Curvature Integrals and Iteration Complexities in SDP and Symmetric Cone Programs

Satoshi Kakihara * Atsumi Ohara † Takashi Tsuchiya‡

September 30, 2011

Abstract

In this paper, we study iteration complexities of Mizuno-Todd-Ye predictor-corrector (MTY-PC) algorithms in SDP and symmetric cone programs by way of curvature integrals. The curvature integral is defined along the central path, reflecting the geometric structure of the central path. The idea to exploit the curvature of the central path for the analysis of iteration complexities is based on the intuitive observation that path-following algorithms trace faster in relatively straight parts than in the rest of the curved parts. Our analyses in this paper are direct extensions to SDP and symmetric cone programs, of Monteiro and Tsuchiya [10], which gave a rigorous asymptotic estimate on iteration complexity of MTY-PC algorithms in Linear programming using the aforementioned curvature integral by tending the opening of the neighborhood to zero. More specifically, we give concrete formulas for curvature integrals in SDP and symmetric cone programs and give asymptotic estimates for iteration complexities. In particular, we conduct numerical experiments in practically large SDP problems from SDPLIB, which suggests that our results serve as a useful analytical tool for various SDP problems.

Keywords: interior-point methods, primal-dual algorithms, path-following methods, iteration complexities, curvature, semidefinite programming, symmetric cone programs

1 Introduction

Let us consider the standard LP problems

$$\begin{array}{llrcl} (P) & \min_x & c^T x & & \\ & & Ax & = & b, \\ & & x & \geq & 0, \end{array}$$

and its Lagrange dual

$$\begin{array}{llrcccl} (D) & \max_{y,s} & b^T y & & & & \\ & & A^T y & + & s & = & c, \\ & & & & s & \geq & 0, \end{array}$$

where $c \in \mathbb{R}^n, A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ are the data, and $x \in \mathbb{R}^n$ and $(y, s) \in \mathbb{R}^m \times \mathbb{R}^n$ are the primal and dual variables, respectively. The central path for (P) and (D) is defined as the solution set for

$$\begin{array}{rcccl} & A^T y & + & s & = & c, \\ Ax & & & & = & b, \\ & Xs & & & = & \nu e, \end{array} \qquad (1)$$

*National Graduate Institute for Policy Studies 7-22-1 Roppongi, Minato-ku, Tokyo 106-8677, Japan. (Email: s-kakihara@grips.ac.jp).

†Electrical and Electronics Engineering, Graduate School of Engineering, University of Fukui, 3-9-1 Bunkyo, Fukui City, Fukui 910-8507, Japan. (Email: ohara@fuee.u-fukui.ac.jp).

‡National Graduate Institute for Policy Studies, 7-22-1 Roppongi, Minato-ku, Tokyo 106-8677, Japan. (Email: tsuchiya@grips.ac.jp). This author is supported in part with Grant-in-Aid for Scientific Research (B), 2007, 19360192 from the Japan Society for the Promotion of Sciences.

where X is the diagonal matrix whose ith diagonal element is x_i and $e = (1, \ldots, 1)^T$. Since the solution set of (1) is unique for each $\nu > 0$, the central path is a well-defined trajectory parameterized by $\nu > 0$. Mizuno-Todd-Ye predictor-corrector (MTY-PC) algorithm [7], which is of our concern in this paper, is a predictor-corrector type path-following algorithm; the algorithm traces the central path to seek an optimal solution, alternately going along the Newton direction in predictor step and returning back to the central path in corrector step. Iteration complexity analysis using the curvature of the central path is motivated from the intuitive observations that path-following algorithms trace the central path faster in relatively straight parts than in the rests of curved parts. The objective of this paper is to give concrete formulas for the curvature integrals along the central path and establish the link between curvature integrals and iteration complexities of MTY-PC algorithms in SDP and symmetric cone programs.

The idea to exploit the curvature of the central path for the analysis of iteration complexities originated from Karmarkar [6] using Riemannian geometry in a certain type of LP. Subsequently, Sonnevend, Stoer and Zhao [15] and Zhao and Stoer [17] gave a concrete formula for the curvature integral and the bound for iteration complexity using this integral in a certain class of LP.

Our analyses in this paper are direct extensions to SDP and symmetric cone programs, of Monteiro and Tsuchiya [10], which gave a rigorous asymptotic estimate to iteration complexity of MTY-PC algorithms in LP using the aforementioned curvature integral by letting the opening of the neighborhood to zero. More specifically, we give concrete formulas for curvature integrals in SDP and symmetric cone programs and give asymptotic estimates for iteration complexities using these integrals. Supplementarily, we conduct numerical experiments in practically large SDP problems from SDPLIB, which suggests that our results serve as a useful analytical tool for various SDP problems.

Ohara and Tsuchiya developed an information geometric framework of interior-point algorithms and demonstrated that the aforementioned curvature integral in the case of LP is written as an information geometric quantity in a rigorous meaning [13]. In an accompanying paper [5], we will extend this result to SDP and symmetric cone programs, namely, the curvature integral developed here is exactly the same quantity as the one in the case of LP in view of the information geometric framework.

This paper is organized as follows. In section 2, we discuss basics of the primal-dual algorithms in SDP along with MTY-PC algorithms. In section 3, we concisely study the iteration complexity via the curvature integrals in SDP case. In section 4, we conduct numerical experiments for SDP problems from SDPLIB, and see how our result is demonstrated in numerical experiments. In section 5, we first discuss symmetric cone programs in view of Euclidean Jordan algebra, and then briefly outline the iteration complexity analysis via the curvature integrals in symmetric cone programs. We end with the conclusion in section 6.

2 Primal-Dual Algorithms in SDP

Let $\mathbb{S}^n$ be a real vector space generated by $n \times n$ real matrices, and its associated inner product is defined by $A \bullet B := \mathrm{Tr}(AB)$ for any $A, B \in \mathbb{S}^n$. $X \succeq 0$ and $X \succ 0$ denote the positive semidefiniteness and positive definiteness of X, respectively. Let $\mathbb{R}^m$ be a m-dimensional real vector space.

The standard primal form of SDP is formulated as

$$\begin{array}{lllll} (P) & \min_X & C \bullet X & & \\ & & A_i \bullet X & = & b_i, \quad i = 1, \ldots, m, \\ & & X & \succeq & 0, \end{array} \tag{2}$$

and its dual is formulated as

$$\begin{array}{llllll} (D) & \max_{y,S} & b^T y & & & \\ & & \sum_{i=1}^m y_i A_i & + \; S & = & C, \\ & & & S & \succeq & 0, \end{array} \tag{3}$$

where $C \in \mathbb{S}^n, A_i \in \mathbb{S}^n, \; i = 1, \ldots, m$ and $b = (b_1, \ldots, b_m)^T \in \mathbb{R}^m$ are the given data, and $X \in \mathbb{S}^n$ and $(y, S) \in \mathbb{R}^m \times \mathbb{S}^n$ are the primal and dual variables, respectively. We assume that $A_i, i = 1, \ldots, m$ are linearly independent.

The *feasible regions* to (2) and (3) are defined by

$$\begin{aligned} \mathcal{P} &:= \{X \in \mathbb{S}^n \mid A_i \bullet X = b_i,\ i = 1, \ldots, m, \quad X \succeq 0\}, &(4)\\ \mathcal{D} &:= \{(y, S) \in \mathbb{R}^m \times \mathbb{S}^n \mid \textstyle\sum_{i=1}^m y_i A_i + S = C, \quad S \succeq 0\}, &(5) \end{aligned}$$

and its corresponding *strictly feasible regions* are defined by

$$\begin{aligned} \mathcal{P}^+ &:= \{X \in \mathcal{P} \mid X \succ 0\}, &(6)\\ \mathcal{D}^+ &:= \{(y, S) \in \mathcal{D} \mid S \succ 0\}. &(7) \end{aligned}$$

We assume that $\mathcal{P}^+ \neq \emptyset$ and $\mathcal{D}^+ \neq \emptyset$.

Using the *(normalized) duality gap*, defined by $\mu(X, S) := X \bullet S/n$, we define the distance measure by

$$d_F(X, S) \;:=\; \left\|X^{1/2} S X^{1/2} - \mu(X, S) I\right\|_F \;=\; \left[\textstyle\sum_{i=1}^n (\lambda_i(XS) - \mu)^2\right]^{1/2}, \tag{8}$$

where $X^{1/2}$ for $X \succ 0$ is the positive semidefinite square root. For a given constant $0 < \beta < 1$, we then define the neighborhood

$$\mathcal{N}(\beta) \;:=\; \{(X, y, S) \in \mathcal{P} \times \mathcal{D} \mid d_F(X, S) \leq \beta\mu(X, S)\}.$$

Let us simplify the notation slightly: we define the operator $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ by

$$(\mathcal{A}X)_i := A_i \bullet X, \quad i = 1, \ldots, m.$$

Then the adjoint of this operator is $\mathcal{A}^* : \mathbb{R}^m \to \mathbb{S}^n$ satisfying

$$\mathcal{A}^* y = \textstyle\sum_{i=1}^m y_i A_i.$$

It is known that the system:

$$\begin{array}{rcccl} & \mathcal{A}^* y & + & S & = C, \\ \mathcal{A}X & & & & = b, \\ & XS & & & = \nu I, \end{array} \tag{9}$$

has a unique solution in $(X, y, S) \in \mathcal{P}^+ \times \mathcal{D}^+$ for any $\nu > 0$ and the set of such solutions are called a *central path*.

Now we introduce Monteiro-Tsuchiya (MT) family of directions [9] and Monteiro-Zhang (MZ) family of directions [11].

MT family of directions are obtained by linearizing the scaled quadratic form of the central path equation:

$$(PXP^T)^{1/2}(P^{-T} S P^{-1})(PXP^T)^{1/2} - \mu I = 0, \tag{10}$$

with a nonsingular matrix P. More specifically, they are obtained from the following system of equations:

$$\begin{array}{cccl} & \mathcal{A}^* \Delta y & +\Delta S & = C - \mathcal{A}^* y - S, \\ \mathcal{A}\Delta X & & & = b - \mathcal{A}X, \\ U\widetilde{S}\widetilde{X}^{1/2} + & \widetilde{X}^{1/2}\widetilde{S}U & +\widetilde{X}^{1/2}\widetilde{\Delta S}\widetilde{X}^{1/2} & = \sigma\mu I - \widetilde{X}^{1/2}\widetilde{S}\widetilde{X}^{1/2}, \\ U\widetilde{X} + & \widetilde{X}U & & = \widetilde{\Delta X}, \end{array} \tag{11}$$

where $\widetilde{X} = PXP^T$, $\widetilde{S} = P^{-T} S P^{-1}$, $\widetilde{\Delta X} = P\Delta X P^T$ and $\widetilde{\Delta S} = P^{-T}\Delta S P^{-1}$. MZ family of directions are obtained from the system:

$$\begin{array}{rcccl} & \mathcal{A}^* \Delta y & +\Delta S & = & C - S - \mathcal{A}^* y, \\ \mathcal{A}\Delta X & & & = & b - \mathcal{A}X, \\ \mathcal{H}_P(\Delta X S & + & X\Delta S) & = & \sigma\mu I - \mathcal{H}_P(XS), \end{array} \tag{12}$$

using an operator $\mathcal{H}_P : \mathbb{R}^{n\times n} \to \mathbb{S}^n$,

$$\mathcal{H}_P(M) = \frac{1}{2}\left[PMP^{-1} + (PMP^{-1})^T\right], \tag{13}$$

with a nonsingular scaling matrix P.

In both families of directions, σ is called the centering parameter; when $\sigma = 0$, the derived direction is the Newton direction or an affine scaling direction and when $\sigma = 1$, it is a centering direction.

We will now state the scheme we will consider in this paper.

MTY Predictor-Corrector Algorithm in SDP:

Fix some constants $\beta \in (0, 1)$.

Let $w^0 = (X^0, y^0, S^0) \in \mathcal{N}\left(\beta^2\right)$ and $\mu_F < \mu_0 := (X^0 \bullet S^0)/n$ be given.

Repeat until $\mu_k \le \mu_f$, **do**

(1) Choose a nonsingular scaling matrix P^k;

(2) Compute the solution $\Delta w^k = (\Delta X^k, \Delta y^k, \Delta S^k)$ of the system (11) in MT direction or (12) in MZ direction with $P = P^k$, $\mu = \mu_k$, $\sigma = 0$ and $(X, y, S) = (X^k, y^k, S^k)$;

(3) Set $z^{k+1} := w^k + \alpha_k \Delta w^k$, where $\alpha_k > 0$ is the largest $\alpha > 0$ such that $w^k + \tilde{\alpha}\Delta w^k \in \mathcal{N}(\beta)$ for all $\tilde{\alpha} \in [0, \alpha]$;

(4) From z^{k+1} compute a point $w^{k+1} = (X^{k+1}, y^{k+1}, S^{k+1}) \in \mathcal{N}\left(\beta^2\right)$ with the same duality gap as w^{k+1};

(4) Set $\mu_{k+1} := (X^{k+1} \bullet S^{k+1})/n$ and increment k by 1.

End do

End

3 Curvature Integrals and Iteration Complexity in SDP

In this section, we discuss the explicit relationship between the curvature integral along the central path and the number of iterations in MTY-PC algorithms in SDP, in case of the infinitesimal opening of the neighborhood.

Let $w_\nu := (X_\nu, y_\nu, S_\nu)$ be the point on the central path whose duality gap is ν and let

$$\widetilde{X} := PXP^T, \quad \widetilde{S} := P^{-T}SP^{-1},$$
$$W_P(w) := \widetilde{X}^{1/2}\widetilde{S}\widetilde{X}^{1/2},$$

where P is an invertible matrix and we drop P from $W_P(\cdot)$ when P is identity. We define the square of the curvature by

$$h_{PD}(\nu) := \frac{\|W''(X_\nu, S_\nu)\|_F}{2\nu}, \tag{14}$$

where W'' is the second-order derivative along the vector $(\dot{X}, \dot{y}, \dot{S}) := \left(\frac{\mathrm{d}X}{\mathrm{d}\nu}, \frac{\mathrm{d}y}{\mathrm{d}\nu}, \frac{\mathrm{d}S}{\mathrm{d}\nu}\right)$. Let $I_{PD}(\nu_f, \nu_i)$ be the integral of the curvature from the duality gap ν_f to ν_i, defined by

$$I_{PD}(\nu_f, \nu_i) := \int_{\nu_f}^{\nu_i} h_{PD}^{1/2}(\nu)\mathrm{d}\nu. \tag{15}$$

The main result of this section, Theorem 3.7, shows that using Jordan product $\circ$ with $A \circ B = 1/2(AB + BA)$ for $\forall A, B \in \mathbb{S}^n$, the curvature integral is represented by

$$I_{PD}(\nu_f, \nu_i) = \int_{\nu_f}^{\nu_i} \left\| \left(X^{-1/2}\dot{X}X^{-1/2}\right) \circ \left(S^{-1/2}\dot{S}S^{-1/2}\right) \right\|_F^{1/2} \mathrm{d}\nu, \tag{16}$$

in the presence of MZ scaling, and we prove the relation:

$$I_{PD}(\nu_f, \nu_i) \approx \sqrt{\beta} \cdot \#_{PD}(\nu_i, \nu_f, \beta), \tag{17}$$

where $\#_{PD}(\nu_i, \nu_f, \beta)$ denotes the number of iteration for MTY-PC algorithm to reduce the duality gap from ν_i to ν_f when the opening of the neighborhood is β.

