{(\nu)})^{\mathrm{T}}\tilde{H}_{BB}\theta_B(x^{(\nu)}) + \eta_B^{*\mathrm{T}}\theta_B(x^{(\nu)}) + F^*. \tag{3.10}$$

Since $x^{(\nu+1)}$ is the point that minimizes F over $\Omega^{(\nu)}$ (cf. (1.4)), we obtain

$$\begin{aligned} F(x^{(\nu+1)}) &= \tfrac{1}{2}\theta_B(x^{(\nu+1)})^{\mathrm{T}}\tilde{H}_{BB}\theta_B(x^{(\nu+1)}) + \tfrac{1}{2}\theta_{\bar{B}}(x^{(\nu+1)})^{\mathrm{T}} H_{\overline{BB}}\theta_{\bar{B}}(x^{(\nu+1)}) \\ &\quad + \eta_B^{*\mathrm{T}}\theta_B(x^{(\nu+1)}) + F^* \le F(y^{(\nu)}) \\ &= \tfrac{1}{2}\theta_B(x^{(\nu)})^{\mathrm{T}}\tilde{H}_{BB}\theta_B(x^{(\nu)}) + \eta_B^{*\mathrm{T}}\theta_B(x^{(\nu)}) + F^*. \end{aligned} \tag{3.11}$$

From proposition 2.4, we have

$$\|\theta_B(x^{(\nu+1)})\| = \|\xi_B(x^{(\nu+1)})\| \le (1 + \mu^{(\nu)})\|\xi_B(x^{(\nu)})\| = (1 + \mu^{(\nu)})\|\theta_B(x^{(\nu)})\|, \tag{3.12}$$

and, together with (3.11), this implies that

$$\tfrac{1}{2}\theta_{\bar{B}}(x^{(\nu+1)})^{\mathrm{T}} H_{\overline{BB}}\theta_{\bar{B}}(x^{(\nu+1)}) \le M\|\xi_E(x^{(\nu)})\|, \tag{3.13}$$

where M is a positive constant, which is the desired result.

We complete the proof by showing $y^{(\nu)} \in \Omega^{(\nu)}$ for sufficiently large ν. To this end, observe the following relation:

$$\begin{aligned}\xi_{\bar{B}}(y^{(\nu)}) - \xi_{\bar{B}}(x^*) &= \theta_{\bar{B}}(y^{(\nu)}) - H_{\overline{BB}}^{-1} H_{B\bar{B}}^{\mathrm{T}}\theta_B(y^{(\nu)})\\ &= \theta_{\bar{B}}(y^{(\nu)}) - H_{\overline{BB}}^{-1} H_{B\bar{B}}^{\mathrm{T}}\theta_B(x^{(\nu)})\\ &= -H_{\overline{BB}}^{-1} H_{B\bar{B}}^{\mathrm{T}}\xi_B(x^{(\nu)}).\end{aligned} \tag{3.14}$$

Since $\xi_E(y^{(\nu)}) - \xi_E(x^{(\nu)}) \in \mathrm{Im}\,(A_E^{\mathrm{T}})$, it follows from (2.3) that

$$\xi_E(y^{(\nu)}) - \xi_E(x^{(\nu)}) = \begin{pmatrix} T_{BR}^{\mathrm{T}} \\ I \end{pmatrix}(\xi_B(y^{(\nu)}) - \xi_B(x^{(\nu)})) = 0. \tag{3.15}$$

With this relation, we see:

$$\begin{aligned}&\|[\xi(x^{(\nu)})]^{-1}(\xi(y^{(\nu)}) - \xi(x^{(\nu)}))\|\\ &\quad\le \|[\xi_E(x^{(\nu)})]^{-1}(\xi_E(y^{(\nu)}) - \xi_E(x^{(\nu)}))\| + \|[\xi_{E^c}(x^{(\nu)})]^{-1}(\xi_{E^c}(y^{(\nu)}) - \xi_{E^c}(x^{(\nu)}))\|\\ &\quad= \|[\xi_{E^c}(x^{(\nu)})]^{-1}\|\,\|\xi_{E^c}(y^{(\nu)}) - \xi_{E^c}(x^{(\nu)})\|\\ &\quad\le \|[\xi_{E^c}(x^{(\nu)})]^{-1}\|\left\|A_{E^c}^{\mathrm{T}}(A_{B\cup\bar{B}}^{\mathrm{T}})^{-1}\begin{pmatrix}\xi_B(y^{(\nu)}) - \xi_B(x^{(\nu)})\\ \xi_{\bar{B}}(y^{(\nu)}) - \xi_{\bar{B}}(x^{(\nu)})\end{pmatrix}\right\|\\ &\quad= \|[\xi_{E^c}(x^{(\nu)})]^{-1}\|\left\|A_{E^c}^{\mathrm{T}}(A_{B\cup\bar{B}}^{\mathrm{T}})^{-1}\begin{pmatrix}0\\ \xi_{\bar{B}}(y^{(\nu)}) - \xi_{\bar{B}}(x^{(\nu)})\end{pmatrix}\right\|\\ &\quad\le \|[\xi_{E^c}(x^{(\nu)})]^{-1}\|\,\|A_{E^c}^{\mathrm{T}}(A_{B\cup\bar{B}}^{\mathrm{T}})^{-1}\|\,\|\xi_{\bar{B}}(y^{(\nu)}) - \xi_{\bar{B}}(x^{(\nu)})\|\\ &\quad\le \|[\xi_{E^c}(x^{(\nu)})]^{-1}\|\,\|A_{E^c}^{\mathrm{T}}(A_{B\cup\bar{B}}^{\mathrm{T}})^{-1}\|\,\{\|\xi_{\bar{B}}(y^{(\nu)}) - \xi_{\bar{B}}(x^*)\| + \|\xi_{\bar{B}}(x^*) - \xi_{\bar{B}}(x^{(\nu)})\|\}.\end{aligned} \tag{3.16}$$

Since the last expression in (3.16) converges to zero because of (3.14) and $\xi(x^{(\nu)}) \to \xi^*$, we have $\|[\xi^{(\nu)}]^{-1}(\xi(y^{(\nu)}) - \xi(x^{(\nu)}))\| \to 0$. This implies $y^{(\nu)} \in \Omega^{(\nu)}$ for sufficiently large ν, and completes the proof. □

The following corollary immediately follows from lemma 3.2 and proposition 2.4.

COROLLARY 3.3

Under the assumptions and the notations of lemma 3.2, there exists a constant $\delta > 0$ such that

$$\theta_{\bar{B}}^{(\nu)\mathrm{T}} H_{\bar{B}\bar{B}} \theta_{\bar{B}}^{(\nu)} \leq \delta \|\xi_{E(\mathscr{X})}^{(\nu)}\|. \tag{3.17}$$

LEMMA 3.4

Let $\{x^{(\nu)}\}$ be the sequence generated by the affine scaling algorithm under the assumption **(A1)**, and let x^* be the limiting point of the sequence which is an interior point of the face $\mathscr{X}$. Choose a slack coordinate $(\xi_{B(\mathscr{X})}, \xi_{\bar{B}(\mathscr{X})})$ associated with $E(\mathscr{X})$, and let $(\theta_B, \theta_{\bar{B}})$ be the local coordinate associated with $(x^*, (\xi_B, \xi_{\bar{B}}))$.

There exists a positive constant δ' such that

$$\|\theta_{\bar{B}}^{(\nu+1)}\| \leq \delta' \|\xi_{E(\mathscr{X})}^{(\nu)}\| \tag{3.18}$$

for all ν.

