PRIMAL-DUAL AFFINE-SCALING ALGORITHMS FAIL FOR SEMIDEFINITE PROGRAMMING

MASAKAZU MURAMATSU AND ROBERT J. VANDERBEI

ABSTRACT. In this paper, we give an example of a semidefinite programming problem in which primal-dual affine-scaling algorithms using the HRVW/KSH/M, MT, and AHO directions fail. We prove that each of these algorithm can generate a sequence converging to a non-optimal solution, and that, for the AHO direction, even its associated continuous trajectory can converge to a non-optimal point. In contrast with these directions, we show that the primal-dual affine-scaling algorithm using the NT direction for the same semidefinite programming problem always generates a sequence converging to the optimal solution. Both primal and dual problems have interior feasible solutions, unique optimal solutions which satisfy strict complementarity, and are nondegenerate everywhere.

1. INTRODUCTION

We consider the standard form semidefinite programming (SDP) problem:

$$\begin{array}{ll} \text{minimize} & C \bullet X \\ \text{subject to} & A_i \bullet X = b_i, \quad i = 1, \dots, m, \quad X \succeq 0, \end{array} \tag{1}$$

and its dual:

$$\begin{array}{ll} \text{maximize} & b^t u \\ \text{subject to} & Z + \sum_{i=1}^m u_i A_i = C, \quad Z \succeq 0, \end{array} \tag{2}$$

where C, X, A_i belong to the space $\mathcal{S}(n)$ of $n \times n$ real symmetric matrices, the operator $\bullet$ denotes the standard inner product in $\mathcal{S}(n)$, i.e., $C \bullet X := \text{tr}(CX) = \sum_{i,j} C_{ij} X_{ij}$, and $X \succeq 0$ means that X is positive semidefinite.

SDP bears a remarkable resemblance to LP. In fact, it is known that several interior-point methods for LP and their polynomial convergence analysis can be naturally extended to SDP (see Alizadeh [1], Jarre [15], Nesterov and Nemirovskii [28, 29], Vandenberghe and Boyd [38]). However, in extending primal-dual interior-point methods from LP to SDP, certain choices have to be made and the resulting search direction depends on these

Date. March 4, 1998.

1991 *Mathematics Subject Classification.* Primary 60E05, Secondary 60C05.

Key words and phrases. Semidefinite Programming, Primal-dual Interior-Point Method, Affine-Scaling Algorithm, Global Convergence Analysis.

research partially supported by the NSF through grant CCR94-03789.

choices. As a result, there can be several search directions for SDP corresponding to a single search direction for LP.

This paper deals with primal-dual interior-point algorithms for SDP based on the following four search directions:

(i) HRVW/KSH/M direction,
(ii) MT direction,
(iii) AHO direction,
(iv) NT direction.

We study a specific simple SDP problem, and for this problem carefully investigate the behavior of the sequence generated by the interior-point methods using these four directions to show how the convergence property of the algorithm varies depending on the choice of direction.

There are two popular classes of interior-point methods: *affine-scaling* algorithm and *path-following* algorithm. Path-following algorithm is characterized by a parametric relaxation of the following optimality conditions for SDP:

$$A_i \bullet X = b_i, \quad i = 1, \ldots, m, \tag{3}$$

$$Z + \sum_{i=1}^{m} u_i A_i = C, \tag{4}$$

$$XZ = \mu I, \tag{5}$$

$$X \succeq 0, \; Z \succeq 0,$$

where $\mu > 0$ is a barrier parameter. In such algorithm, it is necessary to specify a specific choice of μ at any iteration. The particulars vary from paper to paper, and we therefore omit them here. When $\mu \equiv 0$ the corresponding method is called affine-scaling algorithm. Most of the existing SDP literature considers path-following algorithm. In this paper, we restrict our attention to affine-scaling algorithm.

The *affine-scaling* algorithm was originally proposed for LP by Dikin [8], and independently rediscovered by Barnes [5], Vanderbei, Meketon and Freedman [39] and others, after Karmarkar [16] proposed the first polynomial-time interior-point method. Though polynomial-time complexity has not been proved yet for this algorithm, global convergence using so-called long steps was proved by Tsuchiya and Muramatsu [37]. This algorithm is often called the *primal (or dual)* affine-scaling algorithm because the algorithm is based on the primal (or dual) problem only. There is also a notion of *primal-dual* affine-scaling algorithm. In fact, for LP, there are two different types of primal-dual affine-scaling algorithm proposed to date; one by Monteiro, Adler and Resende [23], and the other by Jansen, Roos, and Terlaky [14]. The latter is sometimes called the *Dikin-type* primal-dual affine-scaling algorithm. Both of these papers provide a proof of polynomial-time convergence for the respective algorithm, though the complexity of the former algorithm is much worse than the latter.

All of the affine-scaling algorithms just described can be naturally extended to SDP. Faybusovich [9, 10] dealt with the SDP extension of the primal affine-scaling algorithm. Global convergence of the associated continuous trajectory was proved by Goldfarb and Scheinberg [12]. However, Muramatsu [27] gave an example for which the algorithm fails to converge to an optimal solution for any step size, showing that the primal affine-scaling algorithm for SDP does not have the same global convergence property that one has for LP. For both primal-dual affine-scaling algorithms, de Klerk, Roos and Terlaky [7] proved polynomial-time convergence. However, as was mentioned before, there exist several different search directions in primal-dual interior-point methods for SDP, and each of the primal-dual affine-scaling algorithms studied by de Klerk, Roos and Terlaky was based on a certain specific choice of search direction. Below we discuss in detail how the various search directions arise.

The primal-dual affine-scaling direction proposed by Monteiro, Adler and Resende [23] is the Newton direction for the set of optimality conditions, i.e., primal feasibility, dual feasibility and complementarity. For SDP, the optimality conditions are (3), (4) and

$$XZ = 0. \tag{6}$$

A direct application of Newton's method produces the following equations for ΔX, Δu and ΔZ (throughout this paper, we assume that the current point is primal and dual feasible):

$$\begin{aligned} A_i \bullet \Delta X &= 0, \quad i = 1, \ldots, m, & (7)\\ \Delta Z + \sum_{i=1}^{m} \Delta u_i A_i &= 0, & (8)\\ X \Delta Z + \Delta X Z &= -XZ. & (9) \end{aligned}$$

However, due to (9), the solution of this system does not give a symmetric solution in general (actually ΔZ must be symmetric by (8) but ΔX is generally not symmetric). To date, several ways have been proposed to overcome this difficulty, each producing different directions in general.

In this paper, we study a specific simple SDP problem, and for this problem carefully investigate the behavior of the sequence generated by the primal-dual affine-scaling algorithms using these four directions to show how the convergence property of the algorithm varies depending on the choice of direction.

Now we describe the four directions we deal with in this paper. Note that the papers mentioned below deal exclusively with path-following algorithms, for which the corresponding affine-scaling algorithms can be derived by setting $\mu = 0$.

1.1. The HRVW/KSH/M Direction. This direction is derived by using (7)–(9) as is, and then taking the symmetric part of the resulting ΔX. This method to make a symmetric direction was independently proposed by Helmberg, Rendl, Vanderbei and Wolkowicz [13], Kojima, Shindoh and Hara [18], and Monteiro [21]. Polynomial-time convergence was proved for the path-following algorithms using this direction. For related work, see also the papers of Lin and Saigal [19], Potra and Sheng [32], and Zhang [40]. The HRVW/KSH/M direction is currently very popular for practical implementation because of its computational simplicity. Almost all SDP solvers have an option to use this direction, and some serious solvers (for example, Borchers [6] and Fujisawa and Kojima [11]) use this direction only.

1.2. The MT Direction. Monteiro and Tsuchiya [24] apply Newton's method to the system obtained from (3)–(6) by replacing (6) with

$$X^{1/2} Z X^{1/2} = 0.$$

The resulting direction is guaranteed to be symmetric. It is the solution of (7), (8) and

$$\begin{aligned} VZX^{1/2} + X^{1/2}ZV + X^{1/2}\Delta Z X^{1/2} &= -X^{1/2} Z X^{1/2}, & (10)\\ VX^{1/2} + X^{1/2}V &= \Delta X & (11) \end{aligned}$$

where $V \in \mathcal{S}(n)$ is an auxiliary variable. They proved polynomial-time convergence of the path-following algorithm using this direction. Recently, Monteiro and Zanjacomo [25] discussed a computational aspects of this direction, and gave some numerical experiments.

1.3. The AHO Direction. Alizadeh, Haeberly, and Overton [2] proposed symmetrizing equation (6) by rewriting it as

$$XZ + ZX = 0, \tag{12}$$

and then applying Newton's method to (3), (4) and (12). The resulting direction is a solution of (7), (8) and

$$X\Delta Z + \Delta X Z + Z\Delta X + \Delta Z X = -(XZ + ZX). \tag{13}$$

Several convergence properties including polynomial-time convergence are known for the path-following algorithm using the AHO direction. See for example the work of Kojima, Shida and Shindoh [17], Monteiro [22], and Tseng [36]. The AHO direction however, is not necessarily well-defined on the feasible region as observed by Shida, Shindoh and Kojima [33]; the linear system (7), (8), and (13) can be inconsistent for some problems. In fact, a specific example was given by Todd, Toh and Tütüncü [35]. On the other hand, Alizadeh, Haeberly, and Overton [4] report that the path-following algorithm using the AHO direction has empirically better convergence properties than the one using the HRVW/KSH/M direction.