The outline of the proof for Theorem 3.7 is as follows. We first show that the relation (17) holds for an idealized MTY-PC algorithm in the presence of MT family of directions in Theorem 3.1. Then we extends this to a generic algorithm with MZ family of directions in Theorem 3.4. Next, we prove that the formula (16) holds for HKM dual direction in Lemma 3.5 and then show that this formula is scaling invariant in Lemma 3.6. Finally, combining, Theorem 3.4, Lemma 3.5 and Lemma 3.6, we obtain Theorem 3.7.

We introduce the first theorem which establishes the relation (17) in case of the idealized MTY-PC algorithm with MT family of directions, in which every iteration starts on the central path and returns exactly on the central path.

Theorem 3.1. *[Idealized MTY-PC algorithm in SDP] Let $\beta \in (0, 1/2]$. For given w^0 on the central path and $0 < \nu_f < \mu(w^0)$ denote by $\#_{PD}(\mu(w^0), \nu_f, \beta)$ the number of iterations of the MTY-PC algorithm with $\beta \in (0, 1/2]$ needed to reduce the duality gap from $\nu_i := \mu(w^0)$ to ν_f. Then, using the curvature integral:*

$$I_{PD}(\nu_f, \nu_i) = \int_{\nu_f}^{\nu_i} h_{PD}^{1/2}(\nu)\mathrm{d}\nu, \qquad h_{PD}(\nu) := \left\| \frac{W''(X_\nu, S_\nu)}{2\nu} \right\|_F, \tag{18}$$

we have

$$\lim_{\beta \downarrow 0} \frac{I_{PD}(\nu_f, \nu_i)/\sqrt{\beta}}{\#_{PD}(\nu_i, \nu_f, \beta)} = 1.$$

Before going to the proof, we make some preliminary definitions and discussions. For an invertible matrix P, we define

$$\begin{gathered}
\widetilde{\Delta X} := P\Delta X P^T, \quad \widetilde{\Delta S} := P^{-T}\Delta S P^{-1}, \\
X(\alpha) := X + \alpha\Delta X, \quad S(\alpha) := S + \alpha\Delta S, \\
\widetilde{X}(\alpha) := P(X + \alpha\Delta X)P^T, \quad \widetilde{S}(\alpha) := P^{-T}(S + \alpha\Delta S)P^{-1}, \\
\mu := \frac{X \bullet S}{n} = \frac{\widetilde{X} \bullet \widetilde{S}}{n}, \quad \mu(\alpha) := \frac{X(\alpha) \bullet S(\alpha)}{n} = \frac{\widetilde{X}(\alpha) \bullet \widetilde{S}(\alpha)}{n}.
\end{gathered}$$

We define the function $\phi(\alpha)$, the norm of whose function $\|\phi(\alpha)\|_F$ represents the distance from the central path to $(X(\alpha), y(\alpha), S(\alpha))$.

Definition 3.1. Let $w = (X, y, S) \in \mathcal{N}\left(\beta^2\right)$ and P be an invertible matrix in $\mathbb{R}^{n\times n}$. Let $(\Delta X, \Delta y, \Delta S)$ be either MZ family (12) or MT family (11) of directions at the predictor step obtained with $\sigma = 0$. Then we define

$$\phi(\alpha) := \widetilde{X}(\alpha)^{1/2}\widetilde{S}(\alpha)\widetilde{X}(\alpha)^{1/2} - \mu(\alpha)I.$$

Note that it is easy to see that

$$\phi(\alpha) = \widetilde{X}(\alpha)^{1/2}\widetilde{S}(\alpha)\widetilde{X}(\alpha)^{1/2} - (1-\alpha)\mu I.$$

If w is on the central path, we denote $\phi_c(\alpha) := \phi(\alpha)$. In the next lemma, we compute the first and second order derivative of $\phi(\alpha)$.

Lemma 3.2. *Let $\phi'(\alpha)$ and $\phi''(\alpha)$ be the first and second order derivatives of $\phi(\alpha)$ w.r.t. α. Then we have*

$$\begin{aligned}
\phi'(\alpha) &= U\widetilde{S}(\alpha)\widetilde{X}(\alpha)^{1/2} + \widetilde{X}(\alpha)^{1/2}\widetilde{S}(\alpha)U + \widetilde{X}(\alpha)^{1/2}\widetilde{\Delta S}\widetilde{X}(\alpha)^{1/2} + \mu I, \\
\phi''(\alpha) &= U'\widetilde{S}(\alpha)\widetilde{X}(\alpha)^{1/2} + \widetilde{X}(\alpha)^{1/2}\widetilde{S}(\alpha)U' + 2U\widetilde{\Delta S}\widetilde{X}(\alpha)^{1/2} \\
&\quad +2\widetilde{X}(\alpha)^{1/2}\widetilde{\Delta S}U + 2U\widetilde{S}(\alpha)U,
\end{aligned}$$

where U and U' satisfy

$$\begin{aligned}
\widetilde{X}(\alpha)^{1/2}U + U\widetilde{X}(\alpha)^{1/2} &= \widetilde{\Delta X}, \\
\widetilde{X}(\alpha)^{1/2}U' + U'\widetilde{X}(\alpha)^{1/2} &= -2U^2.
\end{aligned}$$

Proof. The proof of this lemma is found in Lemma 3.2 in [9]. □

In subsequent lemma, we consider the derivatives of $W_P(w)$ on the central path along the vector $(\dot{X}, \dot{y}, \dot{S}) := \left(\frac{\mathrm{d}X}{\mathrm{d}\nu}, \frac{\mathrm{d}y}{\mathrm{d}\nu}, \frac{\mathrm{d}S}{\mathrm{d}\nu}\right)$.

Lemma 3.3. *Let $w = (X, y, S)$ be the point on the central path. We define the tangent vector of the central path at the point w by $\dot{w} := \frac{\mathrm{d}w}{\mathrm{d}\nu}$ or $(\dot{X}, \dot{y}, \dot{S}) := \left(\frac{\mathrm{d}X}{\mathrm{d}\nu}, \frac{\mathrm{d}y}{\mathrm{d}\nu}, \frac{\mathrm{d}S}{\mathrm{d}\nu}\right)$. Let $\dot{\widetilde{X}} = P\dot{X}P^T$ and $\dot{\widetilde{S}} = P^{-T}\dot{S}P^{-1}$. The first and second order directional derivatives of $W_P(w)$ along the tangent vector of the central path $\dot{w} = (\dot{X}, \dot{y}, \dot{S})$ are given by*

$$\begin{aligned} W'(w) &= V\widetilde{S}\widetilde{X}^{1/2} + \widetilde{X}^{1/2}\widetilde{S}V + \widetilde{X}^{1/2}\dot{\widetilde{S}}\widetilde{X}^{1/2}, \\ W''(w) &= V'\widetilde{S}\widetilde{X}^{1/2} + \widetilde{X}^{1/2}\widetilde{S}V' + 2\widetilde{X}^{1/2}\dot{\widetilde{S}}\widetilde{X}^{1/2} \\ &\quad +2\widetilde{X}^{1/2}\dot{\widetilde{S}}V + 2V\widetilde{S}V, \end{aligned}$$

where V and V' satisfy

$$\begin{aligned} \widetilde{X}^{1/2}V + V\widetilde{X}^{1/2} &= \dot{\widetilde{X}}, \\ \widetilde{X}^{1/2}V' + V'\widetilde{X}^{1/2} &= -2V^2. \end{aligned}$$

Proof. The proof of this lemma is very similar to Lemma 3.2. Or see Lemma 2.2 in [9]. □

Interior-point methods reduce the duality gap within the finite interval $[\nu_f, \nu_i]$ to seek an optimal solution, so we define the compact region $\mathcal{F}$ as follows. We later use the fact that variables (X, y, S) and the scaling matrix P, which is a continuous function of (X, y, S), are bounded on $\mathcal{F}$.

Definition 3.2.

$$\mathcal{F} := \mathcal{N}(\beta) \cap \{(X, y, S) \in \mathcal{P} \times \mathcal{D} \mid \nu_f \le \mu(X, S) \le \nu_i\}$$

Now we are ready to prove Theorem 3.1.

Proof. At a certain predictor step, the iterate w_ν begins on the central path. The MT direction Δw_ν is obtained from the system (11) with $\sigma = 0$. Note that the direction on the central path is unique regardless of the scaling matrix P. The necessary and sufficient condition for the point on the predictor step $w_\nu + \alpha\Delta w_\nu$ to be on the boundary of the outer neighborhood $\mathcal{N}(\beta)$ is

$$\|\phi_c(\alpha)\|_F = \beta\mu(\alpha). \tag{19}$$

By applying Taylor's theorem to ϕ_c near 0 up to second degree, (19) is equivalent to

$$\left\|\phi_c(0) + \alpha\phi_c'(0) + \frac{\alpha^2}{2}\phi_c''(0) + L\right\|_F = \beta\mu(\alpha), \tag{20}$$

where $\|L\|_F = \mathcal{O}\left(\alpha^3\right)$. Using the identities $\mu(\alpha) = (1-\alpha)\mu$, $\phi_c'(0) = -\phi_c(0) = 0$ and $\phi_c''(0) = \mu^2 W''$, it follows from (20) that

$$\left\|\frac{\alpha^2}{2}\mu^2 W'' + L\right\|_F = \beta(1-\alpha)\mu. \tag{21}$$

Letting $L' := L/\mu$, the equation (21) is then rewritten as

$$\begin{aligned} \alpha^2\left(\left\|\frac{\mu W''}{2}\right\|_F - \|L'\|_F/\alpha^2\right) \le \alpha^2\left\|\frac{\mu W''}{2} + L'/\alpha^2\right\|_F \le (1-\alpha)\beta, \\ \therefore \quad \alpha^2\left(m - \|L'\|_F/\alpha^2\right) \le \beta. \end{aligned}$$

where m is a lower bound of $\|\mu W''/2\|_F$ on the compact region $\mathcal{F}$. Letting β to 0, we have

$$\alpha^2 \le \frac{\beta}{m - \mathcal{O}(\alpha)} \to 0. \tag{22}$$

This implies that $\alpha = \mathcal{O}\left(\beta^{1/2}\right)$.

This conclusion together with (21) and triangular inequality implies

$$\begin{aligned} \left(1-\mathcal{O}\left(\beta^{1/2}\right)\right)\beta &= \left\|\alpha^2\frac{\mu W''}{2}+L'\right\|_F, \\ \beta\left(1-\mathcal{O}\left(\beta^{1/2}\right)\right)-\|L'\|_F &\leq \alpha^2\left\|\frac{\mu W''}{2}\right\|_F \leq \beta\left(1-\mathcal{O}\left(\beta^{1/2}\right)\right)+\|L'\|_F, \\ \beta\left(1-\mathcal{O}\left(\beta^{1/2}\right)\right)-\mathcal{O}\left(\beta^{3/2}\right) &\leq \alpha^2\left\|\frac{\mu W''}{2}\right\|_F \leq \beta\left(1-\mathcal{O}\left(\beta^{1/2}\right)\right)+\mathcal{O}\left(\beta^{3/2}\right), \\ \therefore\quad \alpha^2\left\|\frac{\mu W''}{2}\right\|_F &= \beta\left(1-\mathcal{O}\left(\beta^{1/2}\right)\right). \end{aligned}$$

We now obtain the square of the curvature $h_{PD}(\mu)$ as

$$h_{PD}(\mu) := \alpha^2\left\|\frac{\mu W''}{2}\right\|_F/\mu^2 = \alpha^2\left\|\frac{W''}{2\mu}\right\|_F = \frac{\beta}{\mu^2}\left(1-\mathcal{O}\left(\beta^{1/2}\right)\right). \tag{23}$$

We are ready to consider the curvature integral on the interval $[\nu_f,\nu_i]$. Let $w^0,\dots,w^K$ be a sequence generated by the MTY-PC algorithm with opening β satisfying the conditions that $\mu_0=\nu_i$ and $\mu_K\leq \nu_f<\mu_{K-1}$, where $\mu_k:=\mu(w^k)$ for all $k=0,\dots,K$.

Using the identity $\mu_{k+1}=(1-\alpha_k)\mu_k$, the mean value theorem and the compactness of the interval $[\mu_f,\mu_i]$, it is easy to see that

$$\begin{aligned} \int_{\mu_{k+1}}^{\mu_k} h_{PD}^{1/2}(\nu)\mathrm{d}\nu &= h_{PD}^{1/2}(\mu_k)(\mu_k-\mu_{k+1})+\mathcal{O}\left((\mu_k-\mu_{k+1})^2\right) \\ &= h_{PD}^{1/2}(\mu_k)(1-\alpha_k)\mu_k+\mathcal{O}\left((\alpha_k\mu_k)^2\right) \\ &= \beta^{1/2}\left(1-\mathcal{O}\left(\beta^{1/2}\right)\right)^{1/2}+\mathcal{O}\left(\beta\right), \end{aligned}$$

where the last equality follows from the fact that $\alpha_k=\mathcal{O}\left(\beta^{1/2}\right)$, and $\mu_k\leq\nu_i$ for all k. The last relation then implies

$$\begin{aligned} &\frac{I(\nu_f,\nu_i)/\sqrt{\beta}}{\#_{PD}(\mu(w^0),\nu_f,\beta)} = \frac{1}{K\beta^{1/2}}\int_{\nu_f}^{\nu_i} h_{PD}^{1/2}(\nu)\mathrm{d}\nu \\ =\ &\frac{1}{K\beta^{1/2}}\left(\sum_{k=0}^{K-2}\int_{\mu_{k+1}}^{\mu_k} h_{PD}^{1/2}(\nu)\mathrm{d}\nu+\int_{\nu_f}^{\mu_{K-1}} h_{PD}^{1/2}(\nu)d\nu\right) \\ =\ &\frac{1}{K\beta^{1/2}}\left((K-1)\beta^{1/2}\left(1-\mathcal{O}\left(\beta^{1/2}\right)\right)^{1/2}+(K-1)\mathcal{O}\left(\beta\right)+\mathcal{O}\left(\beta^{1/2}\right)\right). \end{aligned}$$

Since $\alpha_k=\mathcal{O}\left(\beta^{1/2}\right)$, we have that α_k converges to 0, as β tends to 0, for every $k=0,\dots,K-1$. This observation implies that $K=K(\beta)$ must converge to ∞ as β tends to 0. Using this observation in the previous relation, we conclude that the theorem holds. □

Next, we discuss the relation (17) in case of a generic MTY-PC algorithm. In this algorithm, each iterate starts from the inner neighborhood $\mathcal{N}\left(\beta^2\right)$ and returns to the same inner neighborhood. In the proof of the following theorem, we show that the deviation from the central path does not essentially affect the curvature integral, namely, we can bound the deviation with $\mathcal{O}\left(\beta^2\right)$. For this reason, we reduce the relation in (17) in a generic case to one in an idealized case.