Proof

If such a δ' does not exist, we can find a subsequence $\{x^{(\nu_\tau)}\}$ where

$$\frac{\|\theta_{\bar{B}}^{(\nu_\tau+1)}\|}{\|\xi_E^{(\nu_\tau)}\|} \to \infty. \tag{3.19}$$

We show that there exists a positive constant δ satisfying

$$\frac{g^{(\nu_\tau+1)} G(x^{(\nu_\tau)})^{-1} g^{(\nu_\tau+1)}}{\|\theta_{\bar{B}}^{(\nu_\tau+1)}\|^2} = \frac{\alpha^{(\nu_\tau)\mathrm{T}} P^{(\nu_\tau)} \alpha^{(\nu_\tau)}}{\|\theta_{\bar{B}}^{(\nu_\tau+1)}\|^2} \geq \delta \tag{3.20}$$

for all τ, where $\alpha^{(\nu_\tau)} = [\xi^{(\nu_\tau)}]\eta^{(\nu_\tau+1)}$. If such a δ does not exist, we can choose a subsequence $\{x^{(\nu_\sigma)}\}$ of $\{x^{(\nu_\tau)}\}$ such that

$$\frac{\alpha^{(\nu_\sigma)\mathrm{T}} P^{(\nu_\sigma)} \alpha^{(\nu_\sigma)}}{\|\theta_{\bar{B}}^{(\nu_\sigma+1)}\|^2} \to 0 \tag{3.21}$$

as $\sigma \to \infty$. Let $\tilde{\alpha}^{(\nu_\sigma)} = \alpha^{(\nu_\sigma)} / \|\theta_{\bar{B}}^{(\nu_\sigma+1)}\|$. Since

$$\|\tilde{\alpha}_E^{(\nu_\sigma)}\| \leq \|\eta_E^{(\nu_\sigma+1)}\| \frac{\|\xi_E^{(\nu_\sigma)}\|}{\|\theta_{\bar{B}}^{(\nu_\sigma+1)}\|}, \tag{3.22}$$

we see, by using (3.19),

$$\tilde{\alpha}_E^{(\nu_\sigma)} \to 0 \tag{3.23}$$

as $\sigma \to \infty$. On the other hand, because of $\theta_{\bar{B}} = H_{\overline{BB}}^{-1} \eta_{\bar{B}}$ (cf. (2.12)), $\|\tilde{\alpha}_{\bar{B}}^{(\nu_\sigma)}\|$ is bounded by a constant as follows:

$$\|\tilde{\alpha}_{\bar{B}}^{(\nu_\sigma)}\| \leq \|\eta_{\bar{B}}^{(\nu_\sigma+1)}\| \frac{\|\xi_{E^c}^{(\nu_\sigma)}\|}{\|\theta_{\bar{B}}^{(\nu_\sigma+1)}\|} \leq \|H_{\overline{BB}} \theta_{\bar{B}}^{(\nu_\sigma+1)}\| \frac{\|\xi_{E^c}^{(\nu_\sigma)}\|}{\|\theta_{\bar{B}}^{(\nu_\sigma+1)}\|} \leq \|H_{\overline{BB}}\| \, \|\xi_{E^c}^{(\nu_\sigma)}\|. \tag{3.24}$$

Applying lemma 2.5, we obtain

$$\begin{aligned} \tilde{\alpha}^{(\nu_\sigma)\mathrm{T}} P^{(\nu_\sigma)} \tilde{\alpha}^{(\nu_\sigma)} &= \tilde{\alpha}_E^{(\nu_\sigma)\mathrm{T}} (\hat{P}_{EE}^{(\nu_\sigma)} + \Delta P_{EE}^{(\nu_\sigma)}) \tilde{\alpha}_E^{(\nu_\sigma)} + 2 \tilde{\alpha}_E^{(\nu_\sigma)\mathrm{T}} \Delta P_{E\bar{B}}^{(\nu_\sigma)} \tilde{\alpha}_{\bar{B}}^{(\nu_\sigma)} \\ &\quad + \tilde{\alpha}_{\bar{B}}^{(\nu_\sigma)\mathrm{T}} \{(I + S_{\bar{B}N}^{(\nu_\sigma)} S_{\bar{B}N}^{(\nu_\sigma)\mathrm{T}})^{-1} + \Delta P_{\overline{BB}}^{(\nu_\sigma)}\} \tilde{\alpha}_{\bar{B}}^{(\nu_\sigma)}. \end{aligned} \tag{3.25}$$

Taking note of the fact $\Phi_{E(\mathscr{X})}^{(\nu)} \to 0$ and using (3.23), (3.24), we see the first and the second term on the right hand side converge to zero. Since $\|\Delta P_{\overline{BB}}^{(\nu_\sigma)}\|$ converges to zero while the minimum singular value of $(I + S_{\bar{B}N}^{(\nu_\sigma)} S_{\bar{B}N}^{(\nu_\sigma)\mathrm{T}})^{-1}$ is bounded below by a constant (this follows from the definition of $S_{\bar{B}N}$ and $\xi_{\bar{B}}^{(\nu)}$, $\xi_N^{(\nu)}$ is uniformly bounded below by a positive constant), the third term is bounded as follows:

$$\begin{aligned} &\tilde{\alpha}_{\bar{B}}^{(\nu_\sigma)\mathrm{T}} \{(I + S_{\bar{B}N}^{(\nu_\sigma)} S_{\bar{B}N}^{(\nu_\sigma)\mathrm{T}})^{-1} + \Delta P_{\overline{BB}}^{(\nu_\sigma)}\} \tilde{\alpha}_{\bar{B}}^{(\nu_\sigma)} \\ &\geq \tfrac{1}{2} \tilde{\alpha}_{\bar{B}}^{(\nu_\sigma)\mathrm{T}} (I + S_{\bar{B}N}^{(\nu_\sigma)} S_{\bar{B}N}^{(\nu_\sigma)\mathrm{T}})^{-1} \tilde{\alpha}_{\bar{B}}^{(\nu_\sigma)} \\ &= \frac{1}{2} \frac{\eta_{\bar{B}}^{(\nu_\sigma+1)\mathrm{T}}}{\|\theta_{\bar{B}}^{(\nu_\sigma+1)}\|} ([\xi_{\bar{B}}^{(\nu_\sigma)}]^{-2} + \bar{A}_{\bar{B}} A_N [\xi_N^{(\nu_\sigma)}]^{-2} A_N^{\mathrm{T}} \bar{A}_{\bar{B}}^{\mathrm{T}})^{-1} \frac{\eta_{\bar{B}}^{(\nu_\sigma+1)}}{\|\theta_{\bar{B}}^{(\nu_\sigma+1)}\|} \\ &\geq M_1 \frac{\|\eta_{\bar{B}}^{(\nu_\sigma+1)}\|^2}{\|\theta_{\bar{B}}^{(\nu_\sigma+1)}\|^2} \left\{ \min_{i \notin E(\mathscr{X})} \xi^{(\nu_\sigma)} \right\}^2 \\ &\geq M_2 \frac{\|\eta_{\bar{B}}^{(\nu_\sigma+1)}\|^2}{\|\theta_{\bar{B}}^{(\nu_\sigma+1)}\|^2} = M_2 \frac{\|\eta_{\bar{B}}^{(\nu_\sigma+1)}\|^2}{\|H_{\overline{BB}}^{-1} \eta_{\bar{B}}^{(\nu_\sigma+1)}\|^2} \geq M_3 > 0, \end{aligned} \tag{3.26}$$

where M_1, M_2, M_3 are appropriate positive constants, contradicting (3.21). Thus, we see (3.20) holds for all τ by choosing δ appropriately.

Now the lemma is readily seen as follows. Due to lemma 2.3 and corollary 3.3, we have

$$F(x^{(\nu)}) - F^* = O(\|\xi_E^{(\nu)}\|), \tag{3.27}$$

where $F^* = F(x^*)$. Together with (3.20), we see, by using (1.7),

$$\begin{aligned} 0 \le F(x^{(\nu_\tau+1)}) - F^* &\le F(x^{(\nu_\tau)}) - F^* - \mu^{(\nu_\tau)}\{\alpha^{(\nu_\tau)\mathrm{T}}P^{(\nu_\tau)}\alpha^{(\nu_\tau)}\}^{1/2} \\ &\le M_4\|\xi_E^{(\nu_\tau)}\| - \mu_{\min}\delta^{1/2}\|\theta_{\bar{B}}^{(\nu_\tau+1)}\|, \end{aligned} \tag{3.28}$$

where M_4 is an appropriate positive constant. However, this contradicts (3.19). Thus, (3.18) holds by choosing $\delta' > 0$ appropriately, completing the proof. □

The following corollary immediately follows from lemma 3.4 and proposition 2.4.

COROLLARY 3.5

Under the assumptions and notations of lemma 3.4, there exists a positive constant δ such that

$$\|\theta_{\bar{B}}^{(\nu)}\| \le \delta\|\xi_{E(\mathscr{X})}^{(\nu)}\| \tag{3.29}$$

holds for all ν.

The last two results in this section are proved under the dual nondegeneracy assumption (**A2**).

LEMMA 3.6

Let $\{x^{(\nu)}\}$ be the sequence generated by the affine scaling algorithm under the assumptions (**A1**) and (**A2**), and let x^* be the limiting point of the sequence. Denote by $\mathscr{X}$ the face that contains x^* in its interior. Take the slack coordinate $(\xi_{B(\mathscr{X})}, \xi_{\bar{B}(\mathscr{X})})$ associated with $E(\mathscr{X})$. Then

$$g^{(\nu+1)\mathrm{T}}G(x^{(\nu)})^{-1}g^{(\nu+1)} \ge \delta''\|\xi_{E(\mathscr{X})}^{(\nu)}\|^2, \tag{3.30}$$

where δ'' is a positive constant.