1.4. The NT Direction. Nesterov and Todd [30, 31] proposed primal-dual algorithms for more general convex programming than SDP, which includes SDP as a special case. Their search direction naturally produces a symmetric direction. The direction is the solution of (7), (8) and

$$\Delta X + D\Delta Z D = -X, \tag{14}$$

where $D \in \mathcal{S}(n)$ is a unique solution of

$$DZD = X. \tag{15}$$

Polynomial-time convergence of the corresponding path-following algorithm was proved in their original paper [30]. Also, see the works of Monteiro and Zhang [26], Luo, Sturm and Zhang [20], and Sturm and Zhang [34] for some convergence properties of the algorithms using the NT direction. The primal-dual affine-scaling algorithm studied by de Klerk, Roos and Terlaky [7] was based on this direction. As for numerical computation, Todd, Toh and Tütüncü [35] reported that the path-following algorithm using the NT direction is more robust than algorithms based on the HRVW/KSH/M and AHO directions.

1.5. Notation and Organization. The rest of this paper is organized as follows. In Section 2, we introduce the specific SDP problem we wish to study.

Section 3, deals with the HRVW/KSH/M direction. We consider the long-step primal-dual affine-scaling algorithm. One iteration of the long-step algorithm using search direction $(\Delta X, \Delta u, \Delta Z)$ is as follows:

$$\begin{aligned} X^+ &= X + \lambda\hat{\alpha}\Delta X, \\ u^+ &= u + \lambda\hat{\alpha}\Delta u, \\ Z^+ &= Z + \lambda\hat{\alpha}\Delta Z, \end{aligned}$$

where $\hat{\alpha}$ is defined by

$$\hat{\alpha} := \min(\alpha_P, \alpha_D), \tag{16}$$

where

$$\begin{aligned} \alpha_P &:= \sup\{\, \alpha \mid X + \alpha \Delta X \succeq 0 \,\}, \\ \alpha_D &:= \sup\{\, \alpha \mid Z + \alpha \Delta Z \succeq 0 \,\}, \end{aligned}$$

and λ is a fixed constant less than 1. We prove that, for any fixed λ, there exists a region of initial points such that the long-step primal-dual affine-scaling algorithm using the HRVW/KSH/M direction converges to a non-optimal point.

In Section 4, we prove the same statement as above for the MT direction by showing that the MT direction is identical to the HRVW/KSH/M direction for our example.

In Section 5, we deal with the AHO direction. We consider the continuous trajectory which is a solution of the following autonomous differential equation:

$$\dot{X}_t = \Delta X(X_t, u_t, Z_t), \tag{17}$$

$$\dot{u}_t = \Delta u(X_t, u_t, Z_t), \tag{18}$$

$$\dot{Z}_t = \Delta Z(X_t, u_t, Z_t). \tag{19}$$

We prove that the continuous trajectory of the AHO direction can converge to a non-optimal point.

In Section 6, we show that the long-step primal-dual affine-scaling algorithm using the NT direction generates a sequence converging to the optimal solution for any choice of λ. Note that this result does not mean the global convergence property of the algorithm, but a robust convergence property for the specific problem, for which the other three algorithms can fail to get its optimal solution.

Section 7 provides some concluding remarks.

Note that each section is fairly independent of the others and we use the same symbol $(\Delta X, \Delta u, \Delta Z)$ for different directions; e.g., ΔX in Section 3 refers to the HRVW/KSH/M direction, while in Section 5, it's the AHO direction.

2. The SDP example

The primal-dual pair of SDP problem we deal with in this paper is as follows:

$$\langle P \rangle \left\{ \begin{array}{ll} \text{minimize} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bullet X \\ \text{subject to} & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \bullet X = 2, \quad X \succeq 0, \end{array} \right. \tag{20}$$

$$\langle D \rangle \left\{ \begin{array}{ll} \text{maximize} & 2u \\ \text{subject to} & Z + u \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad Z \succeq 0, \end{array} \right. \tag{21}$$

where $X, Z \in \mathcal{S}(2)$ and $u \in R$. The equality condition of the primal (20) says that the off-diagonal elements of X must be 1 for X to be feasible. Thus, putting

$$X = \begin{pmatrix} x & 1 \\ 1 & y \end{pmatrix}, \tag{22}$$

and noting that $X \succeq 0 \Leftrightarrow x \geq 0,\ y \geq 0,\ xy \geq 1$, we see that problem (20) is equivalent to

$$\begin{array}{ll} \text{minimize} & x + y \\ \text{subject to} & x \geq 0, \quad y \geq 0, \quad xy \geq 1, \end{array} \tag{23}$$

whose optimal solution is $(x, y) = (1, 1)$.

Similarly, from the equality condition of the dual (21), we see that Z can be written as follows:

$$Z = \begin{pmatrix} 1 & -u \\ -u & 1 \end{pmatrix}, \tag{24}$$

and that the dual is equivalent to the following linear program:

$$\begin{array}{ll} \text{maximize} & 2u \\ \text{subject to} & -1 \le u \le 1, \end{array} \tag{25}$$

whose optimal solution is obviously $u = 1$.

Since we assume that the current point is primal and dual feasible in this paper, we see from (22) and (24) that each of the search directions has the following form:

$$\Delta X = \begin{pmatrix} \Delta x & 0 \\ 0 & \Delta y \end{pmatrix}, \quad \Delta Z = \begin{pmatrix} 0 & -\Delta u \\ -\Delta u & 0 \end{pmatrix}. \tag{26}$$

In the following, we put

$$\mathcal{F} := \{ (x, y, u) \mid xy \ge 1, \quad x > 0, \quad y > 0, \quad -1 \le u \le 1 \}.$$

We see that X and Z with (22) and (24) are feasible if and only if $(x, y, u) \in \mathcal{F}$, thus $\mathcal{F}$ is called primal-dual feasible region. We also define the interior of the feasible region:

$$\mathcal{F}^o := \{ (x, y, u) \mid xy > 1, \quad x > 0, \quad y > 0, \quad -1 < u < 1 \}.$$

Obviously, if $(x, y, u) \in \mathcal{F}^o$, then the corresponding X and Z are feasible and positive definite.

It is easy to see that

$$(x^*, y^*, u^*) = (1, 1, 1)$$

is the unique optimal solutions of (23) and (25), hence

$$(X^*, u^*, Z^*) = \left(\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, 1, \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \right)$$

is the unique optimal solutions of (20) and (21). It can also be easily seen that the optimal values of (20) and (21) coincide, that the optimal solutions satisfy strict complementarity, and that the problems are nondegenerate (see Muramatsu [27]; for degeneracy in SDP, see Alizadeh, Haeberly, and Overton [3]). In fact, this example problem was first proposed in Muramatsu [27] to prove that the primal affine-scaling algorithm fails.

3. THE HRVW/KSH/M DIRECTION

In this section, we consider the long-step primal-dual affine-scaling algorithm using the HRVW/KSH/M direction. To calculate the HRVW/KSH/M direction $(\Delta X, \Delta u, \Delta Z)$ at a feasible point (X, u, Z), we first solve the following system:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \bullet \widehat{\Delta X} = 0, \tag{27}$$

$$\Delta Z + \Delta u \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = 0, \tag{28}$$

$$X \Delta Z + \widehat{\Delta X} Z = -XZ. \tag{29}$$

From (27) and (28), we see that $\widehat{\Delta X}$ and ΔZ have the following form:

$$\widehat{\Delta X} = \begin{pmatrix} \Delta x & \Delta w \\ -\Delta w & \Delta y \end{pmatrix}, \quad \Delta Z = \begin{pmatrix} 0 & -\Delta u \\ -\Delta u & 0 \end{pmatrix}.$$

Note that since we apply the HRVW/KSH/M-type method, we do not assume that $\widehat{\Delta X}$ is symmetric here. Then we symmetrize $\widehat{\Delta X}$:

$$\begin{aligned} \Delta X &= (\widehat{\Delta X} + \widehat{\Delta X}^t)/2 \\ &= \frac{1}{2}\left\{\begin{pmatrix} \Delta x & \Delta w \\ -\Delta w & \Delta y \end{pmatrix} + \begin{pmatrix} \Delta x & -\Delta w \\ \Delta w & \Delta y \end{pmatrix}\right\} \\ &= \begin{pmatrix} \Delta x & 0 \\ 0 & \Delta y \end{pmatrix}. \end{aligned} \tag{30}$$

Therefore, ΔX is independent of Δw.

The third equation, (29), can be written componentwise as:

$$\begin{array}{rrrrcl} -\Delta u & -u\Delta w & +\Delta x & & = & u - x, \\ -y\Delta u & -\Delta w & & -u\Delta y & = & uy - 1, \\ -x\Delta u & +\Delta w & -u\Delta x & & = & ux - 1, \\ -\Delta u & +u\Delta w & & +\Delta y & = & u - y. \end{array}$$

Solving these linear equalities, we have

$$\Delta u = \frac{2(1-u^2)}{x+y+2u}, \tag{31}$$

$$\Delta x = \frac{2-xy-x^2}{x+y+2u}, \tag{32}$$

$$\Delta y = \frac{2-xy-y^2}{x+y+2u}. \tag{33}$$

There is also an equation for Δw but we don't write it since it disappears after symmetrization.