Theorem 3.4. *[Generic MTY-PC algorithm in MZ family of directions in SDP] Let $\beta\in(0,1/2]$. For given $w^0\in\mathcal{N}\left(\beta^2\right)$ and $0<\nu_f<\mu(w^0)$ denote by $\#_{PD}(\mu(w^0),\nu_f,\beta)$ the number of iterations of the MTY-PC algorithm with $\beta\in(0,1/2]$ needed to reduce the duality gap from $\nu_i:=\mu(w^0)$ to ν_f. Then, using the curvature integral:*

$$I(\nu_f,\nu_i)=\int_{\nu_f}^{\nu_i} h_{PD}^{1/2}(\nu)\mathrm{d}\nu \quad \textit{and} \quad h_{PD}(\nu):=\left\|\frac{W''(X_\nu,S_\nu)}{2\nu}\right\|_F, \tag{24}$$

we have

$$\lim_{\beta\downarrow 0}\frac{I(\nu_f,\nu_i)/\sqrt{\beta}}{\#_{PD}(\nu_i,\nu_f,\beta)}=1.$$

Proof. As we mentioned, the main point of the proof is the evaluation of the deviation of $\phi(\alpha)$ between a generic algorithm and an idealized algorithm. To see this, we examine the boundary condition (20). $\phi(\alpha)$ can be written as

$$\phi(\alpha)=[\phi(0)-\phi_c(0)]+\phi_c(0)+\alpha\left[\phi'(0)-\phi_c'(0)\right]+\alpha\phi_c'(0)+\frac{\alpha^2}{2}\left[\phi''(0)-\phi_c''(0)\right]+\frac{\alpha^2}{2}\phi_c''(0)+L, \quad (25)$$

where $\|L\|_F=\mathcal{O}\left(\alpha^3\right)$. In the following, we will examine three estimations listed below:

i) $\|\phi(0)-\phi_c(0)\|_F=\mathcal{O}\left(\beta^2\right)$,

ii) $\|\phi'(0)-\phi_c'(0)\|_F=\mathcal{O}\left(\beta^2\right)$,

iii) $\|\phi''(0)-\phi_c''(0)\|_F=\mathcal{O}\left(\beta^2\right)$.

i) The proof of $\|\phi(0)-\phi_c(0)\|_F=\mathcal{O}\left(\beta^2\right)$:
It is easy to see that

$$\begin{aligned}\phi(0)-\phi_c(0) &= \left[\widetilde{X}^{1/2}\widetilde{S}\widetilde{X}^{1/2}-\mu I\right]-\left[\widetilde{X}_\nu^{1/2}\widetilde{S}_\nu\widetilde{X}_\nu^{1/2}-\mu I\right]\\ &= \widetilde{X}^{1/2}\widetilde{S}\widetilde{X}^{1/2}-\widetilde{X}_\nu^{1/2}\widetilde{S}_\nu\widetilde{X}_\nu^{1/2}\\ &= \left(\widetilde{X}^{1/2}-\widetilde{X}_\nu^{1/2}\right)\widetilde{S}\widetilde{X}^{1/2}+\widetilde{X}_\nu^{1/2}\left(\widetilde{S}-\widetilde{S}_\nu\right)\widetilde{X}^{1/2}+\widetilde{X}_\nu^{1/2}\widetilde{S}_\nu\left(\widetilde{X}^{1/2}-\widetilde{X}_\nu^{1/2}\right).\end{aligned}$$

We evaluate $\left\|\widetilde{X}^{1/2}-\widetilde{X}_\nu^{1/2}\right\|_F$ as follows: (C_1,C_2 are some constants)

$$\left\|\widetilde{X}^{1/2}-\widetilde{X}_\nu^{1/2}\right\|_F\le C_1\left\|\widetilde{X}-\widetilde{X}_\nu\right\|_F\le C_2\|X-X_\nu\|_F=\mathcal{O}\left(\beta^2\right).$$

The first and second evaluations follow from the continuity of the operator $(\cdot)^{1/2}$ and scaling matrix P_w, respectively, and the fact that these operations are done in the compact region $\mathcal{F}$. Note that scaling matrix P_w varies continuously over $\mathcal{P}\times\mathcal{D}$. $\|X-X_\nu\|=\mathcal{O}\left(\beta^2\right)$ is from Lemma 3.8. Similarly we have, with some constant C_3,

$$\left\|\widetilde{S}-\widetilde{S}_\nu\right\|_F\le C_3\|S-S_\nu\|_F=\mathcal{O}\left(\beta^2\right).$$

Also, each norm below is bounded on the compact region $\mathcal{F}$:

$$\left\|\widetilde{X}^{1/2}\right\|_F,\left\|\widetilde{X}_\nu^{1/2}\right\|_F,\left\|\widetilde{S}_\nu\right\|_F,\left\|\widetilde{S}\right\|_F. \quad (26)$$

Thus, it is concluded that

$$\|\phi(0)-\phi_c(0)\|_F=\mathcal{O}\left(\beta^2\right). \quad (27)$$

ii) The proof of $\|\phi'(0)-\phi_c'(0)\|_F=\mathcal{O}\left(\beta^2\right)$:
From Lemma 3.2, we have

$$\begin{aligned}\phi'(0)-\phi_c'(0) &= \left[U\widetilde{S}\widetilde{X}^{1/2}+\widetilde{X}^{1/2}\widetilde{S}U+\widetilde{X}^{1/2}\widetilde{\Delta S}\widetilde{X}^{1/2}+\mu I\right]\\ &\quad-\left[U_\nu\widetilde{S}_\nu\widetilde{X}_\nu^{1/2}+\widetilde{X}_\nu^{1/2}\widetilde{S}_\nu U_\nu+\widetilde{X}_\nu^{1/2}\widetilde{\Delta S}_\nu\widetilde{X}_\nu^{1/2}+\mu I\right]\\ &= \left[U\widetilde{S}\widetilde{X}^{1/2}-U_\nu\widetilde{S}_\nu\widetilde{X}_\nu^{1/2}\right]+\left[\widetilde{X}^{1/2}\widetilde{S}U-\widetilde{X}_\nu^{1/2}\widetilde{S}_\nu U_\nu\right]\\ &\quad+\left[\widetilde{X}^{1/2}\widetilde{\Delta S}\widetilde{X}^{1/2}-\widetilde{X}_\nu^{1/2}\widetilde{\Delta S}_\nu\widetilde{X}_\nu^{1/2}\right].\end{aligned}$$

We consider only the first term of the last equation (other terms can be evaluated in the same way), namely,

$$U\widetilde{S}\widetilde{X}^{1/2}-U_\nu\widetilde{S}_\nu\widetilde{X}_\nu^{1/2}=(U-U_\nu)\widetilde{S}\widetilde{X}^{1/2}+U_\nu\left(\widetilde{S}-\widetilde{S}_\nu\right)\widetilde{X}^{1/2}+U_\nu\widetilde{S}_\nu\left(\widetilde{X}^{1/2}-\widetilde{X}_\nu^{1/2}\right).$$

$\|U - U_\nu\|_F$ is evaluated as follows. Using Kronecker product, we have

$$\left(X^{1/2} \otimes I + I \otimes X^{1/2}\right) \mathrm{vec}\, U = \mathrm{vec}\, \Delta X,$$
$$\left(X_\nu^{1/2} \otimes I + I \otimes X_\nu^{1/2}\right) \mathrm{vec}\, U_\nu = \mathrm{vec}\, \Delta X_\nu.$$

Let $A := X_\nu^{1/2} \otimes I + I \otimes X_\nu^{1/2}$, and $\Delta A := \left(X^{1/2} - X_\nu^{1/2}\right) \otimes I + I \otimes \left(X^{1/2} - X_\nu^{1/2}\right)$. Then, using sensitivity analysis formula in Lemma 3.9, we obtain

$$\|U - U_\nu\|_F \leq \frac{\left\|A^{-1}\right\|_F}{1 - \|A^{-1}\|_F \|\Delta A\|_F} \left(\|\Delta X - \Delta X_\nu\|_F + \|\Delta A\|_F \|U_\nu\|_F\right). \tag{28}$$

By Lemma 3.10, we have $\|\Delta X - \Delta X_\nu\|_F = \mathcal{O}(\beta^2)$. A^{-1} exists since $X_\nu^{1/2}$ is positive definite. Also, $\left\|A^{-1}\right\|_F$, $\|\Delta A\|_F$, $\|U_\nu\|_F$ are bounded. From these facts, we obtain $\|U - U_\nu\|_F = \mathcal{O}(\beta^2)$. As we discussed in i), we have

$$\left\|\widetilde{S} - \widetilde{S}_\nu\right\|_F = \mathcal{O}\left(\beta^2\right) \quad \text{and} \quad \left\|X^{1/2} - X_\nu^{1/2}\right\|_F = \mathcal{O}\left(\beta^2\right),$$

and the followings are bounded on the compact region $\mathcal{F}$:

$$\left\|\widetilde{X}^{1/2}\right\|_F, \left\|\widetilde{X}_\nu^{1/2}\right\|_F, \left\|\widetilde{S}_\nu\right\|_F, \left\|\widetilde{S}\right\|_F. \tag{29}$$

Therefore, we have

$$\left\|U\widetilde{S}\widetilde{X}^{1/2} - U_\nu \widetilde{S}_\nu \widetilde{X}_\nu^{1/2}\right\|_F = \mathcal{O}\left(\beta^2\right).$$

Thus we conclude that

$$\|\phi'(0) - \phi_c'(0)\|_F = \mathcal{O}\left(\beta^2\right). \tag{30}$$

iii) The proof of $\|\phi''(0) - \phi_c''(0)\|_F = \mathcal{O}\left(\beta^2\right)$:
Although it is a little complicated, the proof can be done in a similar manner as i) and ii). From Lemma 3.2, we see

$$\begin{aligned}
\phi''(0) - \phi_c''(0) \quad = \quad & \left[U'\widetilde{S}\widetilde{X}^{1/2} + \widetilde{X}^{1/2}\widetilde{S}U' + 2U\widetilde{\Delta S}\widetilde{X}^{1/2} + 2\widetilde{X}^{1/2}\widetilde{\Delta S}U + 2U\widetilde{S}U\right] \\
& - \left[U_\nu'\widetilde{S}_\nu\widetilde{X}_\nu^{1/2} + \widetilde{X}_\nu^{1/2}\widetilde{S}_\nu U_\nu' + 2U_\nu\widetilde{\Delta S}_\nu\widetilde{X}_\nu^{1/2} + 2\widetilde{X}_\nu^{1/2}\widetilde{\Delta S}_\nu U_\nu + 2U_\nu\widetilde{S}_\nu U_\nu\right] \\
= \quad & \left[U'\widetilde{S}\widetilde{X}^{1/2} - U_\nu'\widetilde{S}_\nu\widetilde{X}_\nu^{1/2}\right] + \left[\widetilde{X}^{1/2}\widetilde{S}U' - \widetilde{X}_\nu^{1/2}\widetilde{S}_\nu U_\nu'\right] \\
& + 2\left[U\widetilde{\Delta S}\widetilde{X}^{1/2} - U_\nu\widetilde{\Delta S}_\nu\widetilde{X}_\nu^{1/2}\right] + 2\left[\widetilde{X}^{1/2}\widetilde{\Delta S}U - \widetilde{X}_\nu^{1/2}\widetilde{\Delta S}_\nu U_\nu\right] \\
& + 2\left[U\widetilde{S}U - U_\nu\widetilde{S}_\nu U_\nu\right].
\end{aligned}$$

In above equation, the evaluation of $\|U' - U_\nu'\|_F$ is new, but it is essentially identical to the evaluation of $\|U - U_\nu\|_F$ in ii), and it is easy to see that

$$\|U' - U_\nu'\|_F = \mathcal{O}\left(\beta^2\right).$$

In the rest of the parts of the equation, we go through the similar processes as in i) and ii), and we obtain

$$\|\phi''(0) - \phi_c''(0)\|_F = \mathcal{O}\left(\beta^2\right). \tag{31}$$

Therefore, $\phi(\alpha)$ in (25) is rewritten as

$$\phi(\alpha) = \phi_c(0) + \alpha\phi_c'(0) + \frac{\alpha^2}{2}\phi_c''(0) + L + \mathcal{O}\left(\beta^2\right). \tag{32}$$

In the generic algorithm, the boundary condition (20) of the idealized algorithm is now translated into

$$\left\|\phi_c(0) + \alpha\phi_c'(0) + \frac{\alpha^2}{2}\phi_c''(0) + L + \mathcal{O}\left(\beta^2\right)\right\|_F = \beta\mu(\alpha). \tag{33}$$

Using again the identities $\mu(\alpha) = (1-\alpha)\mu$, $\phi_c'(0) = -\phi_c(0) = 0$ and $\phi_c''(0) = \mu^2 W''$, we have

$$\left\| \frac{\alpha^2}{2}\mu^2 W'' + L + \mathcal{O}\left(\beta^2\right) \right\|_F = \beta(1-\alpha)\mu, \tag{34}$$

which is very similar to (21) in the idealized algorithm. As we did in Theorem 3.1, we first estimate the order of α, then estimate the order of $\|W''\|_F$. Letting $L' := L/\mu$, the equation (34) is then rewritten as

$$\begin{aligned} \alpha^2 \left(\left\| \frac{\mu W''}{2} \right\|_F - \|L'\|_F/\alpha^2 \right) \leq \alpha^2 \left\| \frac{\mu W''}{2} + L'/\alpha^2 \right\|_F \leq (1-\alpha)\beta + \mathcal{O}\left(\beta^2\right), \\ \therefore \quad \alpha^2 \left(m - \|L'\|_F/\alpha^2 \right) \leq \beta + \mathcal{O}\left(\beta^2\right), \end{aligned}$$

where m is a lower bound of $\|\mu W''/2\|_F$ on the region $\mathcal{F}$. Letting β to 0, we have

$$\alpha^2 \leq \frac{\beta + \mathcal{O}\left(\beta^2\right)}{m - \mathcal{O}\left(\alpha\right)} \to 0. \tag{35}$$

This implies that $\alpha = \mathcal{O}\left(\beta^{1/2}\right)$.