Proof

Let

$$\gamma^{(\nu)} = \frac{[\xi^{(\nu)}]\eta^{(\nu+1)}}{\|\xi^{(\nu)}_{E(\mathscr{X})}\| + \|\eta^{(\nu+1)}_{\bar{B}(\mathscr{X})}\|}. \tag{3.31}$$

Take the local coordinate $(\theta_B, \theta_{\bar{B}})$ associated with $(x^*, (\xi_B, \xi_{\bar{B}}))$. From lemma 3.4 and (2.12), we see that

$$\begin{aligned}\|\xi^{(\nu)}_{E(\mathscr{X})}\| + \|\eta^{(\nu+1)}_{\bar{B}}\| &= \|\xi^{(\nu)}_{E(\mathscr{X})}\| + \|H_{\overline{BB}}\theta^{(\nu+1)}_{\bar{B}}\| \\ &\le \|\xi^{(\nu)}_{E(\mathscr{X})}\| + \|H_{\overline{BB}}\|\,\|\theta^{(\nu+1)}_{\bar{B}}\| \\ &= O(\|\xi^{(\nu)}_{E(\mathscr{X})}\|).\end{aligned} \tag{3.32}$$

It is enough to show that

$$\frac{\eta^{(\nu+1)\mathrm{T}}[\xi^{(\nu)}]P^{(\nu)}[\xi^{(\nu)}]\eta^{(\nu+1)}}{(\|\xi^{(\nu)}_{E(\mathscr{X})}\| + \|\eta^{(\nu+1)}_{\bar{B}(\mathscr{X})}\|)^2} = \frac{g^{(\nu+1)\mathrm{T}}G(x^{(\nu)})^{-1}g^{(\nu+1)}}{(\|\xi^{(\nu)}_{E(\mathscr{X})}\| + \|\eta^{(\nu+1)}_{\bar{B}(\mathscr{X})}\|)^2} = \gamma^{(\nu)\mathrm{T}}P^{(\nu)}\gamma^{(\nu)} \tag{3.33}$$

is bounded below by a positive constant.

We have

$$\|\gamma^{(\nu)}\| \le \frac{\|\xi^{(\nu)}_{E(\mathscr{X})}\|\,\|\eta^{(\nu+1)}_{B(\mathscr{X})}\| + \|\xi^{(\nu)}_{\bar{B}(\mathscr{X})}\|\,\|\eta^{(\nu+1)}_{\bar{B}(\mathscr{X})}\|}{\|\xi^{(\nu)}_{E(\mathscr{X})}\| + \|\eta^{(\nu+1)}_{\bar{B}(\mathscr{X})}\|} \le \|\eta^{(\nu+1)}_{E(\mathscr{X})}\| + \|\xi^{(\nu)}_{\bar{B}(\mathscr{X})}\|. \tag{3.34}$$

By contradiction, assume that a subsequence $\{x^{(\nu_\sigma)}\}$ exists where $\gamma^{(\nu_\sigma)\mathrm{T}}P^{(\nu_\sigma)}\gamma^{(\nu_\sigma)} \to 0$ as $\sigma \to \infty$. We can choose a subsequence $\{x^{(\nu_\tau)}\}$ from $\{x^{(\nu_\sigma)}\}$ such that each

$$s_i^{(\nu_\tau)} = \frac{\xi_i^{(\nu_\tau)}}{\|\xi^{(\nu_\tau)}_{E(\mathscr{X})}\| + \|\eta^{(\nu_\tau+1)}_{\bar{B}(\mathscr{X})}\|} \tag{3.35}$$

either converges to a finite value or diverges to infinity. Let Y be the index set consisting of the indices such that the limit of $\{s_i^{(\nu_\tau)}\}$ is zero. By using lemma 2.1, we see that there exists a face $\mathscr{Y}$ such that $E(\mathscr{Y}) = Y$. (If $Y = \emptyset$, we have $\mathscr{Y} = \mathscr{P}$.) Recall that $\mathscr{X}$ is contained in every face that contains the point x^*, and hence that we have $\mathscr{Y} \supseteq \mathscr{X}$. In view of (3.32), we have $\xi^{(\nu_\tau)}_{E(\mathscr{Y})}/\|\xi^{(\nu_\tau)}_{E(\mathscr{X})}\| \to 0$, thus we have $\mathscr{Y} \supset \mathscr{X}$.

Let us choose another slack coordinate $(\xi_{B(\mathscr{Y})}, \xi_{\bar{B}(\mathscr{Y})})$ associated with $E(\mathscr{Y})$. We simply denote $E(\mathscr{Y})$, $B(\mathscr{Y})$, $\bar{B}(\mathscr{Y})$, $N(\mathscr{Y})$ by E, B, $\bar{B}$, N, respectively. (Here we note that the argument below holds even in the "special" case $\mathscr{Y} = \mathscr{P}$.) By using lemma 2.5, we have

$$\gamma^{(\nu_\tau)\mathrm{T}} P^{(\nu_\tau)} \gamma^{(\nu_\tau)} = \gamma_E^{(\nu_\tau)\mathrm{T}} \hat{P}_{EE}^{(\nu_\tau)} \gamma_E^{(\nu_\tau)} + (\gamma_{\bar{B}}^{(\nu_\tau)\mathrm{T}} \gamma_N^{(\nu_\tau)\mathrm{T}}) \hat{P}_{E^cE^c}^{(\nu_\tau)} \begin{pmatrix} \gamma_{\bar{B}}^{(\nu_\tau)} \\ \gamma_N^{(\nu_\tau)} \end{pmatrix} + \gamma^{(\nu_\tau)} \Delta P^{(\nu_\tau)} \gamma^{(\nu_\tau)}. \tag{3.36}$$

Since $\Phi_{E(\mathscr{Y})}^{(\nu_\tau)} \to 0$ and $\|\gamma^{(\nu_\tau)}\|$ is bounded above as was shown in (3.34), it follows from lemma 2.5 that the first and the third terms on the right hand side converge to zero, which implies the convergence of the second term to zero. Substituting the definition of $\gamma^{(\nu_\tau)}$, $S_{\bar{B}N}^{(\nu_\tau)}$, the second term is bounded below by

$$\begin{aligned}
&(\gamma_{\bar{B}}^{(\nu_\tau)} \quad \gamma_N^{(\nu_\tau)})^{\mathrm{T}} \hat{P}_{E^cE^c}^{(\nu_\tau)} \begin{pmatrix} \gamma_{\bar{B}}^{(\nu_\tau)} \\ \gamma_N^{(\nu_\tau)} \end{pmatrix} \\
&= (\gamma_{\bar{B}}^{(\nu_\tau)} + S_{\bar{B}N}^{(\nu_\tau)} \gamma_N^{(\nu_\tau)})^{\mathrm{T}} (I + S_{\bar{B}N}^{(\nu_\tau)} S_{\bar{B}N}^{(\nu_\tau)\mathrm{T}})^{-1} (\gamma_{\bar{B}}^{(\nu_\tau)} + S_{\bar{B}N}^{(\nu_\tau)} \gamma_N^{(\nu_\tau)}) \\
&= (\eta_{\bar{B}}^{(\nu_\tau+1)} + \bar{A}_{\bar{B}} A_N \eta_N^{(\nu_\tau+1)})^{\mathrm{T}} \\
&\quad \times ([s_{\bar{B}}^{(\nu_\tau)}]^{-2} + \bar{A}_{\bar{B}} A_N [s_N^{(\nu_\tau)}]^{-2} A_N^{\mathrm{T}} \bar{A}_{\bar{B}}^{\mathrm{T}})^{-1} (\eta_{\bar{B}}^{(\nu_\tau+1)} + \bar{A}_{B} A_N \eta_N^{(\nu_\tau+1)}) \\
&\geq M \left\{ \min_{i \notin E(\mathscr{Y})} s_i^{(\nu_\tau)} \right\}^2 \|\eta_{\bar{B}}^{(\nu_\tau+1)} + \bar{A}_{\bar{B}} A_N \eta_N^{(\nu_\tau+1)}\|^2,
\end{aligned} \tag{3.37}$$

where M is an appropriate positive constant. Since $\min_{i \notin E(\mathscr{Y})} s_i^{(\nu_\tau)}$ converges to a positive number, we have

$$\|\eta_{\bar{B}(\mathscr{Y})}^{(\nu_\tau+1)} + \bar{A}_{\bar{B}(\mathscr{Y})} A_{N(\mathscr{Y})} \eta_{N(\mathscr{Y})}^{(\nu_\tau+1)}\|^2 \to 0 \tag{3.38}$$

as $\tau \to \infty$. Since $\eta^* = \lim_{\tau\to\infty} \eta^{(\nu_\tau+1)}$, we have

$$A_{\bar{B}(\mathscr{Y})} \eta_{\bar{B}(\mathscr{Y})}^* + A_{\bar{B}(\mathscr{Y})} \bar{A}_{\bar{B}(\mathscr{Y})} A_{N(\mathscr{Y})} \eta_{N(\mathscr{Y})}^* = 0. \tag{3.39}$$