Figure 1 shows the vector field $(\Delta x, \Delta y)$ on the primal feasible region with $u = -0.5$. In fact, since $(\Delta x, \Delta y)$ is independent of u after normalization, u can be arbitrary. From this figure, we can see that when $xy = 1$, $(\Delta x, \Delta y)$ is tangential to the boundary of the primal feasible region, and that its length is not 0 unless the current point is optimal. In the following, we will see that the primal discrete sequence (x, y) can be trapped in the curved boundary, while u remains negative.

Letting the step length $\hat{\alpha}(x, y, u)$ absorb the common factor, we can write one iteration of the primal-dual affine-scaling algorithm in terms of (x, y, u) as follows:

$$x^+ = x + \lambda\hat{\alpha}(x, y, u)(2 - xy - x^2), \tag{34}$$

$$y^+ = y + \lambda\hat{\alpha}(x, y, u)(2 - xy - y^2), \tag{35}$$

$$u^+ = u + 2\lambda\hat{\alpha}(x, y, u)(1 - u^2), \tag{36}$$

where λ is a fixed fraction less than 1 and $\hat{\alpha}(x, y, u)$ is defined by (16). Here, we emphasize the fact that $\hat{\alpha}$, which is originally a function of (X, u, Z), can be regarded as a function of (x, y, u) due to the correspondence (22) and (24). In fact, we identify (x, y, u) and (X, u, Z) in the following.

Now we consider the set

$$\mathcal{G} := \{\, (x, y, u) \in \mathcal{F} \mid u \leq 0, \quad 1 < xy \leq 3/2, \quad x + y \geq 3 \,\}, \tag{37}$$

and investigate the property of the iteration sequence starting in this region. In fact, our aim in this section is to prove the following theorem:

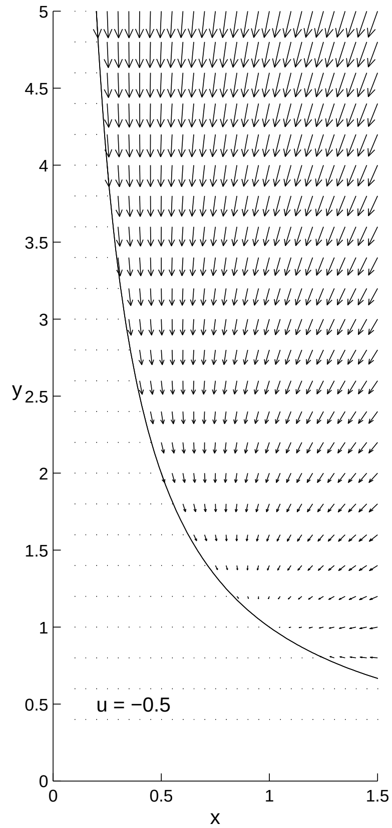

FIGURE 1. Vector Field of the HRVW/KSH/M method

Theorem 1. *Let for any* $2/3 \le \eta < 1$,

$$\mathcal{G}_\eta := \{ (x, y, u) \in \mathcal{G} \mid x \le 1 - \eta \} .$$

If, for the HRVW/KSH/M primal-dual affine-scaling algorithm (34), (35) *and* (36)*, we choose the initial point* $(x^0, y^0, u^0) \in \mathcal{G}$ *to satisfy:*

$$\sqrt{x^0 y^0 - 1} \le \frac{\lambda}{2\sqrt{2}} \min\left(\frac{-u^0}{2}, 1 - \eta - x^0 \right), \tag{38}$$

then the limit point is contained in the closure of $\mathcal{G}_\eta$*.*

Since the closure of $\mathcal{G}_\eta$ does not contain the optimal solution, this theorem implies that the sequence converges to a non-optimal point.

Note also that the condition (38) can be satisfied for all λ and η. In fact, fixing $x^0 < 1-\eta$ and $u^0 < 0$, we can reduce the left hand side arbitrarily by choosing y^0 close to $1/x^0$.

We first show that $\hat{\alpha} = \alpha_P$ on $\mathcal{G}$.

Lemma 2. *If* $(x, y, u) \in \mathcal{G}$*, then* $\hat{\alpha}(x, y, u)$ *is a positive solution of*

$$- R(x, y)\alpha^2 - 2(x + y)\alpha + 1 = 0, \tag{39}$$

where

$$R(x,y)=\frac{(2-xy)(x+y)^2-4}{xy-1}\geq\frac{1}{2(xy-1)}>0 \tag{40}$$

on $\mathcal{G}$.

Proof. Noting that $2(1-u^2)>0$ on the interior feasible region, we have

$$\alpha_D=\sup\left\{\alpha\ \middle|\ u+2\alpha(1-u^2)\leq 1\right\}=\frac{1-u}{2(1-u^2)}=\frac{1}{2(1+u)}>\frac{1}{2}$$

on $\mathcal{G}$.

For the primal problem (20), since $x^+>0$ and $y^+>0$ hold when $x^+y^+\geq 1$, α_P is the solution of $x^+y^+=1$, namely,

$$(x-\alpha_P(2-x(x+y)))\,(y-\alpha_P(2-y(x+y)))=1.$$

Expanding this quadratic equation and dividing by $xy-1$, we have

$$-R(x,y)\alpha_P^2-2(x+y)\alpha_P+1=0. \tag{41}$$

Now we have (40) as

$$R(x,y)=\frac{(2-xy)(x+y)^2-4}{xy-1}\geq\frac{9/2-4}{xy-1}=\frac{1}{2(xy-1)}>0$$

on $\mathcal{G}$. Since the coefficient of α_P^2 and the constant of (41) have the opposite signs, this quadratic equation has one positive solution and one negative, and α_P is the positive solution.

From (41), it follows that

$$-2(x+y)\alpha_P+1=R(x,y)\alpha_P^2>0,$$

from which we have

$$\alpha_P<\frac{1}{2(x+y)}\leq\frac{1}{2}\leq\alpha_D.$$

Therefore, we have $\hat{\alpha}=\alpha_P$ if $(x,y)\in\mathcal{G}$, which is the solution of (41). □

The following two lemmas are used to prove that the sum of $\hat{\alpha}^k$ is bounded, which is essential for the proof of the theorem.

Lemma 3. *We have*

$$\hat{\alpha}(x,y,u)\leq 1/\sqrt{R(x,y)}$$

on $\mathcal{G}$.

Proof. It follows from Lemma 2 that

$$\hat{\alpha}(x,y,u)^2=\frac{1-2(x+y)\hat{\alpha}(x,y,u)}{R(x,y)}\leq\frac{1}{R(x,y)}.$$

Thus we have

$$\hat{\alpha}(x,y,u)\leq 1/\sqrt{R(x,y)}.$$

□

Lemma 4. *Assume that we do one iteration of the primal-dual affine-scaling algorithm ((34),(35) and (36)) from $(x,y,u)\in\mathcal{G}$ to get (x^+,y^+,u^+) with fraction λ. Then we have*

$$\frac{x^+y^+-1}{xy-1}\leq 1-\lambda^2.$$

Proof. The lemma can be seen as follows:

$$\begin{aligned}
\frac{x^+y^+-1}{xy-1} &= 1-2\lambda\hat{\alpha}(x,y,u)(x+y)-R(x,y)\lambda^2\hat{\alpha}(x,y,u)^2 \\
&= 1-2\lambda\hat{\alpha}(x,y,u)(x+y)-\lambda^2(1-2\hat{\alpha}(x,y,u)(x+y)) \\
&= 1-\lambda^2-2\lambda\hat{\alpha}(x,y,u)(x+y)(1-\lambda^2) \\
&< 1-\lambda^2.
\end{aligned}$$

□

Lemma 5. *Assume that the sequence* $\{(x^k, y^k, u^k) \mid k = 0, \ldots, L\}$ *generated by* (34), (35) *and* (36) *is contained in* $\mathcal{G}$. *Then we have*

$$\sum_{k=0}^{L} \hat{\alpha}(x^k, y^k, u^k) \leq \frac{2\sqrt{2(x^0y^0-1)}}{\lambda^2}.$$

Proof. We have

$$\begin{aligned}
\sum_{k=0}^{L}\hat{\alpha}(x^k,y^k,u^k) &\leq \sum_{k=0}^{L}\frac{1}{\sqrt{R(x^k,y^k)}} && \text{(by Lemma 3)} \\
&\leq \sum_{k=0}^{L}\sqrt{2(x^ky^k-1)} && \text{(by (40))} \\
&\leq \sqrt{2(x^0y^0-1)}\sum_{k=0}^{L}\sqrt{(1-\lambda^2)^k} && \text{(by Lemma 4)} \\
&\leq \sqrt{2(x^0y^0-1)}\sum_{k=0}^{L}(1-\lambda^2/2)^k \\
&\leq \sqrt{2(x^0y^0-1)}\sum_{k=0}^{\infty}(1-\lambda^2/2)^k \\
&= \frac{2\sqrt{2(x^0y^0-1)}}{\lambda^2}.
\end{aligned}$$

□

Now we are ready to prove the theorem.