This conclusion together with (34) and triangular inequality implies

$$\begin{aligned} \left(1 - \mathcal{O}\left(\beta^{1/2}\right)\right)\beta &= \left\| \alpha^2 \frac{\mu W''}{2} + L' + \mathcal{O}\left(\beta^2\right) \right\|_F, \\ \beta\left(1 - \mathcal{O}\left(\beta^{1/2}\right)\right) - \|L'\|_F - \mathcal{O}\left(\beta^2\right) &\leq \alpha^2 \left\| \frac{\mu W''}{2} \right\|_F \leq \beta\left(1 - \mathcal{O}\left(\beta^{1/2}\right)\right) + \|L'\|_F + \mathcal{O}\left(\beta^2\right), \\ \beta\left(1 - \mathcal{O}\left(\beta^{1/2}\right)\right) - \mathcal{O}\left(\beta^{3/2}\right) - \mathcal{O}\left(\beta^2\right) &\leq \alpha^2 \left\| \frac{\mu W''}{2} \right\|_F \leq \beta\left(1 - \mathcal{O}\left(\beta^{1/2}\right)\right) + \mathcal{O}\left(\beta^{3/2}\right) + \mathcal{O}\left(\beta^2\right), \\ \therefore \quad \alpha^2 \left\| \frac{\mu W''}{2} \right\|_F &= \beta\left(1 - \mathcal{O}\left(\beta^{1/2}\right)\right). \end{aligned}$$

We now obtain the square of the curvature $h_{PD}(\mu)$ as

$$h_{PD}(\mu) := \alpha^2 \left\| \frac{\mu W''}{2} \right\|_F / \mu^2 = \alpha^2 \left\| \frac{W''}{2\mu} \right\|_F = \frac{\beta}{\mu^2}\left(1 - \mathcal{O}\left(\beta^{1/2}\right)\right). \tag{36}$$

which is same as the one in the idealized algorithm. Applying the same logic of the proof in Theorem 3.1, we obtain the conclusion. □

In the next lemma, we study the concrete formula for the curvature in case of HKM dual direction.

Lemma 3.5. *On the central path, it holds that*

$$\frac{W_P''}{2\mu} = (X^{-1/2}\dot{X}X^{-1/2}) \circ (S^{-1/2}\dot{S}S^{-1/2}), \tag{37}$$

in HKM dual direction $P = X^{-1/2}$.

Proof. Combining the fact that $\widetilde{X} = PXP^T = I, \widetilde{S} = P^{-T}SP^{-1} = \mu I$ and Lemma 3.3, we have

$$\begin{aligned} W_P' &= V\widetilde{S}\widetilde{X}^{1/2} + \widetilde{X}^{1/2}\widetilde{S}V + \widetilde{X}^{1/2}\dot{\widetilde{S}}\widetilde{X}^{1/2}, \\ W_P'' &= V'\widetilde{S}\widetilde{X}^{1/2} + \widetilde{X}^{1/2}\widetilde{S}V' + 2V\dot{\widetilde{S}}\widetilde{X}^{1/2} + 2\widetilde{X}^{1/2}\dot{\widetilde{S}}V + 2V\widetilde{S}V \\ &= 2\mu V' + 2\mu V^2 + 2V\dot{\widetilde{S}} + 2\dot{\widetilde{S}}V, \end{aligned}$$

where V and V' satisfy

$$2V = \dot{\widetilde{X}} \quad \text{and} \quad V' + V^2 = 0.$$

It follows from the identities $\dot{\widetilde{X}} = X^{-1/2}\dot{X}X^{-1/2}$ and $\dot{\widetilde{S}} = \mu S^{-1/2}\dot{S}S^{-1/2}$ that

$$W_P'' = \dot{\widetilde{X}}\dot{\widetilde{S}} + \dot{\widetilde{S}}\dot{\widetilde{X}} = 2\dot{\widetilde{X}} \circ \dot{\widetilde{S}} = 2\mu\left[\left(X^{-1/2}\dot{X}X^{-1/2}\right) \circ \left(S^{-1/2}\dot{S}S^{-1/2}\right)\right].$$

□

Lemma 3.6 states that the equation (37) in Lemma 3.5 holds any scaling matrix P.

Lemma 3.6. *On the central path, $\|W_P''\|_F$ is scaling invariant.*

Proof. Let W_P' be the first-order derivative along the vector $(\Delta X, \Delta y, \Delta S)$ obtained from the system (11) and $W_P'', W_P^{(3)}, W_P^{(4)}$ be second, third, and fourth order derivative along the same vector, respectively. Then we have

$$\begin{aligned}(W_P W_P)' &= W_P' W_P + W_P W_P',\\ (W_P W_P)'' &= W_P'' W_P + 2W_P'^2 + W_P W_P'',\\ (W_P W_P)^{(3)} &= W_P^{(3)} W_P + W_P'' W_P' + 5W_P' W_P'' + W_P W_P^{(3)},\\ (W_P W_P)^{(4)} &= W_P^{(4)} W_P + 2W_P^{(3)} W_P' + 6W_P''^2 + 6W_P' W^{(3)} + W_P W_P^{(4)}.\end{aligned}$$

Since $W_P = \nu I$ and $W_P' = -\nu I$, it follows that

$$(W_P W_P)^{(4)} = 2\nu W_P^{(4)} - 8\nu W_P^{(3)} W_P' + 6W_P''^2. \tag{38}$$

Taking a trace on the equation (38), we have

$$\mathrm{Tr}((W_P W_P)^{(4)}) = 2\nu\, \mathrm{Tr}(W_P^{(4)}) - 8\nu\, \mathrm{Tr}(W_P^{(3)}) + 6\,\mathrm{Tr}(W_P''^2).$$

Since $\mathrm{Tr}(W_P)$ is a scaling invariant, so are $\mathrm{Tr}(W_P^{(3)})$, $\mathrm{Tr}(W_P^{(4)})$ and $\mathrm{Tr}((W_P W_P)^{(4)})$. Therefore, $\|W_P''\|_F = \mathrm{Tr}(W_P''^2)^{1/2}$ is scaling invariant. ☐

Combining Theorem 3.4, Lemma 3.5 and 3.6, we obtain the main theorem in this section:

Theorem 3.7. *Let $\beta \in (0, 1/2]$. For given $w^0 \in \mathcal{N}(\beta)$ and $0 < \nu_f < \mu(w^0)$ denote by $\#_{PD}(\mu(w^0), \nu_f, \beta)$ the number of iterations of the MTY-PC algorithm with $\beta \in (0, 1/2]$ needed to reduce the duality gap from $\nu_i := \mu(w^0)$ to ν_f. Then, using the curvature integral:*

$$I_{PD}(\nu_f, \nu_i) = \int_{\nu_f}^{\nu_i} \left\| \left(X^{-1/2}\dot{X}X^{-1/2}\right) \circ \left(S^{-1/2}\dot{S}S^{-1/2}\right) \right\|_F^{1/2} \mathrm{d}\nu, \tag{39}$$

we have

$$\lim_{\beta \downarrow 0} \frac{I_{PD}(\nu_f, \nu_i)/\sqrt{\beta}}{\#_{PD}(\nu_i, \nu_f, \beta)} = 1.$$

Remark 3.1. As is discussed in Lemma 2.1 of [10], it also holds in SDP that the curvature of the central path is zero in all the parts of the trajectory if it is zero at some point, i.e.,

$$h_{PD}(\nu) = \left\| \left(X^{-1/2}\dot{X}X^{-1/2}\right) \circ \left(S^{-1/2}\dot{S}S^{-1/2}\right) \right\|_F = 0 \text{ for } \forall \nu > 0 \quad \text{if } h_{PD}(\nu_0) = 0 \text{ at some } \nu_0 > 0.$$

Since the derivation of the concrete formula for the curvature is independent of the discussion of the iteration complexity analysis, we can restrict our discussions in Theorem 3.1, 3.4 and 3.7 in case the curvature is larger than zero. The proof is very similar to Lemma 2.1 in [10], and we will discuss it in section 5.

From here to the end of the section, we lay out several lemmas used in the proof of theorems.

Lemma 3.8. *Let $w \in \mathcal{N}\left(\beta^2\right)$ be a feasible point with the duality gap ν and let w_ν be the point on the central path with the same duality gap. Then, asymptotically, the following relation holds:*

$$\|w - w_\nu\| = \mathcal{O}\left(\beta^2\right).$$

Proof. Let

$$F_w(w') = \begin{bmatrix} W_P(w') - W_P(w) \\ \mathcal{A}X' - b \\ \mathcal{A}^* y' + S' - C \end{bmatrix}.$$

Then, letting $R = W_P(w)/\nu$, we have

$$F_w(w) = 0 \text{ and } F_w(w_\nu) = \begin{bmatrix} \nu(I-R) \\ 0 \\ 0 \end{bmatrix}.$$

Hence, from the continuity of F_w, it holds on the neighborhood of the central path that with constants C_1, C_2,

$$\|w - w_\nu\| = C_1 \left\| F_w^{-1}(0) - F_w^{-1}([\nu(I-R), 0, 0]^T) \right\| = C_2 \|\nu(I-R)\|_F = \mathcal{O}\left(\beta^2\right).$$

□

The following lemma is the sensitivity analysis formula, whose proof and discussion are found, for example, in Demmel [2].

Lemma 3.9 (Sensitivity Analysis Formula)**.** *Assume A is a nonsingular matrix which satisfies $\|\Delta A\| \left\|A^{-1}\right\| < 1$ and the following system of equations hold:*

$$\begin{cases} Ax = b \\ (A + \Delta A)(x + \Delta x) = b + \Delta b \end{cases}.$$

Then it holds that

$$\|\Delta x\| \leq \frac{\left\|A^{-1}\right\|}{1 - \|A^{-1}\| \, \|\Delta A\|} \left(\|\Delta b\| + \|\Delta A\| \, \|x\|\right).$$

Note that this formula allows any norms to hold the statement.

Lemma 3.10. *$(X, y, S) \in \mathcal{N}\left(\beta^2\right)$ is given. Let $(\Delta X, \Delta y, \Delta S)$ be MZ family of directions, namely, the solution to the system* (12) *with $\sigma = 0$ and let $(\Delta X_\nu, \Delta y_\nu, \Delta S_\nu)$ be MZ family of directions on the central path with the same duality gap. (Note that on the central path, MZ family of directions and MT family of directions are identical.) Then we have*

$$\|(\Delta X, \Delta y, \Delta S) - (\Delta X_\nu, \Delta y_\nu, \Delta S_\nu)\|_F = \mathcal{O}(\beta^2).$$

Proof. Let

$$\widetilde{\mathcal{A}} := \begin{pmatrix} \mathcal{A} & 0 & 0 \\ 0 & I & \mathcal{A}^* \\ E_\nu & F_\nu & 0 \end{pmatrix}, \quad \widetilde{x} := \begin{bmatrix} \mathrm{vec}(\Delta X_\nu) \\ \mathrm{vec}(\Delta S_\nu) \\ \Delta y_\nu \end{bmatrix}, \quad \widetilde{b} := \begin{bmatrix} 0 \\ 0 \\ \mathrm{vec}(\nu I) \end{bmatrix},$$

$$\widetilde{\Delta\mathcal{A}} := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ E - E_\nu & F - F_\nu & 0 \end{pmatrix}, \quad \widetilde{\Delta x} := \begin{bmatrix} \mathrm{vec}(\Delta X - \Delta X_\nu) \\ \mathrm{vec}(\Delta S - \Delta S_\nu) \\ \Delta y - \Delta y_\nu \end{bmatrix}, \quad \widetilde{\Delta b} := \begin{bmatrix} 0 \\ 0 \\ \mathrm{vec}(\mathcal{H}_p(XS) - \nu I) \end{bmatrix},$$

where

$$E_\nu := \frac{1}{2}\left(P_\nu^{-T} S_\nu \otimes P_\nu + P_\nu \otimes P_\nu^{-T} S_\nu\right), \quad F_\nu := \frac{1}{2}\left(P_\nu X_\nu \otimes P_\nu^{-T} + P_\nu^{-T} \otimes P_\nu X_\nu\right),$$
$$E := \frac{1}{2}\left(P^{-T} S \otimes P + P \otimes P^{-T} S\right), \quad F := \frac{1}{2}\left(P X \otimes P^{-T} + P^{-T} \otimes P X\right).$$

Note that by conventions, we allows above formulations of operators and matrices. Then, by Lemma 3.9, we have

$$\left\|\widetilde{\Delta x}\right\| \leq \frac{\left\|\widetilde{\mathcal{A}}^{-1}\right\|}{1 - \left\|\widetilde{\mathcal{A}}^{-1}\right\| \left\|\widetilde{\Delta\mathcal{A}}\right\|} \left(\left\|\widetilde{\Delta b}\right\| + \left\|\widetilde{\Delta\mathcal{A}}\right\| \|\widetilde{x}\|\right). \tag{40}$$

We first show that $\left\|\widetilde{\Delta\mathcal{A}}\right\|_F = \mathcal{O}\left(\beta^2\right)$. We have

$$\begin{aligned} E - E_\nu &= \frac{1}{2}\left[\left(P^{-T}S \otimes P + P \otimes P^{-T}S\right) - \left(P_\nu^{-T}S_\nu \otimes P_\nu + P_\nu \otimes P_\nu^{-T}S_\nu\right)\right] \\ &= \frac{1}{2}\left[\left(P^{-T}S \otimes P - P_\nu^{-T}S_\nu \otimes P_\nu\right) + \left(P \otimes P^{-T}S - P_\nu \otimes P_\nu^{-T}S_\nu\right)\right] \\ &= \frac{1}{2}\Big[\,\big((P^{-T} - P_\nu^{-T})S \otimes P + P_\nu^{-T}(S - S_\nu) \otimes P + (P_\nu^{-T}S_\nu \otimes (P - P_\nu)\big) + \\ &\qquad \big((P - P_\nu) \otimes P^{-T}S + P_\nu \otimes (P^{-T} - P_\nu^{-T})S + P_\nu \otimes P_\nu^{-T}(S - S_\nu)\big)\,\Big]. \end{aligned}$$