Subtracting (3.39) from

$$A\eta^* = A_{B(\mathscr{Y})} \eta_{B(\mathscr{Y})}^* + A_{B(\mathscr{Y})} \eta_{\bar{B}(\mathscr{Y})}^* + A_{N(\mathscr{Y})} \eta_{N(\mathscr{Y})}^* \tag{3.40}$$

and using the second relation of (2.5), we see

$$A\eta^* \in \mathrm{Im}\,(A_{E(\mathscr{Y})}), \tag{3.41}$$

which implies that the linear function $\eta^{*\mathrm{T}}\xi(x)$ $(= (\partial F(x^*)/\partial x)^{\mathrm{T}}(x - x^*))$ takes a constant value over $\mathscr{Y} \supset \mathscr{X}$ (if $\mathscr{Y} = \mathscr{P}$, then we have $A\eta^* = 0$). However, this is a contradiction to the assumption (**A2**), which comes from assuming the existence of the subsequence such that $\gamma^{(\nu_\sigma)\mathrm{T}}P^{(\nu_\sigma)}\gamma^{(\nu_\sigma)} \to 0$ as $\sigma \to \infty$. This means that $\gamma^{(\nu)\mathrm{T}}P^{(\nu)}\gamma^{(\nu)}$ is uniformly bounded below by a positive constant, completing the proof of the lemma. □

LEMMA 3.7

Let $\{x^{(\nu)}\}$ be the sequence generated by the affine scaling algorithm under the assumptions (**A1**), (**A2**), and let x^* be the limiting point of the sequence. Denote by $\mathscr{X}$ the face that contains x^* in its interior. Choose a slack coordinate $(\xi_{B(\mathscr{X})}, \xi_{\bar{B}(\mathscr{X})})$ associated with $E(\mathscr{X})$. Then, for sufficiently large ν, we have

$$\frac{\eta^{*\mathrm{T}}\xi^{(\nu)}}{\|\xi^{(\nu)}_{E(\mathscr{X})}\|} \geq \delta > 0, \tag{3.42}$$

where δ is a positive constant.

Proof

We take the local coordinate $(\theta_B, \theta_{\bar{B}})$ associated with $(x^*, (\xi_B, \xi_{\bar{B}}))$. Due to lemma 2.3, lemma 3.6 and (1.7), we have

$$\begin{aligned} 0 &\leq \mu^{(\nu)}\delta\|\xi^{(\nu)}_{E(\mathscr{X})}\| \leq \mu^{(\nu)}\{g^{(\nu+1)\mathrm{T}}G(x^{(\nu)})^{-1}g^{(\nu+1)}\}^{1/2} \\ &\leq F(x^{(\nu)}) - F(x^{(\nu+1)}) \\ &\leq F(x^{(\nu)}) - F^* = \tfrac{1}{2}\theta_B^{(\nu)\mathrm{T}}\tilde{H}_{BB}\theta_B^{(\nu)} + \tfrac{1}{2}\theta_{\bar{B}}^{(\nu)\mathrm{T}}H_{\overline{BB}}\theta_{\bar{B}}^{(\nu)} + \eta_B^{*\mathrm{T}}\xi_B^{(\nu)}, \end{aligned} \tag{3.43}$$

where δ is a positive constant. Dividing this relation by $\|\xi_E^{(\nu)}\|$ and using corollary 3.5, we obtain (3.42) for sufficiently large ν, completing the proof. □

4. Proof of global convergence

In [10, 11], we introduced the local Karmarkar potential function to analyze the behavior of the affine scaling algorithm for linear programming in the vicinity of degenerate faces. In this section we show that substantially the same technique can be applied to show global convergence of the affine scaling algorithm for the strictly convex quadratic programming problems. For this purpose, we make use of the following lemma, which is immediately seen from lemma 5.1 of [14].

LEMMA 4.1

Let x^* be a face-optimal point of $\langle\mathbf{D}\rangle$, which is an interior point of the face $\mathscr{Y}$ of $\mathscr{P}$, and choose a slack coordinate $(\xi_{B(\mathscr{Y})}, \xi_{\bar{B}(\mathscr{Y})})$ associated with $E(\mathscr{Y})$. Let

$$\tilde{\eta}^* = (\tilde{\eta}^*_R, \tilde{\eta}^*_B, \tilde{\eta}^*_{\bar{B}}, \tilde{\eta}^*_N) = (0, \nabla_B F(x^*), \nabla_{\bar{B}} F(x^*), 0)$$
$$= (0, \nabla_B F(x^*), 0, 0). \tag{4.1}$$

Consider the linear programming problem

$$\text{minimize } \tilde{\eta}^{*\mathrm{T}}\xi, \quad \text{subject to } \xi = A^{\mathrm{T}}x - b \geq 0, \tag{4.2}$$

and define the local Karmarkar potential function associated with $\mathscr{Y}$ as follows:

$$f_{\mathscr{Y}}(\xi) = |E(\mathscr{Y})| \log(\tilde{\eta}^{*\mathrm{T}}\xi) - \sum_{i \in E(\mathscr{Y})} \log \xi_i$$
$$= |E(\mathscr{Y})| \log(\tilde{\eta}^{*\mathrm{T}}_{E(\mathscr{Y})}\xi_{E(\mathscr{Y})}) - \sum_{i \in E(\mathscr{Y})} \log \xi_i, \tag{4.3}$$

where $E(\mathscr{Y})$ is the set of indices of constraints satisfied with equality on $\mathscr{Y}$, and $|E(\mathscr{Y})|$ is the number of those constraints. We denote by $d\xi^{\mathrm{LP}}$ the unit displacement vector

$$d\xi^{\mathrm{LP}} = -\frac{[\xi]P[\xi]\tilde{\eta}^*}{\{\tilde{\eta}^{*\mathrm{T}}[\xi]P[\xi]\tilde{\eta}^*\}^{1/2}} \tag{4.4}$$

of the affine scaling algorithm for (4.2) represented in the space of the slack variables. If

(i) $\Phi_{E(\mathscr{Y})}(x)$ is sufficiently small,
(ii) x^* is not the optimal solution for $\langle \mathbf{D} \rangle$,
(iii) $\tilde{\eta}^{*\mathrm{T}}(\xi + \mu d\xi^{\mathrm{LP}}) > 0$,
(iv) $\mu \leq 11/80$,

then, we have

$$f_{\mathscr{Y}}(\xi + \mu d\xi^{\mathrm{LP}}) - f_{\mathscr{Y}}(\xi) \leq -\frac{\mu}{8}\left(1 - \sqrt{\frac{(|E(\mathscr{Y})| - 1)}{|E(\mathscr{Y})|}}\right) < 0. \tag{4.5}$$

Proof

The condition (ii) implies that x^* (or $\xi(x^*)$) is not an optimal solution for (4.2). If we replace the condition (iv) $\mu \leq 11/80$ for the upper bound of μ by $\mu \leq 1/8$, the lemma immediately follows from lemma 5.1 of [14]. Since the condition $\mu \leq 1/8$ in lemma 5.1 of [14] is not tight, we may loosen it a little bit to obtain this lemma, as we see by slightly modifying the final argument in the proof of lemma 5.1 of [14]. □

With the help of this lemma, we are now in a position to prove the main result.

THEOREM 1.1

Let $\langle \mathbf{D} \rangle$ be a problem satisfying the assumptions **(A1)** and **(A2)**, and apply the affine scaling algorithm with $0 < \mu_{\min} \leq \mu_{\max} \leq 1/8$. Then, the algorithm either (I) terminates after a finite number of iterations yielding the global minimum point of $F(x)$ over $\mathbb{R}^n$, or (II) generates an infinite sequence that converges to the optimal solution of $\langle \mathbf{D} \rangle$.

Proof

As noted at the end of section 1, we focus on the case (II). We denote by $\{x^{(\nu)}\}$ the infinite sequence generated by the algorithm. Due to lemma 4.1, the sequence converges to the limiting point x^* that is a face-optimal point. Denote by $\mathscr{X}$ the face that contains x^* in its interior. We assume that x^* is not the optimal solution for $\langle \mathbf{D} \rangle$, and derive a contradiction.

Let us take a slack coordinate $(\xi_{B(\mathscr{X})}, \xi_{\bar{B}(\mathscr{X})})$ associated with $E(\mathscr{X})$, and put $\xi^* = \xi(x^*)$. We denote $\eta(x^*)$ by η^*. Note that η^* is a constant vector with $\eta^*_{\bar{B}} = 0$ because x^* is a face-optimal point (cf. proposition 2.2), while $\eta(x)$ changes its value as a function of x, so that $\eta_{\bar{B}}(x)$ is not usually 0.