Proof of Theorem 1. We show that if $(x^l, y^l, u^l) \in \mathcal{G}_\eta$, then $(x^{l+1}, y^{l+1}, u^{l+1}) \in \mathcal{G}_\eta$, from which the theorem follows by induction. We have

$$\begin{aligned}
x^{l+1}-x^0 &= \sum_{k=0}^{l}\lambda\hat{\alpha}(x^k,y^k,u^k)(2-x^ky^k-(x^k)^2) \\
&\leq \lambda\sum_{k=0}^{l}\hat{\alpha}(x^k,y^k,u^k) && \text{(since } x^ky^k \geq 1) \\
&\leq \frac{2\sqrt{2(x^0y^0-1)}}{\lambda} && \text{(by Lemma 5)} \\
&\leq 1-\eta-x^0, && \text{(by (38))}
\end{aligned}$$

which implies

$$x^{l+1} \leq 1-\eta.$$

Similarly, we have

$$\begin{aligned} u^{l+1} - u^0 &= \sum_{k=0}^{l} 2\lambda\hat{\alpha}(x^k, y^k, u^k)(1-(u^k)^2) \\ &\le \lambda \sum_{k=0}^{l} 2\hat{\alpha}(x^k, y^k, u^k) \\ &\le \frac{4\sqrt{2(x^0y^0-1)}}{\lambda} && \text{(by Lemma 5)} \\ &\le -u^0, && \text{(by (38))} \end{aligned}$$

which implies

$$u^{l+1} \le 0.$$

From Lemma 4, $x^{l+1}y^{l+1} \le x^k y^k \le 3/2$ follows. The relation $x^{l+1}y^{l+1} > 1$ is obvious due to the choice of the step-size. Also $x^{l+1} \le 1-\eta \le 1/3$ implies $y^{l+1} \ge 1/x^{l+1} \ge 3$, from which we have $x^{l+1}+y^{l+1} \ge 3$. Therefore, $(x^{l+1}, y^{l+1}, u^{l+1}) \in \mathcal{G}_\eta$, which completes the proof. □

Remark: By replacing $3/2$ with $1+\epsilon$ and 3 with $2+2\epsilon$ in the definition (37) of $\mathcal{G}$, the same analysis provides an initial point arbitrary close to the primal optimal solution but for which convergence is to a non-optimal point.

4. The MT direction

We will show in this section that the MT direction applied with the primal and dual interchanged is identical to the HRVW/KSH/M direction for our primal-dual pair of SDP problems (20) and (21). As is well-known, we can transform the standard form SDP problem to the dual form and vice versa to get the following primal-dual pair $\langle\bar{D}\rangle$, $\langle\bar{P}\rangle$ of SDP problems:

$$\langle\bar{D}\rangle \left\{ \begin{array}{ll} \text{minimize} & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \bullet Z \\ \text{subject to} & \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \bullet Z = 1, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \bullet Z = 1, \quad Z \succeq 0, \end{array} \right. \tag{42}$$

$$\langle\bar{P}\rangle \left\{ \begin{array}{ll} \text{maximize} & -x-y \\ \text{subject to} & X - x\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - y\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ & X \succeq 0, \end{array} \right. \tag{43}$$

which is equivalent to $\langle D\rangle$ and $\langle P\rangle$. In fact, the feasible solutions for $\langle\bar{P}\rangle$ and $\langle\bar{D}\rangle$ are again given by (22) and (24) where $(x, y, u) \in \mathcal{F}$.

According to (7), (8), (10) and (11), the MT direction $(\Delta X, \Delta x, \Delta y, \Delta Z)$ for this primal-dual pair at a feasible solution (X, x, y, Z) is the solution of

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \bullet \Delta Z = 0, \tag{44}$$

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \bullet \Delta Z = 0, \tag{45}$$

$$\Delta X - \Delta x \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \Delta y \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = 0, \tag{46}$$

$$VXZ^{1/2} + Z^{1/2}XV + Z^{1/2}\Delta X Z^{1/2} = -Z^{1/2}XZ^{1/2} \tag{47}$$

$$VZ^{1/2} + Z^{1/2}V = \Delta Z, \tag{48}$$

where $V \in \mathcal{S}(2)$, or equivalently, (26) and (47) and (48). The following lemma shows that the MT direction is the same as the HRVW/KSH/M direction in our problem.

Lemma 6. *For (X, Z) satisfying* (22) *and* (24) *with $(x, y, u) \in \mathcal{F}^o$, the system* (47), (48) *and* (26) *has a unique solution $(\Delta X_M, \Delta Z_M, V_M)$. Let $(\widehat{\Delta X}_H, \Delta X_H, \Delta u_H, \Delta Z_H)$ be the solution of* (26), (29), *and* (30) *for the same (X, Z). Then we have $\Delta X_M = \Delta X_H$ and $\Delta Z_M = \Delta Z_H$.*

Proof. From (29) and (30), it is easy to see that ΔX_H is a unique solution of

$$(X\Delta Z Z^{-1} + Z^{-1}\Delta Z X)/2 + \Delta X = -X. \tag{49}$$

We prove the lemma by showing that (47) and (48) are equivalent to (49) in our case.

In view of (24), we can write

$$Z^{1/2} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

where θ satisfies $\cos\theta > 0$ and $2\cos\theta\sin\theta = -u$. Putting

$$V = \begin{pmatrix} p & q \\ q & r \end{pmatrix},$$

we have

$$VZ^{1/2} = \begin{pmatrix} p\cos\theta + q\sin\theta & p\sin\theta + q\cos\theta \\ q\cos\theta + r\sin\theta & q\sin\theta + r\cos\theta \end{pmatrix}.$$

Due to (48) and (26), the diagonal components of $VZ^{1/2}$ must be 0, i.e.,

$$p\cos\theta + q\sin\theta = 0,$$

$$q\sin\theta + r\cos\theta = 0.$$

Therefore, we have $p = r$, which implies that $VZ^{1/2}$ is symmetric.

Now we have

$$VZ^{1/2} = Z^{1/2}V = \Delta Z/2,$$

from which

$$V = Z^{-1/2}\Delta Z/2 = \Delta Z Z^{-1/2}/2$$

follow. Substituting these relations into (47), we have

$$(Z^{-1/2}\Delta Z X Z^{1/2} + Z^{1/2}X\Delta Z Z^{-1/2})/2 + Z^{1/2}\Delta X Z^{1/2} = -Z^{1/2}XZ^{1/2}.$$

Obviously, $(\Delta X_M, \Delta Z_M)$ is a solution of this system. Multiplying this equation by $Z^{-1/2}$ from the right and left, we have (49). Since the solution of (26) and (49) is unique, the MT direction is unique and identical to the HRVW/KSH/M direction. □

The following theorem is immediate by Lemma 6.

Theorem 7. *Let for any* $2/3 \leq \eta < 1$,

$$\mathcal{G}_\eta := \{\, (x, y, u) \in \mathcal{G} \mid x \leq 1 - \eta \,\}.$$

For the long-step primal-dual affine-scaling algorithm using the MT direction, if, given a step-size parameter λ*, we choose the initial point* $(x^0, y^0, u^0) \in \mathcal{G}$ *to satisfy:*

$$\sqrt{x^0 y^0 - 1} \leq \frac{\lambda}{2\sqrt{2}} \min\left(\frac{-u^0}{2}, 1 - \eta - x^0 \right),$$

then the limit point is contained in the closure of $\mathcal{G}_\eta$.

5. The AHO direction

We deal with the continuous trajectories of the AHO directions on our problem in this section. Let us denote the AHO direction by $(\Delta X, \Delta u, \Delta Z)$. The system for the direction is (27), (28), and (13), or equivalently, (26) and (13). The third equation, (13), can be written componentwise as follows:

$$\begin{array}{rrrcl} -2\Delta u & +2\Delta x & & = & 2(u - x), \\ -(x + y)\Delta u & -u\Delta x & -u\Delta y & = & ux + uy - 2, \\ -2\Delta u & & +2\Delta y & = & 2(u - y). \end{array}$$

Solving these linear equalities, we have

$$\Delta u = \frac{2(1 - u^2)}{x + y + 2u}, \tag{50}$$

$$\Delta x = \frac{2 - xy - x^2 - u(x - y)}{x + y + 2u}, \tag{51}$$

$$\Delta y = \frac{2 - xy - y^2 - u(y - x)}{x + y + 2u}. \tag{52}$$

Figure 2 shows the vector field $(\Delta x, \Delta y)$ on the primal feasible region with $u = -0.5$ and $u = 0.5$. In contrast with the HRVW/KSH/M direction case, the vector field drastically changes depending on u. Namely, when $u = -0.5$ and (x, y) is near the boundary of the primal feasible region, the direction is not nearly tangential to the boundary, aiming at somewhere outside of the feasible region. On the other hand when $u = 0.5$, the direction aims inside. The former observation leads to the convergence of the continuous trajectories of the AHO direction to a non-optimal point (Theorem 9).