Since $\|S\|_F, \|S_\nu\|_F, \|P\|_F, \|P_\nu\|_F$ are bounded on $\mathcal{F}$, and there exists a constant C such that $\|P - P_\nu\|_F = C\|w - w_c\|_F$ and $\|w - w_c\|_F = \mathcal{O}\left(\beta^2\right)$ from Lemma 3.8, it follows that

$$\|E - E_\nu\|_F = \mathcal{O}\left(\beta^2\right).$$

Similarly, we have $\|F - F_\nu\|_F = \mathcal{O}\left(\beta^2\right)$. Thus we obtain

$$\left\|\widetilde{\Delta\mathcal{A}}\right\|_F = \mathcal{O}\left(\beta^2\right).$$

Next, we show $\left\|\widetilde{\Delta b}\right\| = \mathcal{O}\left(\beta^2\right)$. Similarly, it is easy to see that

$$\begin{aligned} \|\mathcal{H}_p(XS) - \nu I\|_F &= \frac{1}{2}\left\|(PXSP^{-1} - \nu I) + (P^{-T}SXP^T - \nu I)\right\|_F \\ &\leq \frac{1}{2}\Bigg[\left\|PX^{1/2}(X^{1/2}SX^{1/2} - \nu I)X^{-1/2}P^{-1}\right\|_F \\ &\qquad + \left\|P^{-T}X^{-1/2}(X^{1/2}SX^{1/2} - \nu I)X^{1/2}P^T\right\|_F\Bigg] \\ &\leq \frac{1}{2}\Bigg[\|P\|_F\left\|X^{1/2}\right\|_F\left\|X^{1/2}SX^{1/2} - \nu I\right\|_F\left\|X^{-1/2}\right\|_F\|P^{-1}\|_F \\ &\qquad + \|P^{-T}\|_F\left\|X^{-1/2}\right\|_F\left\|X^{1/2}SX^{1/2} - \nu I\right\|_F\left\|X^{1/2}\right\|_F\|P^T\|_F\Bigg] = \mathcal{O}(\beta^2). \end{aligned}$$

The last equality comes from the fact that $\left\|X^{1/2}SX^{1/2} - \nu I\right\|_F = \mathcal{O}\left(\beta^2\right)$ and $\|P\|_F$ and $\left\|X^{-1/2}\right\|_F$ are bounded on $\mathcal{F}$.

Since $\left\|\widetilde{\mathcal{A}}^{-1}\right\|$, $\|\widetilde{x}\|$ are bounded, it follows from above discussion that $\left\|\widetilde{\Delta x}\right\| = \mathcal{O}\left(\beta^2\right)$, which implies that

$$\|(\Delta X, \Delta y, \Delta S) - (\Delta X_\nu, \Delta y_\nu, \Delta S_\nu)\|_F = \mathcal{O}(\beta^2).$$

□

4 Numerical Experiments in SDP

In this section, we investigate the relationship between the curvature integral and iteration complexity in MTY-PC algorithms by way of numerical experiments. By construction of the proof of Theorem 3.1 and 3.4, and from the relation described herein:

$$I_{PD}(\nu_f, \nu_i) \approx \sqrt{\beta} \cdot \#_{PD}(\nu_i, \nu_f, \beta), \tag{41}$$

solving the SDP problem with a sufficiently small opening of the neighborhood β is equivalent to computing the approximate value of the curvature integral for the given problem. To see the relation (41) hold for the practically large opening β, we conducted numerical experiments from SDPLIB [1], a collection of SDP problems such as truss topology design problems, problems from control and system theory, max cut problems, etc. We adopt the opening β from 0.125, 0.25, 0.5 to 1.0 and then in the left side of the page, plot data with the number of iteration multiplied by $\sqrt{\beta}$ in x-axis and with the reduced duality gap in y-axis and in the right side plot the same data with the number of iteration multiplied by $\sqrt{\beta}$ in x-axis and with the reduced duality gap in y-axis. If the graph with $\beta = 1$ or $\beta = 0.5$ overlaps with the one with $\beta = 0.125$, our results are given some validity for the practical method to analyze iteration complexities of SDP problems.

We also present the results from LP problems in Netlib for comparison. All the numerical experiments are done in the same manner as in SDP problems.

We will describe the procedure for numerical experiments. We first reformulate the original problem to the augmented SDP problem proposed by Monteiro and Adler [8] so that the initial point starts on the central path with a certain duality gap. Then we implement the algorithm described in Section 2.1 faithfully; the computation of the step size is done by the bisection method and we adopt the Nesterov-Todd direction [16].

Let vec be the vectorization operator. Let $\mathcal{A}^* := [\mathrm{vec}\, A_1 \dots \mathrm{vec}\, A_m]$, $x = \mathrm{vec}\, X$, $s = \mathrm{vec}\, S$, $c = \mathrm{vec}\, C$ and $e = \mathrm{vec}\, I$. Then primal and dual problems of SDP (2) and (3) are rewritten as

$$(P) \qquad \min_x c^T x \quad \text{s.t.} \quad \begin{array}{l} \mathcal{A}x = b, \\ x \in \Omega, \end{array}$$

and

$$(D) \qquad \min_y b^T y \quad \text{s.t.} \quad \begin{array}{l} \mathcal{A}^* y + s = c, \\ s \in \Omega, \end{array}$$

where $x \in \Omega$ denotes that the matrix form of x is positive semidefinite. Let $\widetilde{n} = n + 2$ and $\widetilde{m} = m + 1$. Let $L = L(\mathcal{A}, b, c)$ denote the bit size of the problem (P). Let $\alpha = 2^{4L}$ and $\lambda = 2^{2L}$. Let K_b and K_c be constants defined as follows:

$$K_b := \alpha\lambda(n+1) - \lambda c^T e, \quad K_c := \alpha\lambda. \tag{42}$$

The augmented problem is formulated as follows:

$$(\widetilde{P}) \qquad \min_{x, x_{n+1}, x_{n+2}} c^T x + K_c x_{n+2} \quad \text{s.t.} \quad \begin{array}{l} \mathcal{A}x + (b - \lambda \mathcal{A}e) x_{n+2} = b, \\ (\alpha e - c)^T x + \alpha x_{n+1} = K_b, \\ x \in \Omega, \ x_{n+1} \geq 0, \ x_{n+2} \geq 0, \end{array}$$

$$(\widetilde{D}) \qquad \min_{y, y_{m+1}, s, s_{n+1}, s_{n+2}} b^T y + K_b y_{m+1} \quad \text{s.t.} \quad \begin{array}{l} \mathcal{A}^* y + (\alpha e - c) y_{m+1} + s = c, \\ \alpha y_{m+1} + s_{n+1} = 0, \\ (b - \lambda \mathcal{A}e)^T y + s_{n+2} = K_c, \\ s \in \Omega, \ s_{n+1} \geq 0, \ s_{n+2} \geq 0. \end{array}$$

Then the point $w^0 = (x^0, s^0, y^0)$, defined by

$$x^0 := \begin{pmatrix} \lambda e \\ \lambda \\ 1 \end{pmatrix}, \quad s^0 := \begin{pmatrix} \alpha e \\ \alpha \\ \alpha\lambda \end{pmatrix}, \quad y^0 := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix},$$

is on the central path with duality gap $\alpha\lambda$.

The magnitude of penalty values K_c and K_b in $(\widetilde{P})$ and $(\widetilde{D})$ affects significantly the result of actual computations; if it is too small, the penalty method does not work, and if it is too large, it results in a numerical instability. Based on our heuristics, we choose L between 5 and 12 in our experiments.

We present the following 3 instances of LP problems:

- D2Q06C (Dimension of A: 2172×5167; Optimal value: 1.2278423615×10^5) (Fig. 1)
- PILOT87 (Dimension of A: 2031×4883; Optimal value: 3.0171072827×10^2) (Fig. 2)
- DFL001 (Dimension of A: 6072×12230; Optimal value: 1.12664×10^7) (Fig. 3).

We conduct the following 6 instances of SDP problems:

- CONTROL10 (Dimension: $m = 1326$, $n = 150$; Optimal value: 3.8533×10^1) (Fig. 4)
- CONTROL11 (Dimension: $m = 1596$, $n = 165$; Optimal value: 3.1959×10^1) (Fig. 5)
- EQUALG51 (Dimension: $m = 1001$, $n = 1001$; Optimal value: 4.005601×10^3) (Fig. 6)
- MAXG51 (Dimension: $m = 1000$, $n = 1000$; Optimal value: 4.003809×10^3) (Fig. 7)
- MCP500-4 (Dimension: $m = 500$, $n = 500$; Optimal value: 3.566738×10^3) (Fig. 8)
- TRUSS8 (Dimension: $m = 496$: $n = 628$; Optimal value: -1.331146×10^2) (Fig. 9)

where m denotes the number of constraint matrices and n denotes the rank of variable matrices.

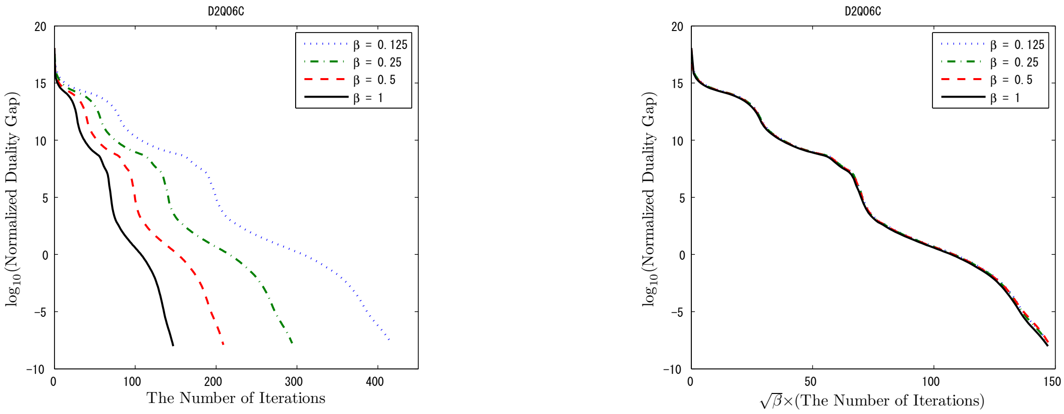

Figure 1: D2Q06C (Dimension of A: 2172×5167; Optimal value: 1.2278423615×10^5)

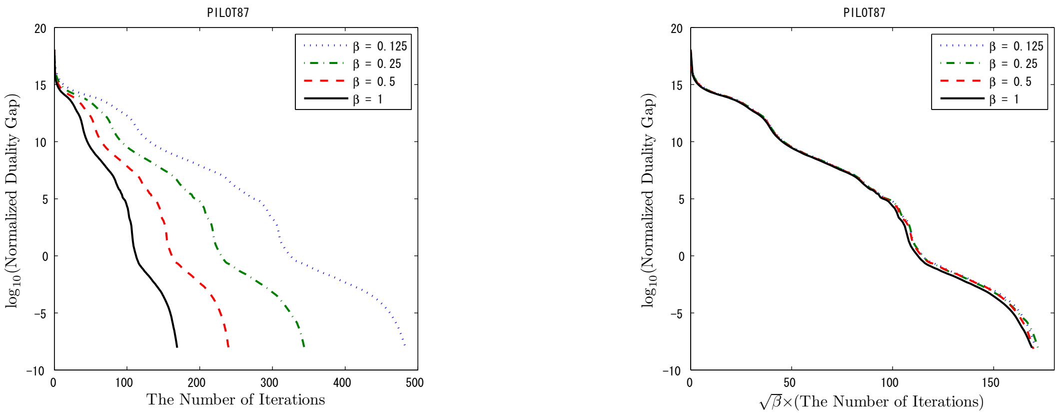

Figure 2: PILOT87 (Dimension of A: 2031×4883; Optimal value: 3.0171072827×10^2)

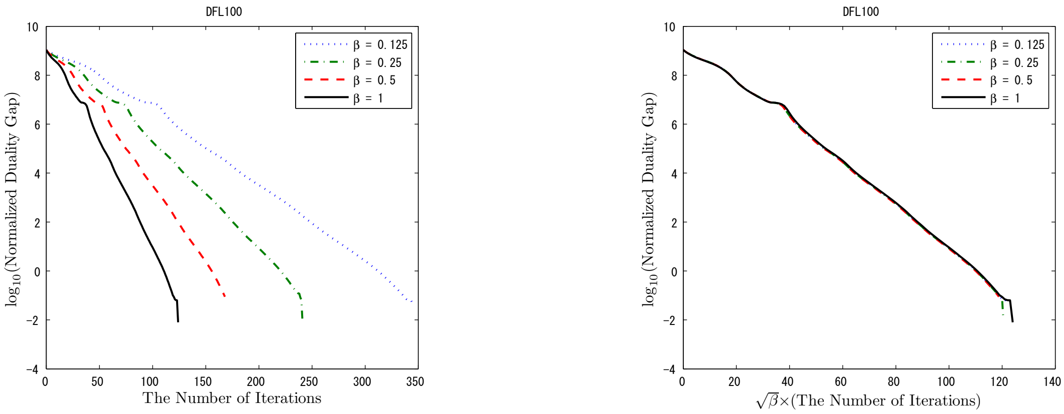

Figure 3: DFL001 (Dimension of A: 6072×12230; Optimal value: 1.12664×10^7)

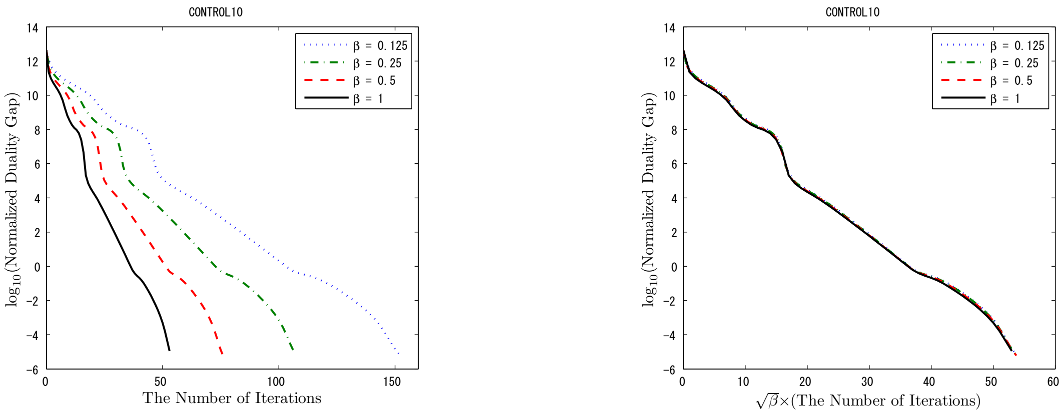

Figure 4: CONTROL10 (Dimension: $m = 1326$, $n = 150$; Optimal value: 3.8533×10^1)