We already observed in lemma 3.7 and lemma 3.4 that

$$\frac{\eta^{*\mathrm{T}}\xi^{(\nu)}}{\|\xi^{(\nu)}_{E(\mathscr{X})}\|} = \frac{\eta_B^{*\mathrm{T}}\xi_B^{(\nu)}}{\|\xi^{(\nu)}_{E(\mathscr{X})}\|} \geq \delta \tag{4.6}$$

and

$$\|H_{\overline{BB}}\theta_{\bar{B}}^{(\nu+1)}\| = \|\eta^{(\nu+1)}_{\bar{B}(\mathscr{X})}\| \leq \delta' \|\xi^{(\nu)}_{E(\mathscr{X})}\| \tag{4.7}$$

hold for sufficiently large ν, where δ, δ' are positive constants. We further introduce the notations

$$\alpha^{*(\nu)} = [\xi^{(\nu)}]\eta^*, \quad \beta^{*(\nu)} = \frac{[\xi^{(\nu)}]\eta^*}{\xi^{(\nu)\mathrm{T}}\eta^*} \tag{4.8}$$

and

$$\alpha^{(\nu)} = [\xi^{(\nu)}]\eta^{(\nu+1)}, \quad \beta^{(\nu)} = \frac{[\xi^{(\nu)}]\eta^{(\nu+1)}}{\xi^{(\nu)\mathrm{T}}\eta^*} = \frac{\alpha^{(\nu)}}{\alpha^{*(\nu)\mathrm{T}}\mathbf{1}}, \tag{4.9}$$

where, by definition, we have

$$\alpha^{*(\nu)}_{E^c} = 0, \quad \beta^{*(\nu)}_{E^c} = 0 \quad \text{and} \quad \alpha^{(\nu)}_N = 0, \quad \beta^{(\nu)}_N = 0. \tag{4.10}$$

(NB: $\alpha^{(\nu)}_{\bar{B}} \neq 0, \beta^{(\nu)}_{\bar{B}} \neq 0$.) Note that $\beta^{(\nu)}$ and $\beta^{*(\nu)}$ are well-defined after a sufficiently large number of iterations because the denominator $\eta^{*\mathrm{T}}\xi^{(\nu)}$ is guaranteed to be strictly positive by lemma 3.7. Also define

$$d\beta^{(\nu)} = \beta^{(\nu)} - \beta^{*(\nu)} = \frac{\alpha^{(\nu)} - \alpha^{*(\nu)}}{\xi^{(\nu)\mathrm{T}}\eta^*} = [\xi^{(\nu)}]\frac{d\eta^{(\nu+1)}}{\xi^{(\nu)\mathrm{T}}\eta^*}, \tag{4.11}$$

where

$$d\eta = \begin{pmatrix} d\eta_R \\ d\eta_B \\ d\eta_{\bar{B}} \\ d\eta_N \end{pmatrix} = \begin{pmatrix} 0 \\ H_{BB}\xi_B + H_{B\bar{B}}d\xi_{\bar{B}} \\ \eta_{\bar{B}} \\ 0 \end{pmatrix} = \eta - \eta^* \tag{4.12}$$

and $d\xi = \xi - \xi^*$. By using (4.6), (4.7), we have

$$\|d\beta^{(\nu)}_{\bar{B}}\| \leq \frac{\|[\xi^{(\nu)}_{\bar{B}}]\| \, \|\eta^{(\nu+1)}_{\bar{B}}\|}{\delta\|\xi^{(\nu)}_E\|} \leq \frac{\|\xi^{(\nu)}_{\bar{B}}\|\delta'}{\delta} \leq M_0 \tag{4.13}$$

for sufficiently large ν, where M_0 is a constant. In the same manner, since $d\eta^{(\nu+1)} \to 0$, it is not difficult to see that

$$\|d\beta^{(\nu)}_E\| \to 0 \tag{4.14}$$

as $\nu \to \infty$. It follows from (4.6), (4.7) and (4.14) that $\|\beta^{*(\nu)}\|$ and $\|\beta^{(\nu)}\|$ are bounded as follows as $\nu \to \infty$:

$$\|\beta^{*(\nu)}\| = \frac{\|[\xi^{(\nu)}]\eta^*\|}{\xi^{(\nu)\mathrm{T}}\eta^*} \leq \frac{\|\xi^{(\nu)}_B\| \, \|\eta^*_B\|}{\delta\|\xi^{(\nu)}_E\|} \leq M_1, \tag{4.15}$$

$$\|\beta^{(\nu)}\| = \|\beta^{*(\nu)}\| + \|d\beta^{(\nu)}\| \leq M_1 + 1.1M_0, \tag{4.16}$$

where M_1 is a positive constant.

In terms of the notations above, at the current iterate $x^{(\nu)}$, the unit displacement vector $d\xi^{\mathrm{LP}(\nu)}$ of the affine scaling algorithm for "the LP problem"

$$\text{minimize } \eta^{*\mathrm{T}}\xi, \quad \text{subject to } \xi = A^{\mathrm{T}}x - b \geq 0 \tag{4.17}$$

is written, in the space of the slack variables, as follows:

$$d\xi^{\mathrm{LP}(\nu)} = -[\xi^{(\nu)}]\frac{P^{(\nu)}[\xi^{(\nu)}]\eta^*}{\{\eta^{*\mathrm{T}}[\xi^{(\nu)}]P^{(\nu)}[\xi^{(\nu)}]\eta^*\}^{1/2}} = -[\xi^{(\nu)}]\frac{P^{(\nu)}\alpha^{*(\nu)}}{\{\alpha^{*(\nu)\mathrm{T}}P^{(\nu)}\alpha^{*(\nu)}\}^{1/2}}$$

$$= -[\xi^{(\nu)}]\frac{P^{(\nu)}\beta^{*(\nu)}}{\{\beta^{*(\nu)\mathrm{T}}P^{(\nu)}\beta^{*(\nu)}\}^{1/2}}. \tag{4.18}$$

On the other hand, the unit displacement vector $d\xi^{\mathrm{QP}(\nu)}$ of the affine scaling algorithm for $\langle\mathbf{D}\rangle$ is written as

$$d\xi^{\mathrm{QP}(\nu)} = -[\xi^{(\nu)}]\frac{P^{(\nu)}[\xi^{(\nu)}]\eta^{(\nu+1)}}{\{\eta^{(\nu+1)\mathrm{T}}[\xi^{(\nu)}]P^{(\nu)}[\xi^{(\nu)}]\eta^{(\nu+1)}\}^{1/2}} = -[\xi^{(\nu)}]\frac{P^{(\nu)}\alpha^{(\nu)}}{\{\alpha^{(\nu)\mathrm{T}}P^{(\nu)}\alpha^{(\nu)}\}^{1/2}}$$

$$= -[\xi^{(\nu)}]\frac{P^{(\nu)}\beta^{(\nu)}}{\{\beta^{(\nu)\mathrm{T}}P^{(\nu)}\beta^{(\nu)}\}^{1/2}}. \tag{4.19}$$

Applying lemma 2.5 and taking note of the fact that $\beta_N^{(\nu)} = 0$, we see that $P^{(\nu)}\beta^{(\nu)}$ and $P^{(\nu)}\beta^{*(\nu)}$ are written as

$$P\beta = \begin{pmatrix} \hat{P}_{EE} + \Delta P_{EE} \\ \Delta P_{E\bar{B}}^{\mathrm{T}} \\ \Delta P_{EN}^{\mathrm{T}} \end{pmatrix}\beta_E + \begin{pmatrix} \Delta P_{E\bar{B}} \\ (I + S_{\bar{B}N}S_{\bar{B}N}^{\mathrm{T}})^{-1} + \Delta P_{\overline{BB}} \\ S_{\bar{B}N}^{\mathrm{T}}(I + S_{\bar{B}N}S_{\bar{B}N}^{\mathrm{T}})^{-1} + \Delta P_{\bar{B}N}^{\mathrm{T}} \end{pmatrix}\beta_{\bar{B}} \tag{4.20}$$

and

$$P\beta^* = \begin{pmatrix} \hat{P}_{EE} + \Delta P_{EE} \\ \Delta P_{E\bar{B}}^{\mathrm{T}} \\ \Delta P_{EN}^{\mathrm{T}} \end{pmatrix}\beta_E^*, \tag{4.21}$$

where we omit the upper index ν indicating the number of iterative steps. Putting

$$\tilde{P}_{EE} = \hat{P}_{EE} + \Delta P_{EE}, \tag{4.22}$$

we have

$$\beta^{\mathrm{T}} P\beta = \beta_E^{\mathrm{T}} \tilde{P}_{EE}\beta_E + 2\beta_E^{\mathrm{T}} \Delta P_{E\bar{B}}\beta_{\bar{B}} + \beta_{\bar{B}}^{\mathrm{T}}\{(I + S_{\bar{B}N}S_{\bar{B}N}^{\mathrm{T}})^{-1} + \Delta P_{\overline{BB}}\}\beta_{\bar{\beta}}, \tag{4.23}$$

$$\beta^{*\mathrm{T}} P\beta^* = \beta_E^{*\mathrm{T}} \tilde{P}_{EE}\beta_E^*, \tag{4.24}$$

and

$$\beta^{*\mathrm{T}} P\beta = \beta_E^{*\mathrm{T}} \tilde{P}_{EE}\beta_E + \beta_E^{*\mathrm{T}} \Delta P_{E\bar{B}}\beta_{\bar{B}}. \tag{4.25}$$