We deal with the trajectory (17), (18) and (19) in the space of (x, y, u) by using the one-to-one correspondence (22) and (24). Furthermore, since the trajectory is not changed if we multiply each right-hand side by a common positive factor, we can multiply by $x + y + 2u$ which is greater than 0, to get

$$\dot{x}_t = 2 - x_t y_t - x_t^2 - u_t(x_t - y_t), \tag{53}$$

$$\dot{y}_t = 2 - x_t y_t - y_t^2 - u_t(y_t - x_t), \tag{54}$$

$$\dot{u}_t = 2(1 - u_t^2). \tag{55}$$

The equation (55) can be easily solved as follows:

$$u_t = \frac{(1 + u_0)e^{4t} - (1 - u_0)}{(1 + u_0)e^{4t} + (1 - u_0)}, \tag{56}$$

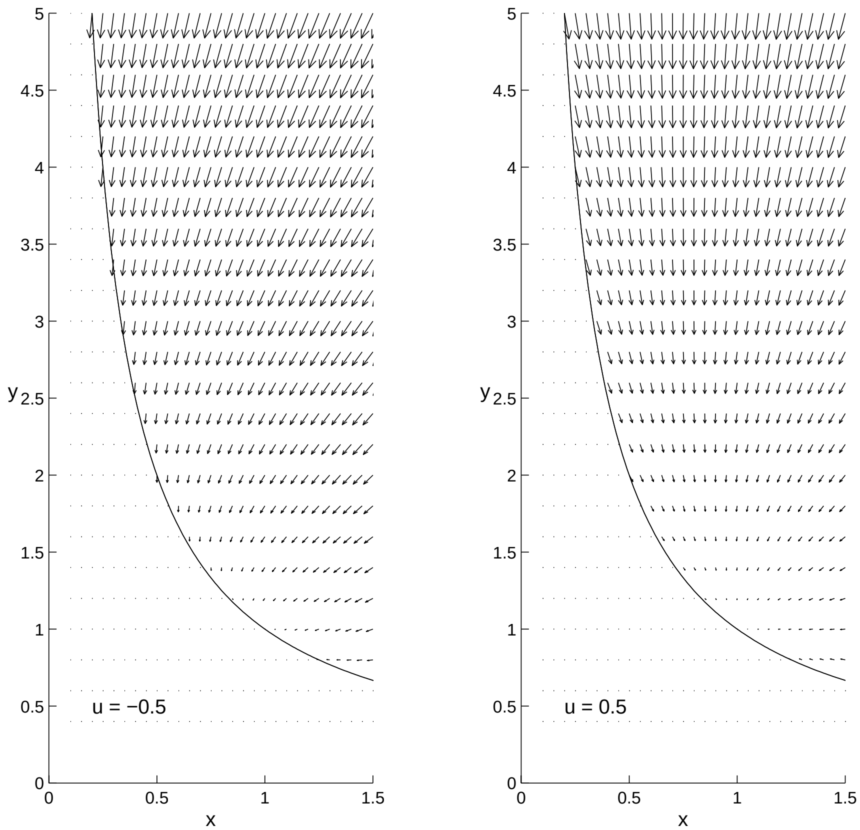

FIGURE 2. Vector Fields of the AHO method

where u_0 is the initial value of u_t.

The following properties of the vector field can easily be observed.

Lemma 8. *We have*

$$\begin{aligned} x_t\dot{y}_t + y_t\dot{x}_t &= u_t(x_t - y_t)^2 - 2(x_t y_t - 1)(x_t + y_t) && (57)\\ &\leq u_t(x_t - y_t)^2 && (58)\\ x_t\dot{y}_t - y_t\dot{x}_t &= (x_t - y_t)(2 + u_t(x_t + y_t)). && (59) \end{aligned}$$

Proof. We omit subscript t in this proof for simplicity. The former equation can be seen as:

$$\begin{aligned} x\dot{y} + y\dot{x} &= x(2 - xy - y^2 - uy + ux) + y(2 - xy - x^2 - ux + uy)\\ &= 2(x + y) - 2xy^2 - 2x^2y - 2uxy + ux^2 + uy^2\\ &= 2(x + y) - 2xy(x + y) + u(x - y)^2\\ &= u(x - y)^2 - 2(xy - 1)(x + y)\\ &\leq u(x - y)^2. \end{aligned}$$

The latter equation can be seen as:

$$\begin{aligned} x\dot{y} - y\dot{x} &= x(2 - xy - y^2 - uy + ux) - y(2 - xy - x^2 - ux + uy) \\ &= 2x - 2y + ux^2 - uy^2 \\ &= (x - y)(2 + u(x + y)). \end{aligned}$$

□

Now we restrict our attention to the set

$$\mathcal{H} := \{ (x, y, u) \in \mathcal{F} \mid u \leq -1/2, \quad y \geq 16x \}. \tag{60}$$

We then introduce the following change of variables:

$$r = \sqrt{xy}, \tag{61}$$

$$\theta = \frac{1}{2} \log \frac{y}{x}. \tag{62}$$

The inverse mapping is:

$$\begin{aligned} x &= re^{-\theta}, \\ y &= re^{\theta}. \end{aligned}$$

Putting

$$\Phi(x, y, u) := (r, \theta, u),$$

we can easily see that

$$\Phi(\mathcal{H}) = \{ (r, \theta, u) \mid -1 \leq u \leq -1/2, \quad r \geq 1, \quad \theta \geq \log 4 \}.$$

Now consider the trajectory in the new coordinate system:

$$(r_t, \theta_t, u_t) := \Phi(x_t, y_t, u_t) \tag{63}$$

starting from $(r_0, \theta_0, u_0) \in \Phi(\mathcal{H})$, and define

$$\begin{aligned} \hat{t} &:= \sup\{ t > 0 \mid (r_t, \theta_t, u_t) \in \Phi(\mathcal{F}) \} \\ \bar{t} &:= \sup\{ t > 0 \mid (r_t, \theta_t, u_t) \in \Phi(\mathcal{H}) \}. \end{aligned}$$

We use $(\hat{x}, \hat{y}, \hat{u})$, $(\hat{r}, \hat{\theta}, \hat{u})$, $(\bar{x}, \bar{y}, \bar{u})$, $(\bar{r}, \bar{\theta}, \bar{u})$ for $(x_{\hat{t}}, y_{\hat{t}}, u_{\hat{t}})$, $(r_{\hat{t}}, \theta_{\hat{t}}, u_{\hat{t}})$, $(x_{\bar{t}}, y_{\bar{t}}, u_{\bar{t}})$, $(r_{\bar{t}}, \theta_{\bar{t}}, u_{\bar{t}})$, respectively, for notational simplicity.

We will prove the following theorem in this section:

Theorem 9. *Let the initial point (x_0, y_0, u_0) be in $\mathcal{H}$ and let $(r_0, \theta_0, u_0) = \Phi(x_0, y_0, u_0)$ denote the corresponding point in $\Phi(\mathcal{H})$. If*

$$r_0^2 - 1 < \log \frac{1 - u_0}{3(1 + u_0)},$$

then $(x_{\hat{t}}, y_{\hat{t}}, u_{\hat{t}}) \in \mathcal{H}$ whereas $(x^*, y^*, u^*) \notin \mathcal{H}$.

The following lemma elucidates the behavior of the continuous trajectories on $\Phi(\mathcal{H})$

Lemma 10. *For $0 \leq t \leq \bar{t}$, we have*

$$r_t^2 \leq r_0^2 - 4t, \tag{64}$$

$$\theta_t \leq \theta_0. \tag{65}$$

Proof. It follows from (61) and (62) that

$$\dot{r}_t = \frac{x_t\dot{y}_t + y_t\dot{x}_t}{2r_t}, \tag{66}$$

$$\dot{\theta}_t = \frac{x_t\dot{y}_t - y_t\dot{x}_t}{2r_t^2}. \tag{67}$$

We have from (66) that

$$\begin{aligned} \frac{d}{dt}(r_t^2) &= x_t\dot{y}_t + y\dot{x}_t \\ &\le u_t(x_t - y_t)^2 \quad \text{(Use (58))} \\ &< -4 \quad \text{(Since } y - x > 3 \text{ and } u \le 1/2 \text{ on } \mathcal{H}\text{).} \end{aligned}$$

Therefore, we have

$$r_t^2 - r_0^2 < -4t.$$

The second assertion of the lemma can be easily derived from (59) and (67), since $x - y < 0$ and $2 + u(x+y) \le 2 - (x+y)/2 < 0$ on $\mathcal{H}$. □

Now we prove the theorem.

Proof of Theorem 9. Obviously, if $\bar{r} = 1$, $\bar{\theta} > \log 4$, and $\bar{u} < -1/2$, then the solution cannot be extended in the feasible region any more, i.e., $\hat{t} = \bar{t}$. Since $\bar{\theta} > \theta_0 \ge \log 4$ follows from (65), we will show that $\bar{r} = 1$ and $\bar{u} < -1/2$ in the following.