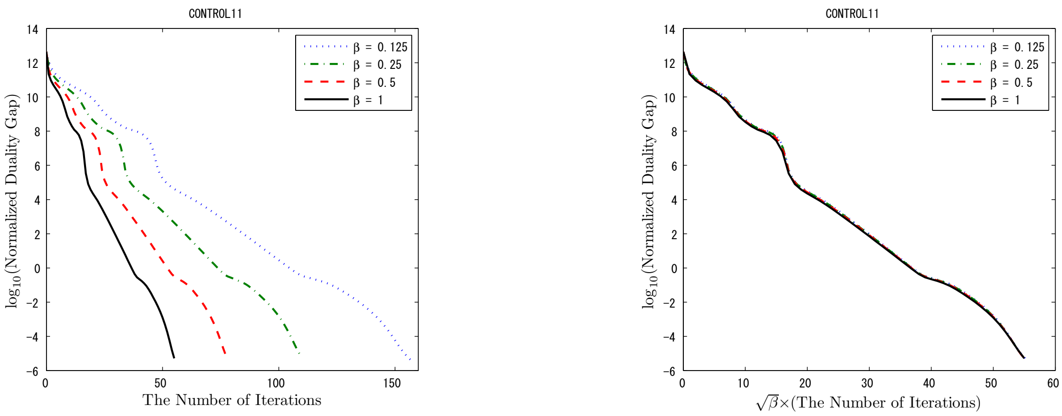

Figure 5: CONTROL11 (Dimension: $m = 1596$, $n = 165$; Optimal value: 3.1959×10^1)

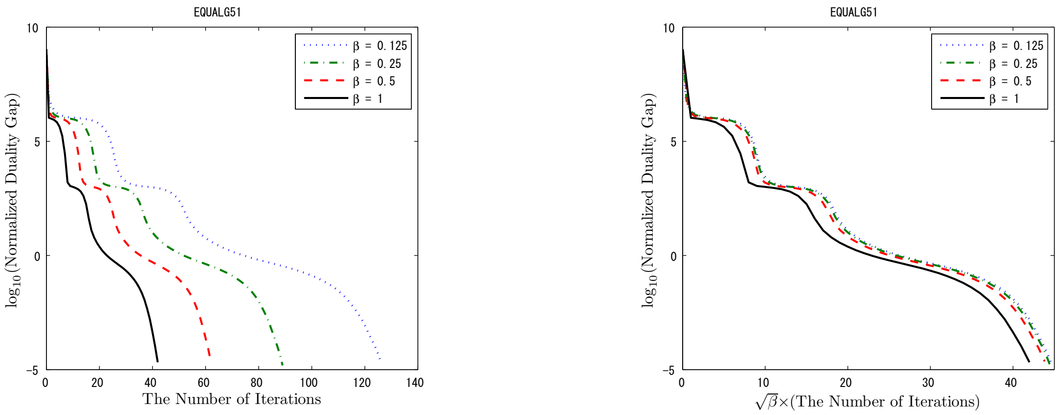

Figure 6: EQUALG51 (Dimension: $m = 1001$, $n = 1001$; Optimal value: 4.005601×10^3)

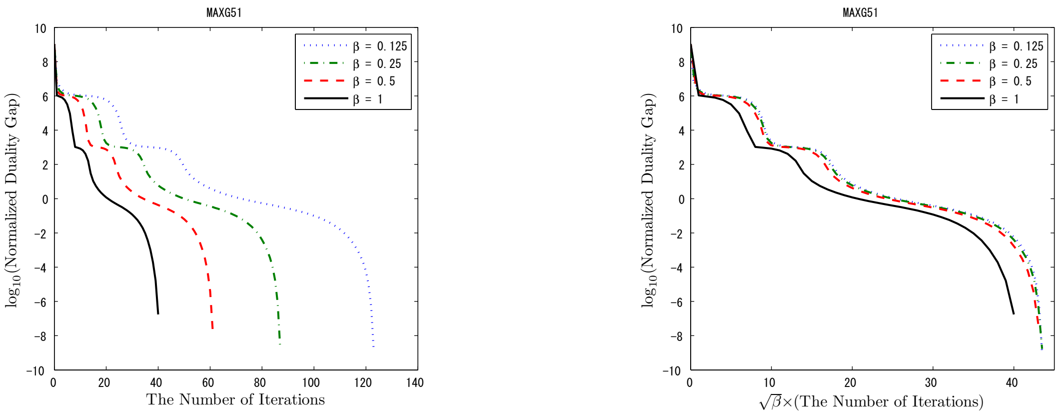

Figure 7: MAXG51 (Dimension: $m = 1000$, $n = 1000$; Optimal value: 4.003809×10^3)

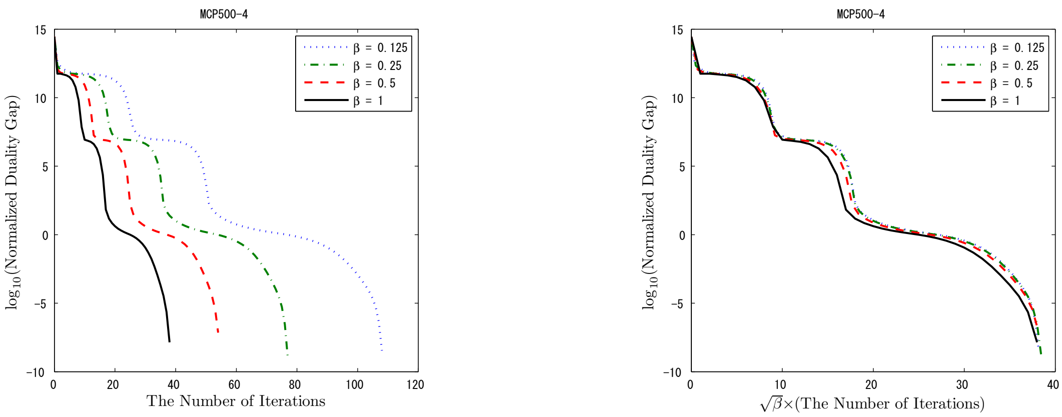

Figure 8: MCP500-4 (Dimension: $m = 500$, $n = 500$; Optimal value: 3.566738×10^3)

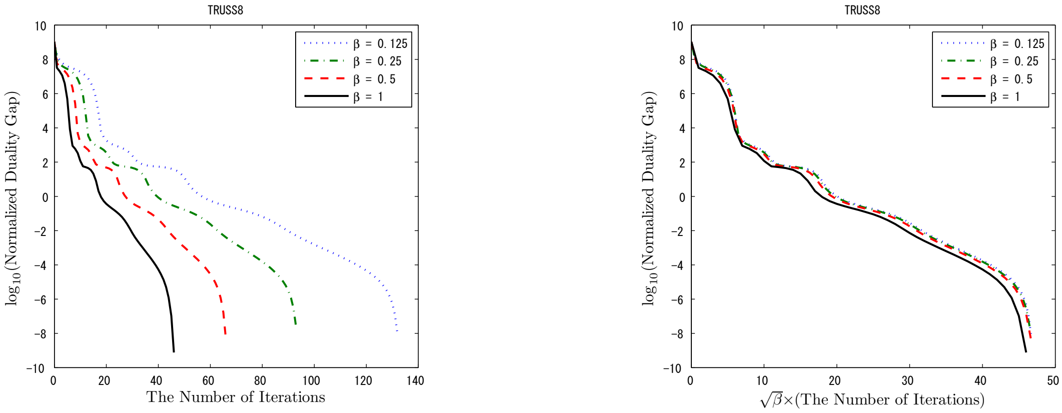

Figure 9: TRUSS8 (Dimension: $m = 496$: $n = 628$; Optimal value: -1.331146×10^2)

Let's see the left hand side of the figures. In the intervals where the large reduction of the duality gap is observed, the central path is relatively straight, whereas in the intervals where little or no reduction is observed, the central path is strongly curved.

In the right hand side of the figures, we multiplies the number of iterations by $\sqrt{\beta}$, and plot the same data. It is observed that in all figures, graphs with $\beta = 0.125, 0.25, 0.5$ overlap very closely. In some figures, the graphs with $\beta = 1$ do not overlap with others. In these figures, it is seen that the separated parts of the graphs are all strongly curved, with very little reduction of the duality gaps being observed. It is considered that MTY-PC algorithms with large opening can go through the strongly curved parts less number of iterations than ones with smaller opening, even if they are plotted with a scale measure. In fact, in CONTROL10 and CONTROL11, whose duality gaps decrease at a constant rate, all graphs overlap precisely.

It is also observed that SDP problems demonstrate the very similar characteristics to LP problems. Yet, compared to LP problems, SDP problems are required roughly half the number of iterations to reduce the duality gap. For example, in case of DFL001 in LP problem, the ratio of the reduced duality gap to the number of iteration is almost constant, and it takes more than a hundred iterations to reduce ten digits of duality gap.

Although there are some degree of irregularity in the strongly curved parts, the approximation (41) holds for practically large opening of the neighborhood, which indicates that our results serve as a useful analytical tool for various SDP problems.

5 Notes on the Extension to Symmetric Cone Programs

In this section, we briefly discuss the symmetric cone programs and then prove symmetric cone programs' case for Theorem 3.7 in SDP.

5.1 Symmetric Cone Programs

In this subsection, we first introduce Euclidean Jordan Algebra, the language to command the structure of symmetric cone programs, and then briefly explain MTY-PC algorithm in symmetric cone programs. The notions and algorithms of symmetric cone programs are specified by word-for-word substitution from SDP [4, 12, 14].

Let V be an n-dimensional real vector space with a bilinear mapping $(x, y) \mapsto x \circ y$ from $V \times V$ into V. *Jordan Algebra* is a vector space V with a multiplication $\circ$ satisfying

1. $x \circ y = y \circ x$
2. $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$

where $x^2 = x \circ x$. Note that in general it is not associative, i.e., $x \circ (y \circ z) \neq (x \circ y) \circ z$, but it is power associative, i.e., $x^{k+l} = x^k \circ x^l$ for nonnegative integers k and l. In this paper, we assume that there exists a unique identity $e \in V$ such that $x \circ e = e \circ x = x$ for all $x \in V$.

If there exists a positive definite bilinear form $\langle \cdot, \cdot \rangle$ on V which is associative, namely,

$$\langle x \circ y, z\rangle = \langle x, y \circ z\rangle \text{ for } \forall x, y, z \in V, \tag{43}$$

then, Jordan algebra is *Euclidean*. In what follows, we assume that V is Euclidean Jordan algebra.

$c \in V$ is called *idempotent* if $c^2 = c$. Idempotents c and c' are orthogonal if $c \circ c' = 0$. An idempotent is *primitive* if it is non-zero and cannot be expressed by the sum of other two non-zero idempotents. The number of primitive idempotents is less than or equal to the dimension of V, so we define the *rank* of V by the maximum possible number of primitive orthogonal idempotents in V.

$c_1, \ldots, c_k$ is a complete system of orthogonal primitive idempotents or *Jordan frame* if each c_i is an idempotent and if

$$c_i \circ c_j = 0,\ i \neq j, \quad \sum\nolimits_{i=1}^{k} = e. \tag{44}$$

Theorem 5.1 (Spectral Decomposition). *Suppose V has rank r. Then for $x \in V$, there exists Jordan frame $\{c_1, \ldots, c_r\}$ and spectral values $\{\lambda_1, \ldots, \lambda_r\}$ such that*

$$x = \sum\nolimits_{i=1}^{r} \lambda_i c_i. \tag{45}$$

The spectral values λ_i are uniquely determined by x. Furthermore,

$$\det x = \prod\nolimits_{i=1}^{r} \lambda_i, \qquad \operatorname{trace} x = \sum\nolimits_{i=1}^{r} \lambda_i. \tag{46}$$

If $x = \sum_i \lambda_i c_i$ is the spectral decomposition, then

$$\begin{aligned} x^2 &= x \circ x = \sum\nolimits_{i=1}^{r} \lambda^2 c_i, \\ \|x\|^2 &= \sum\nolimits_{i=1}^{r} \lambda^2. \end{aligned}$$

Furthermore, we can define x^{-1} and $x^{1/2}$ as follows:

$$\begin{aligned} x^{-1} &:= \sum\nolimits_{i=1}^{r} \lambda^{-1} c_i, \text{ if } \lambda_i \neq 0 \text{ for } \forall i, \\ x^{1/2} &:= \sum\nolimits_{i=1}^{r} \lambda^{1/2} c_i, \text{ if } \lambda_i > 0 \text{ for } \forall i. \end{aligned}$$

From this, we have $x^{-1} \circ x = x \circ x^{-1} = e$ and $x^{1/2} \circ x^{1/2} = x$.

Using spectral decomposition, *symmetric cones* are defined as the set of elements in V whose λ_i's are all nonnegative. Or equivalently, they are defined as the cones of squares of Euclidean Jordan Algebra V. It is known that symmetric cones are self-dual, i.e.,

$$\Omega = \Omega^* := \{y \in V \mid \langle y, x\rangle \geq 0, \quad \forall x \in V\},$$

and their automorphism group, defined by

$$G(\Omega) := \{g \in GL(V) | g\Omega = \Omega\}, \tag{47}$$

with the general linear group $GL(V)$, acts transitively on their interiors.

We introduce a linear form L_x and two types of quadratic forms Q_x and $Q_{x,y}$. Let L_x be a linear mapping of V defined by $L_x y := x \circ y$ for $\forall x, y \in V$ and let $Q_x := 2L_x^2 - L_{x^2}$ and $Q_{x,y} := L_x L_y + L_y L_x - L_{x\circ y}$. Then the following proposition holds. (For proofs, please see [3])

Proposition 5.2. *i)* $Q_x^{-1} = Q_{x^{-1}}$ *if* x *is invertible,*

ii) $Q_x x^{-1} = x$ *if* x *is invertible,*

iii) $Q_x^{1/2} = Q_{x^{1/2}}$ *if the spectral values of* x *are all positive,*

iv) $(Q_x y)^{-1} = Q_{x^{-1}} y^{-1}$ *if* x *and* y *are invertible,*

v) $Q_{Q_x y} = Q_x Q_y Q_x$*,*

vi) $Q_{x^2,y} = Q_x L_{Q_x^{-1} y} Q_x$ *if* x *is invertible,*

vii) $Q_x L_{x^{-1}} = L_x$ *if* x *is invertible.*

We list some important formulas for directional derivatives. (D_u denotes the directional derivative operator along the vector u.)