Now, let us consider the local Karmarkar potential function $f_{\mathscr{X}}$ defined as in (4.3) associated with $\mathscr{X}$ and the LP problem (4.17), and observe the asymptotic reduction of the local Karmarkar potential function by one iteration of (1.5), which is written as

$$\begin{aligned}
&f_{\mathscr{X}}(\xi^{(\nu)} + \mu^{(\nu)} d\xi^{\mathrm{QP}(\nu)}) - f_{\mathscr{X}}(\xi^{(\nu)}) \\
&= \left(|E(\mathscr{X})| \log \eta_{E(\mathscr{X})}^{*\mathrm{T}} (\xi_{E(\mathscr{X})}^{(\nu)} + \mu^{(\nu)} d\xi^{\mathrm{QP}(\nu)}) - \sum_{i\in E(\mathscr{X})} \log (\xi_i^{(\nu)} + \mu^{(\nu)} d\xi_i^{\mathrm{QP}(\nu)}) \right) \\
&\quad - \left(|E(\mathscr{X})| \log \eta_{E(\mathscr{X})}^{*\mathrm{T}} \xi_{E(\mathscr{X})}^{(\nu)} - \sum_{i\in E(\mathscr{X})} \log \xi_i^{(\nu)} \right) \\
&= |E(\mathscr{X})| \log \left(1 - \mu^{(\nu)} \beta^{*(\nu)\mathrm{T}} \frac{P^{(\nu)}\beta^{(\nu)}}{\{\beta^{(\nu)\mathrm{T}} P^{(\nu)}\beta^{(\nu)}\}^{1/2}} \right) \\
&\quad - \sum_{i\in(\mathscr{X})} \log \left\{ 1 - \mu^{(\nu)} \frac{P^{(\nu)}\beta^{(\nu)}}{\{\beta^{(\nu)\mathrm{T}} P^{(\nu)}\beta^{(\nu)}\}^{1/2}} \right\}_i .
\end{aligned} \tag{4.26}$$

Note that $f_{\mathscr{X}}(\xi^{(\nu)}), f_{\mathscr{X}}(\xi^{(\nu+1)})$ are well-defined because of lemma 3.7 for sufficiently large ν. We will show that (4.26) is bounded from above asymptotically by a negative constant under the assumptions of theorem 1.1 and the condition that x^* is not the optimal solution for $\langle\mathbf{D}\rangle$. For this purpose, we rewrite (4.26) as

$$\begin{aligned}
&f_{\mathscr{X}}(\xi^{(\nu)} + \mu^{(\nu)} d\xi^{\mathrm{QP}(\nu)}) - f_{\mathscr{X}}(\xi^{(\nu)}) \\
&= |E(\mathscr{X})| \log \left(1 - \mu^{*(\nu)} \beta^{*(\nu)\mathrm{T}} \frac{P^{(\nu)}\beta^{(\nu)}}{\{\beta^{*(\nu)\mathrm{T}} P^{(\nu)}\beta^{*(\nu)}\}^{1/2}} \right) \\
&\quad - \sum_{i\in E(\mathscr{X})} \log \left\{ 1 - \mu^{*(\nu)} \frac{P^{(\nu)}\beta^{(\nu)}}{\{\beta^{*(\nu)\mathrm{T}} P^{(\nu)}\beta^{*(\nu)}\}^{1/2}} \right\}_i ,
\end{aligned} \tag{4.27}$$

by introducing the new step-size

$$\mu^{*(\nu)} = \mu^{(\nu)} \frac{\{\beta^{*(\nu)\mathrm{T}} P^{(\nu)} \beta^{*(\nu)}\}^{1/2}}{\{\beta^{(\nu)\mathrm{T}} P^{(\nu)} \beta^{(\nu)}\}^{1/2}}. \tag{4.28}$$

In the following we give a bound for the reduction of $f_{\mathscr{X}}$ per iteration. We show this in several steps of observations.

(1) We have

$$\Phi^{(\nu)}_{E(\mathscr{X})} \to 0 \tag{4.29}$$

as $\nu \to \infty$, where $\Phi^{(\nu)}_{E(\mathscr{X})} = O(\|\xi^{(\nu)}_{E(\mathscr{X})}\|)$.

Proof

This immediately follows from the fact that $\{x^{(\nu)}\}$ converges to an interior point of $\mathscr{X}$.

(2) For sufficiently large ν,

$$\frac{\eta^{*\mathrm{T}}\xi^{(\nu+1)}}{\eta^{*\mathrm{T}}\xi^{(\nu)}} = 1 - \mu^{*(\nu)} \frac{\beta^{*(\nu)\mathrm{T}} P^{(\nu)} \beta^{(\nu)}}{\{\beta^{*(\nu)\mathrm{T}} P^{(\nu)} \beta^{*(\nu)}\}^{1/2}} \tag{4.30}$$

is bounded below by a positive constant, say, ζ.

Proof

If such ζ does not exist, we can take a subsequence $\{x^{(\nu_\tau)}\}$ of $\{x^{(\nu)}\}$ such that

$$\frac{\eta^{*\mathrm{T}}\xi^{(\nu_\tau+1)}}{\eta^{*\mathrm{T}}\xi^{(\nu_\tau)}} \to 0. \tag{4.31}$$

On the other hand, since

$$\|\xi^{(\nu+1)}_{E(\mathscr{X})}\| \geq (1-\mu^{(\nu)})\|\xi^{(\nu)}_{E(\mathscr{X})}\| \geq (1-\mu_{\max})\|\xi^{(\nu)}_{E(\mathscr{X})}\|, \tag{4.32}$$

we have

$$\frac{\eta^{*\mathrm{T}}\xi^{(\nu_\tau+1)}}{\|\xi_E^{(\nu_\tau+1)}\|} = \frac{\eta^{*\mathrm{T}}\xi^{(\nu_\tau+1)}}{\eta^{*\mathrm{T}}\xi^{(\nu_\tau)}} \frac{\|\xi_E^{(\nu_\tau)}\|}{\|\xi_E^{(\nu_\tau+1)}\|} \frac{\eta^{*\mathrm{T}}\xi^{(\nu_\tau)}}{\|\xi_E^{(\nu_\tau)}\|} \leq \frac{\eta^{*\mathrm{T}}\xi^{(\nu_\tau+1)}}{\eta^{*\mathrm{T}}\xi^{(\nu_\tau)}} \frac{\|\eta^*\|}{1-\mu^{(\nu)}} \leq \frac{\eta^{*\mathrm{T}}\xi^{(\nu_\tau+1)}}{\eta^{*\mathrm{T}}\xi^{(\nu_\tau)}} \frac{\|\eta^*\|}{1-\mu_{\max}}. \tag{4.33}$$

Then we obtain

$$\frac{\eta^{*\mathrm{T}}\xi^{(\nu_\tau+1)}}{\|\xi_E^{(\nu_\tau+1)}\|} \to 0 \tag{4.34}$$

as $\tau \to \infty$, which, however, is a contradiction to lemma 4.7.

(3) For sufficiently large ν, we have

$$\beta^{*(\nu)\mathrm{T}} P^{(\nu)} \beta^{*(\nu)} = \beta_E^{*(\nu)\mathrm{T}} \tilde{P}_{EE}^{(\nu)} \beta_E^{*(\nu)} \geq \frac{1}{2|E(\mathscr{X})|}. \tag{4.35}$$

Since $\beta_E^{*(\nu)\mathrm{T}} \tilde{P}_{EE}^{(\nu)} \beta_E^{*(\nu)} = \beta_E^{*(\nu)\mathrm{T}} (\hat{P}_{EE}^{(\nu)} + \Delta P_{EE}^{(\nu)}) \beta_E^{*(\nu)}$, $\beta^{*(\nu)}$ is bounded, and $\Delta P_{EE}^{(\nu)} \to 0$ as $\Phi_E^{(\nu)} \to 0$, it is enough to show $\beta_E^{*(\nu)\mathrm{T}} \hat{P}_{EE}^{(\nu)} \beta_E^{*(\nu)} \geq 1/|E(\mathscr{X})|$, which can be seen easily by taking note of $\mathbf{1}_E^{\mathrm{T}} \hat{P}_{EE}^{(\nu)} \beta_E^{*(\nu)} = \mathbf{1}_E^{\mathrm{T}} \beta_E^{*(\nu)} = 1$ (cf. lemma 2.5).