Since $r_t \ge 1$ on $\Phi(\mathcal{H})$, we have from (64) that t must satisfy

$$t \le \frac{r_0^2 - 1}{4}$$

as far as $(r_t, \theta_t, u_t) \in \Phi(\mathcal{H})$. In other words, we have

$$\bar{t} \le \frac{r_0^2 - 1}{4} < \frac{1}{4}\log\frac{1 - u_0}{3(1 + u_0)}. \tag{68}$$

On the other hand, in view of (56), we have

$$\begin{aligned} & u_t < 1/2 \\ \Leftrightarrow\ & \frac{(1+u_0)e^{4t} - (1-u_0)}{(1+u_0)e^{4t} + (1-u_0)} < -1/2 \\ \Leftrightarrow\ & (1+u_0)e^{4t} - (1-u_0) < -((1+u_0)e^{4t} + (1-u_0))/2 \\ \Leftrightarrow\ & 3(1+u_0)e^{4t}/2 < (1-u_0)/2 \\ \Leftrightarrow\ & t < \frac{1}{4}\log\frac{1-u_0}{3(1+u_0)}. \end{aligned}$$

Therefore, from (68), we have $\bar{u} < -1/2$, and since $(\bar{r}, \bar{\theta}, \bar{u})$ is at the boundary of $\Phi(\mathcal{H})$, we have $r_{\bar{t}} = 1$. □

6. THE NT DIRECTION

In this section, we prove that the long-step primal-dual affine-scaling algorithm using the NT direction generates a sequence converging to an optimal point for our SDP problem. We denote the NT direction by $(\Delta X, \Delta u, dZ)$. To calculate the NT direction, we first calculate the scaling matrix D. From (22), (24) and (15), we see that

$$D = \frac{1}{\rho\sigma} \begin{pmatrix} \rho x + 1 & \rho + u \\ \rho + u & \rho y + 1 \end{pmatrix}, \tag{69}$$

where

$$\sigma = \sigma(x, y, u) = \sqrt{x + y - 2u + 2\sqrt{(1 - u^2)(xy - 1)}}, \tag{70}$$

$$\rho = \rho(x, y, u) = \sqrt{\frac{1 - u^2}{xy - 1}}. \tag{71}$$

Solving (7), (8) and (14) with (69), we have

$$\begin{aligned} \Delta x &= \frac{(\rho + u)^2 - (\rho x + 1)^2}{\rho\phi} && (72) \\ &= \frac{(1 - x^2)\rho^2 + 2(u - x)\rho - (1 - u^2)}{\rho\phi}, && (73) \\ \Delta y &= \frac{(\rho + u)^2 - (\rho y + 1)^2}{\rho\phi}, && (74) \\ &= \frac{(1 - y^2)\rho^2 + 2(u - y)\rho - (1 - u^2)}{\rho\phi} && (75) \\ \Delta u &= \frac{\rho^2\sigma^2}{\phi} && (76) \\ &= \frac{\rho^2(x + y - 2u) + 2\rho(1 - u^2)}{\phi}, && (77) \end{aligned}$$

where

$$\phi = \phi(x, y, u) = (\rho y + 1)(\rho x + 1) + (\rho + u)^2. \tag{78}$$

The figure 3 shows the vector fields $(\Delta x, \Delta y)$ on the primal feasible region with $u = -0.5$ and $u = 0.5$. Note that the direction depends on u like the AHO direction does, but the length of the direction differs from that of the AHO direction; the length of the direction becomes small if (x, y) is close to the boundary, and in fact, it vanishes at the boundary. Now imagine that a discrete sequence generated by the primal-dual affine-scaling algorithm using the NT direction approaches the boundary $\{ (x, y, u) \in \mathcal{F} \mid xy = 1 \}$. Intuitively saying, the movement in the (x, y) space becomes small if xy is close to 1, and instead u is improved so much that u converges to 1, which is the optimal solution. Then the primal direction $(\Delta x, \Delta y)$ aims inside the feasible region, and the stagnation ends.

The following is what we prove in this section.

Theorem 11. *For any step-size parameter* $0 < \lambda < 1$*, and any initial point* (x^0, y^0, u^0)*, the sequence generated by the long-step primal-dual affine-scaling algorithm for the primal-dual pair of SDP problems* (20) *and* (21) *using the NT direction converges to the optimal point.*

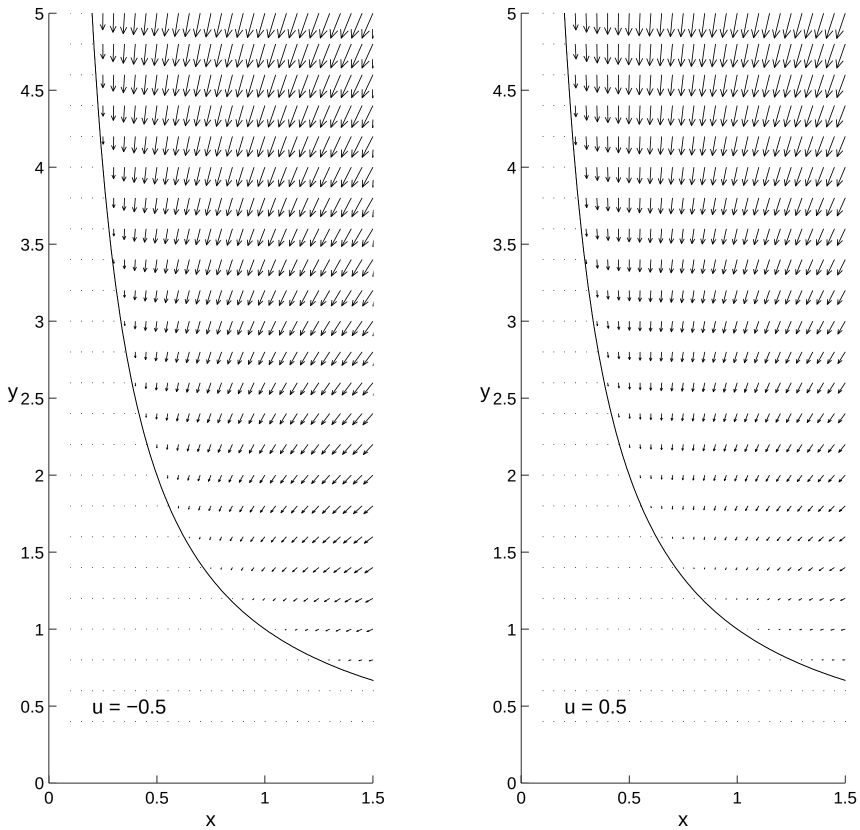

FIGURE 3. Vector Fields of the NT method

First, we observe that the duality gap $X \bullet Z = x + y - 2u$ is monotonically decreasing.

Lemma 12. *We have*

$$x^+ + y^+ - 2u^+ = (1 - \lambda\hat{\alpha})(x + y - 2u),$$

or equivalently,

$$\Delta x + \Delta y - 2\Delta u = -(x + y - 2u). \tag{79}$$

This is a standard calculation, thus we omit the proof. Note that if the duality gap does not converge to 0, then $\prod_{k=1}^{\infty}(1 - \lambda\hat{\alpha}^k) > 0$, which implies $\hat{\alpha}^k \to 0$. On the other hand, if $x^k + y^k - 2u^k \to 0$, then, since the optimal solution is unique, the sequence (x^k, y^k, u^k) converges to the optimal solution $(1, 1, 1)$. We use these relations in the following extensively.

Next lemma shows that the sequence (x^k, y^k, u^k) converges, and the search direction is bounded along the sequence.

Lemma 13. *We have* $(x^k, y^k, u^k) \to (x^\infty, y^\infty, u^\infty)$, *and* $(\Delta x^k, \Delta y^k, \Delta u^k)$ *is bounded.*

Proof. From (77), $\Delta u^k > 0$ follows. Since $\{u^k\}$ is an increasing sequence and bounded by 1, the limit u^∞ exists.

We have from Lemma 12 that

$$x^k + y^k \leq x^0 + y^0 - 2u^0 + 2u^k \leq x^0 + y^0 - 2u^0 + 2,$$

which implies that (x^k, y^k) is bounded, since $x^k > 0$ and $y^k > 0$.

By definition (78), we have

$$\phi \geq xy\rho + 1 \geq \rho + 1 \geq 1.$$

Therefore, since (x^k, y^k, u^k) is bounded, we have

$$\begin{aligned} |\Delta x^k| &= (\phi^k)^{-1} \left|\rho^k(1-(x^k)^2) + 2(u^k - x^k) - (1-(u^k)^2)/\rho^k\right| \\ &\leq \frac{\rho^k \left|1-(x^k)^2\right|}{1+(\rho^k)^2} + 2\left|u^k - x^k\right| + \sqrt{(x^k y^k - 1)(1-(u^k)^2)} \\ &\leq \left|1-(x^k)^2\right|/2 + 2\left|u^k - x^k\right| + \sqrt{(x^k y^k - 1)(1-(u^k)^2)} \\ &\leq M \end{aligned}$$

for some positive constant M. We see in the same way that Δy^k is bounded, and, from (79), that Δu^k is also bounded.

If $x^k + y^k - 2u^k \to 0$, obviously the sequence converges to the optimal solution. Therefore, we deal with the case that $x^k + y^k - 2u^k \to \hat{\delta} > 0$. Then Lemma 12 implies that there exists some $\delta > 0$ such that

$$\prod_{k=0}^{\infty}(1 - \lambda\hat{\alpha}^k) \geq \delta.$$

Taking logarithm of the both sides, we have

$$\log\delta \leq \sum_{k=0}^{\infty}\log(1-\lambda\hat{\alpha}^k) \leq -\lambda\sum_{k=0}^{\infty}\hat{\alpha}^k.$$

Using this inequality, we have

$$\sum_{k=0}^{l}|x^{k+1} - x^k| \leq \sum_{k=0}^{l}\lambda\hat{\alpha}^k|\Delta x^k| \leq -M\log\delta < \infty$$

for all l, which implies that $\{x^k\}$ is a Cauchy sequence. Thus $\{x^k\}$ converges. The convergence of $\{y^k\}$ can be shown in the same way. □

Using the lemma above, we prove that the dual iterates converges to its optimal.