Lemma 5.3. *i)* $D_u(x^{-1}) = -Q_x^{-1} u$

ii) $D_u(Q_x) = 2Q_{x,u}$.

Proof. We only make a proof of ii). For the first one, please see [3]. For $\forall y \in V$, it holds that

$$Q_x y = 2x \circ (x \circ y) - x^2 \circ y.$$

By differentiating above equation, we have

$$\begin{aligned} D_u(Q_x y) &= 2u \circ (x \circ y) + 2x \circ (u \circ y) - (2x \circ u) \circ y \\ &= 2\{u \circ (x \circ y) + x \circ (u \circ y) - (x \circ u) \circ y\} \\ &= 2\{L_u L_x + L_x L_u - L_{x\circ u}\} y = 2Q_{x,u} y. \end{aligned}$$

□

We also state the lemma used later.

Lemma 5.4. *Let* $p \in V$ *is an invertible and* $x, y \in V$*. Then*

$$x \circ y = 0 \Longleftrightarrow Q_p x \circ Q_{p^{-1}} y = 0.$$

Proof. Since $Q_p L_{p^{-1}} = L_p$ and Q_p and $L_{p^{-1}}$ commute, we have

$$p^{-1} \circ Q_p x = L_{p^{-1}} Q_p x = p \circ x.$$

Taking the directional derivative of both sides of the above equation w.r.t. p along u, we have

$$(-Q_{p^{-1}} u) \circ Q_p x + 2p^{-1} Q_{p,u} x = u \circ x.$$

Since L_x and L_y commute and $x \circ y = 0$, we obtain the desired result. □

We are now ready to discuss symmetric cone programs. Let V be an n-dimensional Euclidean Jordan algebra with rank r and Ω its associated symmetric cone. We consider a primal-dual pair of the symmetric cone program:

$$(P) \qquad \min_x \langle c, x\rangle \qquad \text{s.t.} \quad \begin{array}{l} \langle a_i, x\rangle = b_i, \quad i = 1, \ldots, m, \\ x \in \Omega, \end{array} \tag{48}$$

and

$$(D) \qquad \min_{y,s} b^T y \qquad \text{s.t.} \quad \begin{array}{l} \sum_{i=1}^m y_i a_i + s = c, \\ s \in \Omega, \end{array} \tag{49}$$

where $c \in V,\ a_i \in V,\ i = 1, \ldots, m$ and $b = (b_1, \ldots, b_m)^T \in \mathbb{R}^m$ are the given data, and $x \in V$ and $(y, s) \in \mathbb{R}^m \times V$ are the primal and dual variables, respectively.

The feasible regions of (48) and (49) are defined as

$$\mathcal{P} := \{x \in V \mid \langle a_i, x\rangle = b_i,\ i = 1, \ldots, m, \quad x \in \Omega\}, \tag{50}$$

$$\mathcal{D} := \left\{(y, s) \in \mathbb{R}^m \times V \mid \sum_{i=1}^m y_i a_i + s = c, \quad s \in \Omega\right\}, \tag{51}$$

and its corresponding strictly feasible regions are defined as

$$\mathcal{P}^+ := \{x \in \mathcal{P} \mid x \in \mathrm{int}(\Omega)\}, \tag{52}$$

$$\mathcal{D}^+ := \{(y, s) \in \mathcal{D} \mid s \in \mathrm{int}(\Omega)\}, \tag{53}$$

where $\mathrm{int}(\Omega)$ denotes the relative interior of Ω. We assume that $\mathcal{P}^+ \neq \emptyset$ and $\mathcal{D}^+ \neq \emptyset$.

Using the normalized duality gap $\mu(x, s) := \langle s, x\rangle / r$, the distance measure is defined as

$$d(x, s) := \|Q_{x^{1/2}} s - \mu(x, s) e\| = \left[\sum_{i=1}^r (\lambda_i(x \circ s) - \mu)^2\right]^{1/2}, \tag{54}$$

and, for a given constant $0 < \beta < 1$, we then define the neighborhood:

$$\mathcal{N}(\beta) := \{(x, y, s) \in \mathcal{P}^+ \times \mathcal{D}^+ \mid d(x, s) \leq \beta \mu(x, s)\}.$$

Using the notations:

$$(\mathcal{A}x)_i := \langle a_i, x\rangle, \quad i = 1, \ldots, m,$$

and letting $\mathcal{A}^*$ be an adjoint operator of $\mathcal{A}$, the system:

$$\begin{array}{rcccl} & \mathcal{A}^* y & + & s & = c, \\ \mathcal{A}x & & & & = b, \\ & x \circ s & & & = \nu e, \end{array} \tag{55}$$

has a unique solution in $(x, y, s) \in \mathcal{P}^+ \times \mathcal{D}^+$ for any $\nu > 0$ and the set of such solutions are called the *central path*.

Based on this, we introduce Monteiro-Tsuchiya (MT) family of directions [9] and Monteiro-Zhang (MZ) family of directions [11].

MT family of directions in symmetric cone programs are given by

$$\begin{array}{rcl} & \mathcal{A}^*\Delta y \qquad +\Delta s & = c - \mathcal{A}^* y - s, \\ \mathcal{A}\Delta x & & = b - \mathcal{A}x, \\ 2Q_{(gx)^{1/2},u}(g^{-*}s) + & +Q_{(gx)^{1/2}}(g^{-*}\Delta s) & = \sigma\mu e - Q_{(gx)^{1/2}}(g^{-*}s), \\ 2(gx)^{1/2} \circ u & & = g\Delta x, \end{array} \tag{56}$$

with $g \in G(\Omega)$. On the other hand, MZ family of directions are given by

$$\begin{array}{rcccl} & \mathcal{A}^*\Delta y & & +\Delta s & = c - s - \mathcal{A}^* y, \\ \mathcal{A}\Delta x & & & & = b - \mathcal{A}x, \\ (g^{-*}s) \circ (g\Delta x) & & + & (gx) \circ (g^{-*}\Delta s) & = \sigma\mu e - (gx) \circ (g^{-*}s). \end{array} \tag{57}$$

Note that σ is a centering parameter in both family of directions. The MTY-PC algorithm in symmetric cone programs is identical to one in SDP except that it solves (56) or (57) instead of (11) or (12) for computing the direction.

5.2 Curvature Integrals and Iteration Complexity in Symmetric Cone Programs

Similar to the procedure in SDP case, we first investigate an asymptotic behavior of iteration complexities of an idealized MTY-PC algorithm in symmetric cone programs with MT family of directions in Theorem 5.5, and then extends this result to a generic MTY-PC algorithm with MZ family of directions in Theorem 5.8. We give a concrete formula for curvature integral in Lemma 5.9 and show that it is scaling invariant in Lemma 5.10. Combining Theorem 5.8 and Lemma 5.9 and 5.10, we obtain the final Theorem 5.11, which gives a concrete formula for the curvature integral and the rigorous asymptotic estimate for the iteration complexity of MTY-PC algorithms in symmetric cone programs by way of the aforementioned curvature integral.

For $g \in G(\Omega)$, an automorphism group of Ω, we define

$$\widetilde{x} := gx, \quad \widetilde{s} := g^{-*}s \tag{58}$$

$$\widetilde{\Delta x} := g\Delta x, \quad \widetilde{\Delta s} := g^{-*}\Delta x \tag{59}$$

$$x(\alpha) := x + \alpha\Delta x, \quad s(\alpha) := s + \alpha\Delta s \tag{60}$$

$$\widetilde{x}(\alpha) := g(x + \alpha\Delta x), \quad \widetilde{s}(\alpha) := g^{-*}(s + \alpha\Delta s) \tag{61}$$

$$W_g(w) := Q_{\widetilde{x}^{1/2}}\widetilde{s} \tag{62}$$

$$\mu := \frac{\langle x, s\rangle}{r} = \frac{\langle \widetilde{x}, \widetilde{s}\rangle}{r}, \quad \mu(\alpha) := \frac{\langle x(\alpha), s(\alpha)\rangle}{r} = \frac{\langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle}{r} \tag{63}$$

We drop g from $W_g(\cdot)$ when g is an identity.

We first introduce theorem in regard to an idealized MTY-PC algorithm, in the sense that every iterate starts on the central path and returns exactly on the central path.

Theorem 5.5. *[Idealized MTY-PC algorithm in symmetric cone programs] Let $\beta \in (0, 1/2]$. For given w^0 on the central path and $0 < \nu_f < \mu(w^0)$ denote by $\#_{PD}(\mu(w^0), \nu_f, \beta)$ the number of iterations of the MTY-PC algorithm with $\beta \in (0, 1/2]$ needed to reduce the duality gap from $\nu_i := \mu(w^0)$ to ν_f. Then, using the curvature integral:*

$$I_{PD}(\nu_f, \nu_i) = \int_{\nu_f}^{\nu_i} h_{PD}^{1/2}(\nu)\mathrm{d}\nu, \qquad h_{PD}(\nu) := \left\| \frac{W''(x_\nu, s_\nu)}{2\nu} \right\|, \tag{64}$$

we have

$$\lim_{\beta \downarrow 0} \frac{I_{PD}(\nu_f, \nu_i)/\sqrt{\beta}}{\#_{PD}(\nu_i, \nu_f, \beta)} = 1.$$

Before going to a proof, we make some preliminary definitions and discussions.

Definition 5.1. Let $w = (x, y, s) \in \mathcal{N}\left(\beta^2\right)$ and $g \in G(\Omega)$. Let $(\Delta x, \Delta y, \Delta s)$ be either MZ or MT family of directions at the predictor step obtained with $\sigma = 0$. Then we define

$$\phi(\alpha) := Q_{\widetilde{x}(\alpha)^{1/2}}\widetilde{s}(\alpha) - \mu(\alpha)e.$$

Note that it is easy to compute that

$$\phi(\alpha) = Q_{\widetilde{x}(\alpha)^{1/2}}\widetilde{s}(\alpha) - (1-\alpha)\mu e.$$

If w is on the central path, we denote $\phi_c(\alpha) := \phi(\alpha)$. In the next two lemmas, we compute derivative of ϕ w.r.t. α and W_P along Δw. Since the proofs are very similar to Lemma 3.2 and 3.3 in SDP, we omit the proofs.

Lemma 5.6. *Let $\phi'(\alpha)$ and $\phi''(\alpha)$ be the first and second order derivatives of $\phi(\alpha)$ w.r.t. α. Then we have*

$$\begin{aligned}\phi'(\alpha) &= 2Q_{\widetilde{x}(\alpha)^{1/2},u}\widetilde{s}(\alpha) + Q_{\widetilde{x}(\alpha)^{1/2}}\widetilde{\Delta s}(\alpha) + \mu e,\\ \phi''(\alpha) &= 2L_{u^2}\widetilde{s}(\alpha) + 2Q_{\widetilde{x}(\alpha)^{1/2},u'}\widetilde{s}(\alpha) + 4Q_{\widetilde{x}(\alpha)^{1/2},u}\widetilde{\Delta s},\end{aligned}$$

where u and u' satisfy

$$2\widetilde{x}(\alpha)^{1/2} \circ u = \widetilde{\Delta x}, \qquad \widetilde{x}(\alpha)^{1/2} \circ u' = -u^2.$$

Lemma 5.7. *Let $W_g(w) := Q_{(gx)^{1/2}}(g^{-*}s) = Q_{\widetilde{x}^{1/2}}\widetilde{s}$. Let $\dot{\widetilde{w}} := (\dot{\widetilde{x}}, \dot{\widetilde{y}}, \dot{\widetilde{s}}) = \left(\frac{\mathrm{d}\widetilde{x}}{\mathrm{d}\nu}, \frac{\mathrm{d}\widetilde{y}}{\mathrm{d}\nu}, \frac{\mathrm{d}\widetilde{s}}{\mathrm{d}\nu}\right)$. The first- and second-order directional derivatives of $W_g(w)$ along the vector $\dot{\widetilde{w}} = (\dot{\widetilde{x}}, \dot{\widetilde{y}}, \dot{\widetilde{s}})$ are given by*

$$\begin{aligned}W_g'(w) &= 2Q_{\widetilde{x}^{1/2},v}\widetilde{s} + Q_{\widetilde{x}^{1/2}}\dot{\widetilde{s}},\\ W_g''(w) &= 2Q'_{\widetilde{x}^{1/2},v}\widetilde{s} + 4Q_{\widetilde{x}^{1/2},v}\dot{\widetilde{s}},\end{aligned}$$

where v and v' satisfy

$$2\widetilde{x} \circ v = \dot{\widetilde{x}}, \qquad \widetilde{x}^{1/2} \circ v' = -v^2.$$

We define the compact region where interior-point methods are conducted.

Definition 5.2.

$$\mathcal{F} := \mathcal{N}(\beta) \cap \{(x, y, s) \in \mathcal{P} \times \mathcal{D} : \nu_f \le \mu(x, s) \le \nu_i\}$$

Now we are ready to prove Theorem 5.5.