(4)

$$\lim_{\nu\to\infty} \frac{\beta_E^{*(\nu)\mathrm{T}} (P^{(\nu)}\beta^{(\nu)})_E}{\beta_E^{*(\nu)\mathrm{T}} \tilde{P}_{EE}^{(\nu)} \beta_E^{*(\nu)}} = \lim_{\nu\to\infty} \frac{\beta^{*(\nu)\mathrm{T}} (P^{(\nu)}\beta^{(\nu)})}{\beta^{*(\nu)\mathrm{T}} P^{(\nu)} \beta^{*(\nu)}} = 1. \tag{4.36}$$

Proof

This easily follows from (4.35), (4.25), $d\beta_{E(\mathscr{X})}^{(\nu)} \to 0$, $\Phi_{E(\mathscr{X})}^{(\nu)} \to 0$, and the boundedness of $\|\beta^{(\nu)}\|$.

(5)

$$0 < \eta' < \frac{\mu^{*(\nu)}}{\mu^{(\nu)}} = \frac{\{\beta^{*(\nu)\mathrm{T}} P^{(\nu)} \beta^{*(\nu)}\}^{1/2}}{\{\beta^{(\nu)\mathrm{T}} P^{(\nu)} \beta^{(\nu)}\}^{1/2}} \leq 1 + \delta^{(\nu)} \tag{4.37}$$

asymptotically as $\nu \to \infty$, where η' is a constant and $\delta^{(\nu)}$ is a sequence that tends to zero as $\nu \to \infty$.

Proof

As shown in (4.23), we have

$$\begin{aligned} \beta^{(\nu)\mathrm{T}} P^{(\nu)} \beta^{(\nu)} = {} & \beta_E^{(\nu)\mathrm{T}} \tilde{P}_{EE}^{(\nu)} \beta_E^{(\nu)} + 2\beta_E^{(\nu)\mathrm{T}} \Delta P_{E\bar{B}}^{(\nu)} \beta_{\bar{B}}^{(\nu)} \\ & + \beta_{\bar{B}}^{(\nu)\mathrm{T}} \{(I + S_{\bar{B}N}^{(\nu)} S_{\bar{B}N}^{(\nu)\mathrm{T}})^{-1} + \Delta P_{\bar{B}\bar{B}}^{(\nu)}\} \beta_{\bar{B}}^{(\nu)}. \end{aligned} \tag{4.38}$$

We observed in (4.16) that $\|\beta^{(\nu)}\|$ is bounded above by a constant. Since $\Phi_{E(\mathscr{X})}^{(\nu)} \to 0$, we see $\|\Delta P_{E\bar{B}}^{(\nu)}\| \to 0$ and $\|\Delta P_{\bar{B}\bar{B}}^{(\nu)}\| \to 0$ as $\nu \to \infty$. Taking account of the fact that

$d\beta_E^{(\nu)} \to 0$ and $\|(I+S_{\bar{B}N}^{(\nu)}S_{\bar{B}N}^{(\nu)\mathrm{T}})^{-1}\| \le 1$, we have

$$\begin{aligned}
\beta^{*(\nu)\mathrm{T}}P^{(\nu)}\beta^{*(\nu)} - \epsilon^{(\nu)} &= \beta_E^{*(\nu)\mathrm{T}}\tilde{P}_{EE}^{(\nu)}\beta_E^{*(\nu)} - \epsilon^{(\nu)} \\
&\le \beta^{(\nu)\mathrm{T}}P^{(\nu)}\beta^{(\nu)} \\
&\le \beta_E^{*(\nu)\mathrm{T}}\tilde{P}_{EE}^{(\nu)}\beta_E^{*(\nu)} + \|\beta_{\bar{B}}^{(\nu)}\|^2\,\|(I+S_{\bar{B}N}^{(\nu)}S_{\bar{B}N}^{(\nu)\mathrm{T}})^{-1}\| + \epsilon^{(\nu)} \\
&= \beta^{*(\nu)\mathrm{T}}P^{(\nu)}\beta^{*(\nu)} + \|\beta_{\bar{B}}^{(\nu)}\|^2\,\|(I+S_{\bar{B}N}^{(\nu)}S_{\bar{B}N}^{(\nu)\mathrm{T}})^{-1}\| + \epsilon^{(\nu)}, \qquad (4.39)
\end{aligned}$$

where $\epsilon^{(\nu)}$ tends to zero as $\nu \to \infty$, which, together with $\beta_E^{*(\nu)\mathrm{T}}\tilde{P}_{EE}^{(\nu)}\beta_E^{*(\nu)} \ge 1/(2|E(\mathscr{X})|)$, immediately implies (4.37).

(6)

$$\lim_{\nu\to\infty} \frac{1-\mu^{*(\nu)}\dfrac{\beta^{*(\nu)\mathrm{T}}(P^{(\nu)}\beta^{(\nu)})}{\{\beta^{*(\nu)\mathrm{T}}P^{(\nu)}\beta^{*(\nu)}\}^{1/2}}}{1-\mu^{*(\nu)}\{\beta^{*(\nu)\mathrm{T}}P^{(\nu)}\beta^{*(\nu)}\}^{1/2}} \to 1. \qquad (4.40)$$

Proof

This follows from observations (2) and (4).

(7) For each $i \in E(\mathscr{X})$, we have

$$\lim_{\nu\to\infty} \frac{1-\mu^{*(\nu)}\dfrac{(P^{(\nu)}\beta^{(\nu)})_i}{\{\beta^{*(\nu)\mathrm{T}}P^{(\nu)}\beta^{*(\nu)}\}^{1/2}}}{1-\mu^{*(\nu)}\dfrac{(P^{(\nu)}\beta^{*(\nu)})_i}{\{\beta^{*(\nu)\mathrm{T}}P^{(\nu)}\beta^{*(\nu)}\}^{1/2}}} \to 1. \qquad (4.41)$$

Proof

It is not difficult to see that $(P^{(\nu)}\beta^{(\nu)} - P^{(\nu)}\beta^{*(\nu)})_i$ tends to zero as $\nu \to \infty$ for each $i \in E(\mathscr{X})$. Since $\mu^{(\nu)} \le \mu_{\max} < 1$ and $\|P^{(\nu)}\beta^{*(\nu)}/\{\beta^{*(\nu)}P^{(\nu)}\beta^{*(\nu)}\}^{1/2}\| = 1$, the denominator of (4.41) is bounded below by a positive constant. Taking note of these facts and using (4.35), we see (4.41).

(8) For sufficiently large ν, we have

$$
\begin{aligned}
&f_{\mathscr{X}}(\xi^{(\nu)}+\mu^{(\nu)}d\xi^{\mathrm{QP}(\nu)})-f_{\mathscr{X}}(\xi^{(\nu)})\\
&\quad=|E(\mathscr{X})|\log\left(1-\mu^{(\nu)}\beta^{*(\nu)\mathrm{T}}\frac{P^{(\nu)}\beta^{(\nu)}}{\{\beta^{(\nu)}P^{(\nu)}\beta^{(\nu)}\}^{1/2}}\right)\\
&\qquad-\sum_{i\in E(\mathscr{X})}\log\left\{1-\mu^{(\nu)}\frac{P^{(\nu)}\beta^{(\nu)}}{\{\beta^{(\nu)}P^{(\nu)}\beta^{(\nu)}\}^{1/2}}\right\}_i\le-\delta'',
\end{aligned}
\tag{4.42}
$$

where δ'' is a positive constant.