Lemma 14. *We have* $u^k \to 1$.

Proof. Let us assume that $u^\infty < 1$. Since $(x^\infty, y^\infty, u^\infty)$ cannot be an interior point, we have $x^k y^k \to 1$.

If $\hat{\alpha}^k = \alpha_D^k$ occurs infinitely many times, then obviously $u^k \to 1$, which contradicts the assumption. Thus we can assume that $\alpha_P^k = \hat{\alpha}^k$ for sufficiently large k and that $\alpha_P^k \to 0$.

On the other hand, we have

$$\begin{aligned}
\alpha_P^k &= \sup\{\alpha \mid X + \alpha\Delta X \succeq 0\} \\
&= \sup\left\{\alpha \,\middle|\, D^{-1}XD^{-1} + \alpha D^{-1}\Delta X D^{-1} \succeq 0\right\} \\
&= \sup\{\alpha \mid Z + \alpha(-Z - \Delta Z) \succeq 0\} \\
&= \sup\left\{\alpha \,\middle|\, \begin{pmatrix} 1-\alpha & -(1-\alpha)u + \alpha\Delta u \\ -(1-\alpha)u + \alpha\Delta u & 1-\alpha \end{pmatrix} \succeq 0\right\} \\
&= \sup\left\{\alpha \,\middle|\, \alpha \le 1,\ (1-\alpha)^2 - ((1-\alpha)u - \alpha\Delta u)^2 \ge 0\right\}.
\end{aligned}$$

Therefore, α_P^k satisfies

$$(1-\alpha_P^k)^2 - ((1-\alpha_P^k)u^k - \alpha_P^k\Delta u^k)^2 = 0. \tag{80}$$

Since Δu^k is bounded, we have $\alpha_P^k \Delta u^k \to 0$, which implies that the left hand side of (80) converges to $1-(u^\infty)^2 > 0$, while the right hand side is 0. This is a contradiction, and we have $u^\infty = 1$. □

Now we know that $u^k \to 1$, and (x^k, y^k) is converging. We will prove $(x^k, y^k) \to (1, 1)$ in the following. To show this, we first show that the limit point is on the boundary of the primal feasible region.

Lemma 15. *We have $x^\infty y^\infty = 1$.*

Proof. Assume that $x^\infty y^\infty = 1 + \delta > 1$. In this case, we have $\rho^k \to 0$ and $\phi^k \to 2$ from definitions (71) and (78), and also $\hat{\alpha}^k \to 0$ from Lemma 12. Since

$$\begin{aligned}
\alpha_D^k &= \frac{1-u^k}{\Delta u^k} = \frac{(1-u^k)\phi^k}{(\rho^k)^2(\sigma^k)^2} \\
&= \frac{\phi^k(x^k y^k - 1)}{(1+u^k)\left(x^k + y^k - 2u^k + 2\sqrt{(1-(u^k)^2)(x^k y^k - 1)}\right)} \\
&\to \delta/(x^\infty + y^\infty - 2) > 0,
\end{aligned}$$

we see that $\alpha_P^k = \hat{\alpha}^k$ for sufficiently large k and that $\alpha_P^k \to 0$.

For α_P^k, we have

$$\begin{aligned}
\alpha_P^k &= \sup\{\alpha \mid X + \alpha\Delta X \succeq 0\} \\
&= \sup\left\{\alpha \,\middle|\, \begin{pmatrix} x^k + \alpha\Delta x^k & 1 \\ 1 & y^k + \alpha\Delta y^k \end{pmatrix} \succeq 0\right\} \\
&= \sup\left\{\alpha \,\middle|\, (x^k + \alpha\Delta x^k)(y^k + \alpha\Delta y^k) \ge 1\right\} \\
&= \sup\left\{\alpha \,\middle|\, x^k y^k - 1 + \alpha(x^k\Delta y^k + y^k\Delta x^k) + \alpha^2\Delta x^k\Delta y^k \ge 0\right\}.
\end{aligned}$$

Therefore, $\hat{\alpha}_P^k$ satisfies

$$\Delta x^k\Delta y^k(\alpha_P^k)^2 + (x^k\Delta y^k + y^k\Delta x^k)\alpha_P^k + x^k y^k - 1 = 0. \tag{81}$$

However, since (x^k, y^k) and $(\Delta x^k, \Delta y^k)$ are bounded, the left hand side of (81) goes to $x^\infty y^\infty - 1 = \delta > 0$, while the right hand side is 0. This is a contradiction, and we have $x^\infty y^\infty = 1$. □

The following relation can be seen by a straightforward calculation.

Lemma 16. *We have*

$$(x\Delta y + y\Delta x)\phi = 2u(x+y-2u) - \bar{\delta}(x,y,u),$$

where $\bar{\delta}(x,y,u) \to 0$ *when* $xy \to 1$ *and* $u \to 1$.

Proof. We have

$$\begin{aligned}
& x\Delta y + y\Delta x \\
&= \phi^{-1}\left\{x\left(\rho(1-y^2) + 2(u-y) - \sqrt{(1-u^2)(xy-1)}\right)\right. \\
&\quad \left. + y\left(\rho(1-x^2) + 2(u-x) - \sqrt{(1-u^2)(xy-1)}\right)\right\} \\
&= \phi^{-1}\left\{\rho(x+y)(1-xy) + 2u(x+y) - 4xy - (x+y)\sqrt{(1-u^2)(xy-1)}\right\} \\
&= \phi^{-1}\left\{2u(x+y-2u) - 4(1-u^2) - 4(xy-1)\right. \\
&\quad \left. -2(x+y)\sqrt{(1-u^2)(xy-1)}\right\}.
\end{aligned}$$

Therefore, putting

$$\bar{\delta}(x,y,u) := -4(1-u^2) - 4(xy-1) - 2(x+y)\sqrt{(1-u^2)(xy-1)},$$

we have the lemma. □

Now we are ready to prove that the optimality of (x^∞, y^∞). Obviously, this lemma together with Lemma 17 proves Theorem 11.

Lemma 17. *We have* $(x^k, y^k) \to (1,1)$.

Proof. It can be seen that

$$\begin{aligned}
\frac{x^{k+1}y^{k+1} - 1}{x^k y^k - 1} &= \frac{(x^k + \lambda\hat{\alpha}^k\Delta x^k)(y^k + \lambda\hat{\alpha}^k\Delta y^k) - 1}{x^k y^k - 1} \\
&= 1 + \frac{\lambda\hat{\alpha}^k(x^k\Delta y^k + y^k\Delta x^k) + \lambda^2(\hat{\alpha}^k)^2\Delta x^k\Delta y^k}{x^k y^k - 1} \\
&= 1 + \frac{\lambda\hat{\alpha}^k\left(\phi^k(x^k\Delta y^k + y^k\Delta x^k) + \lambda\hat{\alpha}^k\phi^k\Delta x^k\Delta y^k\right)}{\phi^k(x^k y^k - 1)}. \qquad (82)
\end{aligned}$$

We claim that $\phi^k\Delta x^k\Delta y^k$ is bounded. Assume by contradiction, $\phi^k\Delta x^k\Delta y^k$ is not bounded. Then we can take a diverging subsequence, i.e., there exists a subsequence $L \subset \{0,1,2,\ldots\}$ such that $\lim_{k\in L}\phi^k|\Delta x^k\Delta y^k| \to \infty$. Since Δx^k and Δy^k are bounded, we have $\lim_{k\in L}\phi^k \to \infty$, and from the definition of ϕ^k, $\lim_{k\in L}\rho^k \to \infty$, too. Therefore this is a contradiction because, for $k \in L$,

$$\begin{aligned}
\phi^k\Delta x^k\Delta y^k &= \frac{\left((\rho^k+u^k)^2 - (\rho^k x^k + 1)^2\right)\left((\rho^k+u^k)^2 - (\rho^k y^k + 1)^2\right)}{(\rho^k)^2\phi^k} \\
&= \frac{\left((1+u^k/\rho^k)^2 - (x^k + 1/\rho^k)^2\right)\left((1+u^k/\rho^k)^2 - (y^k + 1/\rho^k)^2\right)}{(y^k + 1/\rho^k)(x + 1/\rho^k) + (1 + u^k/\rho^k)^2} \\
&\to (1-x^\infty)(1-y^\infty)/2.
\end{aligned}$$

Assume that $(x^k, y^k) \not\to (1, 1)$. Then Lemma 12 implies that $\hat{\alpha}^k \to 0$. From Lemmas 14, 15, and 16, we have that

$$(x^k \Delta y^k + y^k \Delta x^k)\phi^k \geq x^\infty + y^\infty - 2 > 0 \tag{83}$$

for sufficiently large k, while, since $\phi^k \Delta x^k \Delta y^k$ is bounded and $\hat{\alpha}^k \to 0$, $\hat{\alpha}^k \phi^k \Delta x^k \Delta y^k \to 0$. Therefore,

$$(x^k \Delta y^k + y^k \Delta x^k)\phi^k + \lambda \hat{\alpha}^k \phi^k \Delta x^k \Delta y^k > 0$$

holds for sufficiently large k. This and (82) imply that

$$x^{k+1} y^{k+1} - 1 > x^k y^k - 1$$

i.e., $x^k y^k - 1$ is increasing for sufficiently large k. This contradicts the fact that $x^k y^k \to 1$. Therefore, we have $(x^k, y^k) \to (1, 1)$. □

7. Concluding Remarks

The practical success of interior-point methods for LP relies heavily on the ability to take the long steps, i.e., stepping a fixed fraction of the way to the boundary for the next iterate. Even when convergence has not been proved, it is necessary in practice to take such a long step. For LP, these long steps are very successful, and every implementation uses bold step-length parameters.

These bold choices of step-length parameters are supported by the robustness of the primal-dual affine-scaling algorithm (not the Dikin-type variant). It is known that the continuous trajectories associated with the primal-dual affine-scaling algorithm converge to the optimal solution, and there is no evidence so far that the long-step primal-dual affine-scaling algorithm fails to find the optimal solution.

However in SDP, the situation is different; even a continuous trajectory can converge to a non-optimal point. The results of this paper suggest that, for finding the optimal solution, such bold steps as are taken in the LP case should not be taken at least for the HRVW/KSH/M, MT and AHO directions; otherwise, jamming may occur.

It seems that the algorithm corresponding to the NT direction is more robust than those corresponding to the other directions. The same observation was reported by Todd, Toh and Tütüncü [35].

Acknowledgements

We thank Professor Michael Overton of New York University for showing us the bad local behavior of the HRVW/KSH/M method, which motivated us to do this research. We also thank Dr. Masayuki Shida of Kanagawa University for many stimulating discussions which inspired us to develop the results for the AHO direction, and kindly pointing out that the MT direction is identical to the HRVW/KSH/M direction in our example.

References

[1] F. Alizadeh. Interior point methods in semidefinite programming with application to combinatorial optimization. *SIAM Journal on Optimization*, 5:13–51, 1995.

[2] F. Alizadeh, J. P. A. Haeberly, and M. L. Overton. Primal-dual interior-point methods for semidefinite programming. Technical report, 1994.

[3] F. Alizadeh, J. P. A. Haeberly, and M. L. Overton. Complementarity and nondegeneracy in semidefinite programming. Technical report, 1995.

[4] F. Alizadeh, J. P. A. Haeberly, and M. L. Overton. Primal-dual inteior-point methods for semidefinite programming: Convergence rates, stability and numerical results. Technical report, 1995.

[5] E.R. Barnes. A variation on Karmarkar's algorithm for solving linear programming problems. *Mathematical Programming*, 36:174–182, 1986.

[6] B. Borchers. CSDP, a C library for semidefinite programming. Technical report, 1997.

[7] E. de Klerk, C. Roos, and T. Terlaky. Polynomial primal-dual affine scaling algorithms in semidefinite programming. Technical Report 96-42, Delft University of Technology, 1996.

[8] I.I. Dikin. Iterative solution of problems of linear and quadratic programming. *Doklady Akademii Nauk SSSR*, 174:747–748, 1967. (Translated in: *Soviet Mathematics Doklady* 8(1967)674-675.).

[9] L. Faybusovich. Dikin's algorithm for matrix linear programming problems. In J. Henry and J. Pavon, editors, *Lecture notes in control and information sciences 197*, pages 237–247. Springer–Verlag, 1994.

[10] L. Faybusovich. On a matrix generalization of affine-scaling vector fields. *SIAM Journal on Matrix Analysis and Application*, 16(3):886–897, 1995.

[11] K. Fujisawa and M. Kojima. SDPA (semidefinite programming algorithm) user's manual. Technical Report B-308, Tokyo Institute of Technology, 1995.

[12] D. Goldfarb and K. Scheinberg. Interior point trajectories in semidefinite programming. Technical report, Dept. of IEOR, Columbia Univsersity, New York, NY, USA, 1996.

[13] C. Helmberg, F. Rendl, R.J. Vanderbei, and H. Wolkowicz. An interior point method for semidefinite programming. *SIAM Journal on Optimization*, 6:342–361, 1996.

[14] B. Jansen, C. Roos, and T. Terlaky. A polynomial primal-dual dikin-type algorithm for linear programming. Technical Report 93-36, Faculty of Technical Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands, 1993.

[15] F. Jarre. An interior-point method for minimizing the maximum eigenvalue of a linear combination of matrices. *SIAM Journal on Control and Optimization*, 31:1360–1377, 1993.

[16] N. K. Karmarkar. A new polynomial time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[17] M. Kojima, M. Shida, and S. Shindoh. A predictor-correctoe interior-point algorithm for the semidefinite linear complementarity problem using Alizadeh-Haeberly-Overton search direction. Technical report, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152, Japan, 1996.

[18] M. Kojima, S. Shindoh, and S. Hara. Interior point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. *SIAM Journal on Optimization*, 7(1):86–125, 1997.

[19] C-J. Lin and R. Saigal. A predictor-corrector method for semi-definite programming. Technical report, Dept. of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, Michigan 48109-2177, 1995.

[20] Z-Q. Luo, J. F. Sturm, and S. Zhang. Superlinear convergence of a symmetric primal-dual path-following algorithms for semidefinite programming. Technical Report 9607/A, Econometric Institute, Erasmus University, Rotterdam, The Netherlands, 1996.

[21] R. D. C. Monteiro. Primal-dual path-following algorithms for semidefinite programming. Technical report, School of Industrial and System Engineering, Georgia Tech, Atlanta, GA30332, 1995. (To appear in SIAM Journal on Optimization).

[22] R. D. C. Monteiro. Polynomial convergence of primal-dual algorithms for semidefinite programming based on Monteriro and Zhang family of directions. Technical report, School of Industrial and System Engineering, Georgia Tech, Atlanta, GA30332, 1996. (To appear in SIAM Journal on Optimization).

[23] 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. of Operations Research*, 15:191–214, 1990.

[24] R. D. C. Monteiro and T. Tsuchiya. Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming. Technical report, Georgia Institute of Technology, Atlanta, Georgia, USA, 1996.

[25] R. D. C. Monteiro and P. Zanjacomo. Implementation of primal-dual methods for semidefinite programming based on monteiro and tsuchiya newton directions and their variants. Technical report, Georgia Institute of Technology, Atlanta, Georgia, USA, 1997.

[26] R. D. C. Monteiro and Y. Zhang. A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming. Technical report, 1996.

[27] M. Muramatsu. Affine scaling algorithm fails for semidefinite programming. Technical Report 16, Department of Mechanical Engineering, Sophia University, 1996.

[28] Y. E. Nesterov and A. S. Nemirovskii. Optimization over positive semidefinite matrices: Mathematical background and user's manual. Technical report, Central Economic & Mathematical Institute, USSR Acad. Sci. Moscow, USSR, 1990.

[29] Y. E. Nesterov and A. S. Nemirovskii. *Interior Point Polynomial Methods in Convex Programming : Theory and Algorithms*. SIAM Publications, Philadelphia, 1993.

[30] Y. E. Nesterov and M. J. Todd. Primal-dual interior-point methods for self-scaled cones. Technical Report 1125, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, 14853-3801, 1995.

[31] Y. E. Nesterov and M. J. Todd. Self-scaled barriers and interior-point methods for convex programming. Technical Report 1091, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, 14853-3801, 1995.

[32] F. A. Potra and R. Sheng. A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. Technical Report 78, Department of Mathematics, The University of Iowa, 1995.

[33] M. Shida, S. Shindoh, and M. Kojima. Existence of search directions in interior-point algorithms for the sdp and the monotone SDLCP. Technical report, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152, Japan, 1996.

[34] J. F. Sturm and S. Zhang. Symmetric primal-dual path-following algorithms for semidefinite programming. Technical Report 9554/A, Ecometric Institute, Erasmus University Rotterdam, The Netherlands, 1995.

[35] M. J. Todd, K. C. Toh, and R. H. Tütüncü. On the Nesterov-Todd direction in semidefinite programming. Technical report, School of Operations Research and Industrial Engineering, Cornell University, 1996.

[36] P. Tseng. Analysis of infeasible path-following methods using the Alizadeh-Haeberly-Overton direction for the monotone semidefinite lcp. Technical report, Department of Mathematics, University of Washington, Seattle, Washington 98195, U.S.A., 1996.

[37] T. Tsuchiya and M. Muramatsu. Global convergence of a long-step affine scaling algorithm for degenerate linear programming problems. *SIAM Journal on Optimization*, 5(3):525–551, 1995.

[38] L. Vandenberghe and S. Boyd. Primal-dual potential reduction method for problems involving matrix inequalities. *Mathematical Programming*, 69:205–236, 1995.

[39] R. J. Vanderbei, M. S. Meketon, and B. F. Freedman. A modification of Karmarkar's linear programming algorithm. *Algorithmica*, 1:395–407, 1986.

[40] Y. Zhang. On extending primal-dual interior-point algorithms from linear programming to semidefinite programming. Technical Report TR95-20, Dept. of Math/Stat, University of Maryland, Baltimore County, Baltimore, Maryland, USA, 1995.

MASAKAZU MURAMATSU, PRINCETON UNIVERSITY, PRINCETON, NJ 08544

ROBERT J. VANDERBEI, PRINCETON UNIVERSITY, PRINCETON, NJ 08544