Proof

From (4.27), (4.40) and (4.41), we see

$$
\begin{aligned}
&f_{\mathscr{X}}(\xi^{(\nu)}+\mu^{(\nu)}d\xi^{\mathrm{QP}(\nu)})-f_{\mathscr{X}}(\xi^{(\nu)})\\
&\quad=|E(\mathscr{X})|\log\left(1-\mu^{(\nu)}\beta^{*(\nu)\mathrm{T}}\frac{P^{(\nu)}\beta^{(\nu)}}{\{\beta^{(\nu)}P^{(\nu)}\beta^{(\nu)}\}^{1/2}}\right)\\
&\qquad-\sum_{i\in E(\mathscr{X})}\log\left\{1-\mu^{(\nu)}\frac{P^{(\nu)}\beta^{(\nu)}}{\{\beta^{(\nu)}P^{(\nu)}\beta^{(\nu)}\}^{1/2}}\right\}_i\\
&\quad=|E(\mathscr{X})|\log\left(1-\mu^{*(\nu)}\beta^{*(\nu)\mathrm{T}}\frac{P^{(\nu)}\beta^{(\nu)}}{\{\beta^{*(\nu)}P^{(\nu)}\beta^{*(\nu)}\}^{1/2}}\right)\\
&\qquad-\sum_{i\in E(\mathscr{X})}\log\left\{1-\mu^{*(\nu)}\frac{P^{(\nu)}\beta^{(\nu)}}{\{\beta^{*(\nu)}P^{(\nu)}\beta^{*(\nu)}\}^{1/2}}\right\}_i\\
&\quad\le|E(\mathscr{X})|\log(1-\mu^{*(\nu)}\{\beta^{*(\nu)}P^{(\nu)}\beta^{*(\nu)}\}^{1/2})\\
&\qquad-\sum_{i\in E(\mathscr{X})}\log\left\{1-\mu^{*(\nu)}\frac{P^{(\nu)}\beta^{*(\nu)}}{\{\beta^{*(\nu)}P^{(\nu)}\beta^{*(\nu)}\}^{1/2}}\right\}_i+\epsilon\\
&\quad\le f_{\mathscr{X}}(\xi^{(\nu)}+\mu^{*(\nu)}d\xi^{\mathrm{LP}(\nu)})-f_{\mathscr{X}}(\xi^{(\nu)})+\epsilon
\end{aligned}
\tag{4.43}
$$

holds for sufficiently large ν, where ϵ is an (arbitrary) positive constant given in advance.

Recall that we assumed that x^* is not the optimal solution of $\langle \mathbf{D} \rangle$, and apply lemma 4.1 to obtain an upper bound for $f_{\mathscr{X}}(\xi^{(\nu)} + \mu^{*(\nu)} d\xi^{\mathrm{LP}(\nu)}) - f_{\mathscr{X}}(\xi^{(\nu)})$ in the rightmost hand side of (4.43). Taking account of the observation (5) and the assumption that $0 < \mu_{\min} \le \mu^{(\nu)} \le \mu_{\max} \le 1/8$, we see that $\mu^{*(\nu)}$ is asymptotically bounded from below by a strictly positive number and from above by 11/80. Furthermore, it is easily checked by using observations (2) and (6) that

$$\begin{aligned}\liminf_{\nu\to\infty} \frac{\eta^{*\mathrm{T}}(\xi^{(\nu)} + \mu^{*(\nu)} d\xi^{\mathrm{LP}(\nu)})}{\eta^{*\mathrm{T}}\xi^{(\nu)}} &= \liminf_{\nu\to\infty} \left(1 + \mu^{*(\nu)} \eta^{*\mathrm{T}} \frac{d\xi^{\mathrm{LP}(\nu)}}{\eta^{*\mathrm{T}}\xi^{(\nu)}} \right) \\ &= \liminf_{\nu\to\infty} (1 - \mu^{*(\nu)} \{\beta^{*(\nu)\mathrm{T}} P^{(\nu)} \beta^{*(\nu)}\}^{1/2}) \\ &= \liminf_{\nu\to\infty} \left(1 - \mu^{*(\nu)} \frac{\beta^{*(\nu)\mathrm{T}} P^{(\nu)} \beta^{(\nu)}}{\{\beta^{*(\nu)\mathrm{T}} P^{(\nu)} \beta^{*(\nu)}\}^{1/2}} \right) \ge \zeta, \quad (4.44)\end{aligned}$$

where $\zeta > 0$ is a constant, then the condition (iii) of lemma 4.1 is satisfied if ν is sufficiently large. Now we may apply lemma 4.1 to obtain

$$f_{\mathscr{X}}(\xi^{(\nu)} + \mu^{*(\nu)} d\xi^{\mathrm{LP}(\nu)}) - f_{\mathscr{X}}(\xi^{(\nu)}) < -\delta''' \tag{4.45}$$

for sufficiently large ν, where δ''' is a positive constant. Since we can choose ϵ in (4.43) sufficiently small in advance, the inequality (4.42) holds for sufficiently large ν, by choosing $\delta'' > 0$ appropriately. This completes the proof of the reduction of $f_{\mathscr{X}}$ when ν is sufficiently large.

Now, we are ready to complete the proof of theorem 1.1. Recall that we assumed that the limiting point x^* is not the optimal solution of $\langle \mathbf{D} \rangle$. Since x^* is not the optimal solution, from the observation (8), we see

$$f_{\mathscr{X}}(\xi^{(\nu)}) \to -\infty \tag{4.46}$$

as $\nu \to \infty$. By using the well-known inequality

$$\exp(f_{\mathscr{X}}(\xi^{(\nu)})) \ge \left(|E(\mathscr{X})| \frac{\eta^{*\mathrm{T}}\xi^{(\nu)}}{\|\xi^{(\nu)}_{E(\mathscr{X})}\|_1} \right)^{|E(\mathscr{X})|}, \tag{4.47}$$

we obtain

$$\lim_{\nu\to\infty} \frac{\eta^{*\mathrm{T}}\xi^{(\nu)}}{\|\xi^{(\nu)}_{E(\mathscr{X})}\|_1} = 0. \tag{4.48}$$

However, this is a contradiction to lemma 3.7 (cf. (3.42)), coming from assuming that x^* is not the optimal solution. Thus, x^* has to be the optimal solution of $\langle \mathbf{D} \rangle$. This completes the proof of the theorem. □

5. Concluding remark

So far we demonstrated the global convergence of the affine scaling algorithm for strictly convex quadratic programming problems satisfying the dual nondegeneracy condition. We hope that the dual nondegeneracy condition is removed by developing this approach.

Acknowledgements

The author wishes to thank Prof. Kunio Tanabe and Prof. Shinji Mizuno of the Institute of Statistical Mathematics and Mr. Masakazu Muramatsu of the Graduate University of Advanced Studies for their stimulating discussions. He would like to thank the anonymous referees for their valuable comments and suggestions.

The revision was carried out during the author's stay at the Department of Computational and Applied Mathematics of Rice University, Houston, USA. He would like to thank Prof. J. Dennis and the colleagues there for providing the excellent research environment he is enjoying.

This research was supported in part by a Grant-in-Aid for Encouragement of Young Scientists 03740126 (1991) and Overseas Research Scholars (1992) of the Ministry of Education, Science and Culture of Japan.

References

[1] I. Adler et al., An implementation of Karmarkar's algorithm for linear programming, Math. Progr. 44 (1989) 297–335.

[2] E.R. Barnes, A variation on Karmarkar's algorithm for solving linear programming problems, Math. Progr. 36 (1986) 174–182.

[3] I.I. Dikin, Iterative solution of problems of linear and quadratic programming, Sov. Math. Doklady 8 (1967) 674–675.

[4] I.I. Dikin and V.I. Zorkaltsev, *Iterative Solutions of Mathematical Programming Problems* (Nauka, Novosibirsk, 1980).

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

[6] J. Moré, The Levenberg–Marquardt algorithm: Implementation and theory, in: *Numerical Analysis*, ed. G.A. Watson (Springer, Berlin, 1978) pp. 105–116.

[7] A. Schrijver, *Theory of Linear and Integer Programming* (Wiley, Chichester, England, 1986).

[8] J. Sun, A convergence proof of an affine-scaling algorithm for convex quadratic programming without nondegeneracy assumptions, Technical Report, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208, USA (1990).

[9] P. Tseng and Z.-Q. Luo, On the convergence of the affine-scaling algorithm, Math. Progr. Series A 56 (1992) 301–319.
[10] T. Tsuchiya, Global convergence of the affine scaling methods for degenerate linear programming problems, Math. Progr. Series B 52 (1991) 377–404.
[11] T. Tsuchiya, Global convergence property of the affine scaling methods for primal degenerate linear programming problems, Math. Oper. Res. 17 (1992) 527–557.
[12] R.J. Vanderbei et al., A modification of Karmarkar's linear programming algorithm, Algorithmica 1 (1986) 395–407.
[13] Y. Ye, An extension of Karmarkar's algorithm and the trust region method for quadratic programming, in: *Progress in Mathematical Programming*, ed. N. Megiddo (Springer, 1989) pp. 49–63.
[14] Y. Ye, A new complexity result on minimization of a quadratic function over a sphere constraint, Technical Report, Department of Management Sciences, The University of Iowa, Iowa City, IA 52242, USA (November, 1990).
[15] Y. Ye and E. Tse, An extension of Karmarkar's projective algorithm for convex quadratic programming, Math. Progr. 44 (1989) 157–179.