Proof. At a certain predictor step, the iterate w_ν is on the central path. The direction is computed from the system (56). Note that the direction on the central path is unique. The necessary and sufficient condition for $w_\nu + \alpha\Delta w_\nu$ to be on the boundary of the neighborhood $\mathcal{N}(\beta)$ is

$$\|\phi_c(\alpha)\| = \beta\mu(\alpha). \tag{65}$$

By applying Taylor's theorem to ϕ_c near 0 up to second degree, (65) is equivalent to

$$\left\|\phi_c(0) + \alpha\phi_c'(0) + \frac{\alpha^2}{2}\phi_c''(0) + L\right\| = \beta\mu(\alpha), \tag{66}$$

where $\|L\| = \mathcal{O}\left(\alpha^3\right)$. Using the identities $\mu(\alpha) = (1-\alpha)\mu$, $\phi_c'(0) = -\phi_c(0) = 0$ and $\phi_c''(0) = \mu^2 W''$, it follows from (66) that

$$\left\|\frac{\alpha^2}{2}\mu^2 W'' + L\right\| = \beta(1-\alpha)\mu, \tag{67}$$

which is identical to (21) in Theorem 3.1. Then, following exactly the same process in the proof of Theorem 3.1, we obtain $\alpha = \mathcal{O}\left(\beta^{1/2}\right)$ and have the square of the integrand of the curvature integral as

$$h_{PD}(\mu) := \alpha^2 \left\|\frac{\mu W''}{2}\right\| / \mu^2 = \alpha^2 \left\|\frac{W''}{2\mu}\right\| = \frac{\beta}{\mu^2}\left(1 - \mathcal{O}\left(\beta^{1/2}\right)\right). \tag{68}$$

Using this $h_{PD}(\mu)$, the rest of the proof is identical to one of Theorem 3.1 in SDP case. □

Theorem 5.8. *[Generic MTY-PC algorithm in MZ family of directions in symmetric cone programs] Let $\beta \in (0, 1/2]$. For given $w^0 \in \mathcal{N}\left(\beta^2\right)$ and $0 < \nu_f < \mu(w^0)$ denote by $\#_{PD}(\mu(w^0), \nu_f, \beta)$ the number of iterations of the MTY-PC algorithm with $\beta \in (0, 1/2]$ needed to reduce the duality gap from $\nu_i := \mu(w^0)$ to ν_f. Then, using the curvature integral:*

$$I_{PD}(\nu_f, \nu_i) = \int_{\nu_f}^{\nu_i} h_{PD}^{1/2}(\nu)\mathrm{d}\nu, \qquad h_{PD}(\nu) := \left\|\frac{W''(x_\nu, s_\nu)}{2\nu}\right\|, \tag{69}$$

we have

$$\lim_{\beta\downarrow 0} \frac{I_{PD}(\nu_f, \nu_i)/\sqrt{\beta}}{\#_{PD}(\nu_i, \nu_f, \beta)} = 1.$$

Proof. Using Taylor's expansions, we rewrite $\phi(\alpha)$ as follows:

$$\phi(\alpha) = [\phi(0) - \phi_c(0)] + \phi_c(0) + \alpha\,[\phi'(0) - \phi_c'(0)] + \alpha\phi_c'(0) + \frac{\alpha^2}{2}\,[\phi''(0) - \phi_c''(0)] + \frac{\alpha^2}{2}\phi_c''(0) + L, \quad (70)$$

with $\|L\| = \mathcal{O}\left(\alpha^3\right)$. As we did in SDP, we will prove the followings:

i) $\|\phi(0) - \phi_c(0)\| = \mathcal{O}\left(\beta^2\right)$,

ii) $\|\phi'(0) - \phi_c'(0)\| = \mathcal{O}\left(\beta^2\right)$,

iii) $\|\phi''(0) - \phi_c''(0)\| = \mathcal{O}\left(\beta^2\right)$.

i) The proof of $\|\phi(0) - \phi_c(0)\| = \mathcal{O}\left(\beta^2\right)$:
We can rewrite as follows:

$$\begin{aligned}\phi(0) - \phi_c(0) &= [Q_{\widetilde{x}^{1/2}}\widetilde{s} - \mu e] - \left[Q_{\widetilde{x}_\nu^{1/2}}\widetilde{s}_\nu - \mu e\right] \\ &= \left(Q_{\widetilde{x}^{1/2}} - Q_{\widetilde{x}_\nu^{1/2}}\right)\widetilde{s} + Q_{\widetilde{x}_\nu^{1/2}}\left(\widetilde{s} - \widetilde{s}_\nu\right).\end{aligned}$$

Then, we have

$$\begin{aligned}Q_{\widetilde{x}^{1/2}} - Q_{\widetilde{x}_\nu^{1/2}} &= \left(2L^2_{\widetilde{x}^{1/2}} - L_{\widetilde{x}}\right) - \left(2L^2_{\widetilde{x}_\nu^{1/2}} - L_{\widetilde{x}_\nu}\right) \\ &= 2\left(L^2_{\widetilde{x}^{1/2}} - L^2_{\widetilde{x}_\nu^{1/2}}\right) - \left(L_{\widetilde{x}} - L_{\widetilde{x}_\nu}\right).\end{aligned}$$

For constants $C_i, i = 1, \ldots, 4$, it follows that

$$\left\|L^2_{\widetilde{x}^{1/2}} - L^2_{\widetilde{x}_\nu^{1/2}}\right\| \le C_1\left\|L_{\widetilde{x}^{1/2}} - L_{\widetilde{x}_\nu^{1/2}}\right\| \le C_2\left\|\widetilde{x}^{1/2} - \widetilde{x}_\nu^{1/2}\right\| \le C_3\,\|\widetilde{x} - \widetilde{x}_\nu\| \le C_4\,\|x - x_\nu\| = \mathcal{O}\left(\beta^2\right).$$

These evaluation are due to the continuity of the operator $(\cdot)^{1/2}$, $L_{(\cdot)}$, and the scaling operator and $\|w - w_c\| = \mathcal{O}\left(\beta^2\right)$ (Lemma 5.12). From above processes, it is easy to see that

$$\left\|Q_{\widetilde{x}^{1/2}} - Q_{\widetilde{x}_\nu^{1/2}}\right\| = \mathcal{O}\left(\beta^2\right).$$

Since we consider only on the compact set $\mathcal{F}$, we can bound the followings:

$$\|\widetilde{s}\|, \left\|Q_{\widetilde{x}_\nu^{1/2}}\right\|. \quad (71)$$

Combining all discussions above, we conclude that

$$\|\phi(0) - \phi_c(0)\| = \mathcal{O}\left(\beta^2\right). \quad (72)$$

ii) The proof of $\|\phi'(0) - \phi_c'(0)\| = \mathcal{O}\left(\beta^2\right)$:
We have

$$\begin{aligned}\phi'(0) - \phi_c'(0) &= \left[2Q_{\widetilde{x}^{1/2},u}\widetilde{s} + Q_{\widetilde{x}^{1/2}}\widetilde{\Delta s} + \mu e\right] - \left[2Q_{\widetilde{x}_\nu^{1/2},u_\nu}\widetilde{s}_\nu + Q_{\widetilde{x}_\nu^{1/2}}\widetilde{\Delta s}_\nu + \mu e\right] \\ &= 2\left[Q_{\widetilde{x}^{1/2},u}\widetilde{s} - Q_{\widetilde{x}_\nu^{1/2},u_\nu}\widetilde{s}_\nu\right] + \left[Q_{\widetilde{x}^{1/2}}\widetilde{\Delta s} - Q_{\widetilde{x}_\nu^{1/2}}\widetilde{\Delta s}_\nu\right] \\ &= 2\left[\left(Q_{\widetilde{x}^{1/2},u} - Q_{\widetilde{x}_\nu^{1/2},u_\nu}\right)\widetilde{s} + Q_{\widetilde{x}_\nu^{1/2},u_\nu}\left(\widetilde{s} - \widetilde{s}_\nu\right)\right] \\ &\quad + \left[\left(Q_{\widetilde{x}^{1/2}} - Q_{\widetilde{x}_\nu^{1/2}}\right)\widetilde{\Delta s} + Q_{\widetilde{x}_\nu^{1/2}}\left(\widetilde{\Delta s} - \widetilde{\Delta s}_\nu\right)\right].\end{aligned}$$

We evaluate first term only. Other terms can be done in a similar way.

$$\begin{aligned}Q_{\widetilde{x}^{1/2},u} - Q_{\widetilde{x}_\nu^{1/2},u_\nu} &= \left[L_{\widetilde{x}^{1/2}}L_u + L_uL_{\widetilde{x}^{1/2}} - L_{\widetilde{x}^{1/2}\circ u}\right] - \left[L_{\widetilde{x}_\nu^{1/2}}L_{u_\nu} + L_{u_\nu}L_{\widetilde{x}_\nu^{1/2}} - L_{\widetilde{x}_\nu^{1/2}\circ u_\nu}\right] \\ &= \left[L_{\widetilde{x}^{1/2}}L_u - L_{\widetilde{x}_\nu^{1/2}}L_{u_\nu}\right] + \left[L_uL_{\widetilde{x}^{1/2}} - L_{u_\nu}L_{\widetilde{x}_\nu^{1/2}}\right] - \left[L_{\widetilde{x}^{1/2}\circ u} - L_{\widetilde{x}_\nu^{1/2}\circ u_\nu}\right].\end{aligned}$$

The first term of the above equation is evaluated by

$$\left\| L_{\widetilde{x}^{1/2}} L_u - L_{\widetilde{x}_\nu^{1/2}} L_{u_\nu} \right\| \le \left\| \left(L_{\widetilde{x}^{1/2}} - L_{\widetilde{x}_\nu^{1/2}} \right) L_u \right\| + \left\| L_{\widetilde{x}_\nu^{1/2}} \left(L_u - L_{u_\nu} \right) \right\|.$$

The problem here is the evaluation of $\|L_u - L_{u_\nu}\|$. Since

$$\left\| L^{-1}_{\widetilde{x}^{1/2}} - L^{-1}_{\widetilde{x}_\nu^{1/2}} \right\| \le C \left\| L_{\widetilde{x}^{1/2}} - L_{\widetilde{x}_\nu^{1/2}} \right\| = \mathcal{O}\left(\beta^2\right), \quad (C : \text{constant}),$$

$\|\Delta x - \Delta x_\nu\| = \mathcal{O}\left(\beta^2\right)$ (Lemma 5.12) and

$$\begin{aligned} \|u - u_\nu\| &= \frac{1}{2} \left\| L^{-1}_{\widetilde{x}^{1/2}} \widetilde{\Delta x} - L^{-1}_{\widetilde{x}_\nu^{1/2}} \widetilde{\Delta x}_\nu \right\| \\ &= \frac{1}{2} \left\| \left(L^{-1}_{\widetilde{x}^{1/2}} - L^{-1}_{\widetilde{x}_\nu^{1/2}} \right) \widetilde{\Delta x} + L^{-1}_{\widetilde{x}_\nu^{1/2}} \left(\widetilde{\Delta x} - \widetilde{\Delta x}_\nu \right) \right\| = \mathcal{O}\left(\beta^2\right), \end{aligned}$$

we have

$$\|L_u - L_{u_\nu}\| = \mathcal{O}\left(\beta^2\right).$$

Therefore,

$$\left\| Q_{\widetilde{x}^{1/2},u} - Q_{\widetilde{x}_\nu^{1/2},u_\nu} \right\| = \mathcal{O}\left(\beta^2\right).$$

To sum up, we have

$$\|\phi'(0) - \phi'_c(0)\| = \mathcal{O}\left(\beta^2\right). \tag{73}$$

iii) The proof of $\|\phi''(0) - \phi''_c(0)\| = \mathcal{O}\left(\beta^2\right)$:
We have

$$\begin{aligned} \phi''(0) - \phi''_c(0) &= \left[2L_{u^2}\widetilde{s} + 2Q_{\widetilde{x}^{1/2},u'}\widetilde{s} + 4Q_{\widetilde{x}^{1/2},u}\widetilde{\Delta s} \right] \\ &\quad - \left[2L_{u_\nu^2}\widetilde{s}_\nu + 2Q_{\widetilde{x}_\nu^{1/2},u'_\nu}\widetilde{s}_\nu + 4Q_{\widetilde{x}_\nu^{1/2},u_\nu}\widetilde{\Delta s}_\nu \right] \\ &= 2\left[L_{u^2}\widetilde{s} - L_{u_\nu^2}\widetilde{s}_\nu \right] + 2\left[Q_{\widetilde{x}^{1/2},u'}\widetilde{s} - Q_{\widetilde{x}_\nu^{1/2},u'_\nu}\widetilde{s}_\nu \right] \\ &\quad + 4\left[Q_{\widetilde{x}^{1/2},u}\widetilde{\Delta s} - Q_{\widetilde{x}_\nu^{1/2},u_\nu}\widetilde{\Delta s}_\nu \right]. \end{aligned}$$

According to the experiences from i) and ii), the problem here is the evaluation of $\|u' - u'_\nu\|$. Since $\left\|u^2 - u_\nu^2\right\| \le C\|u - u_\nu\| = \mathcal{O}\left(\beta^2\right)$ with a constant C, we have

$$\|u' - u'_\nu\| = \left\| L^{-1}_{\widetilde{x}^{1/2}} u^2 - L^{-1}_{\widetilde{x}_\nu^{1/2}} u_\nu^2 \right\| = \left\| \left(L^{-1}_{\widetilde{x}^{1/2}} - L^{-1}_{\widetilde{x}_\nu^{1/2}} \right) u^2 + L^{-1}_{\widetilde{x}_\nu^{1/2}} \left(u^2 - u_\nu^2 \right) \right\| = \mathcal{O}\left(\beta^2\right).$$

Thus, we conclude that

$$\|\phi''(0) - \phi''_c(0)\| = \mathcal{O}\left(\beta^2\right). \tag{74}$$

Therefore, $\phi(\alpha)$ in (70) is rewritten as

$$\phi(\alpha) = \phi_c(0) + \alpha\phi'_c(0) + \frac{\alpha^2}{2}\phi''_c(0) + L + \mathcal{O}\left(\beta^2\right). \tag{75}$$

Using the identities $\mu(\alpha) = (1-\alpha)\mu$, $\phi'_c(0) = -\phi_c(0) = 0$ and $\phi''(0) = \mu^2 W''$, the boundary condition is now

$$\left\| \frac{\alpha^2}{2}\mu^2 W'' + L + \mathcal{O}\left(\beta^2\right) \right\| = \beta(1-\alpha)\mu, \tag{76}$$

which is very similar to (67) in the idealized algorithm. As we did in the SDP case, we estimate $\alpha = \mathcal{O}\left(\beta^{1/2}\right)$ and

$$h_{PD}(\mu) := \alpha^2 \left\| \frac{\mu W''}{2} \right\| / \mu^2 = \alpha^2 \left\| \frac{W''}{2\mu} \right\| = \frac{\beta}{\mu^2}\left(1 - \mathcal{O}\left(\beta^{1/2}\right)\right).$$

Applying the same logic of the proof in SDP case, we obtain the conclusion. □

In what follows, we make a proof of a series of lemmas corresponding to Lemma 3.5 and 3.6.

Lemma 5.9. *On the central path, it follows in HKM dual direction ($g = Q_{x^{-1/2}}$) that*

$$\frac{W_g''}{2\mu} = Q_{x^{-1/2}}\dot{x} \circ Q_{s^{-1/2}}\dot{s}. \tag{77}$$

Proof. In HKM dual direction, we have the identities: $\widetilde{x} = e$, $\widetilde{s} = \mu e$, $\dot{\widetilde{x}} = 2u$, $u^2 + u' = 0$. Therefore, we have

$$\begin{aligned} W_g'' &= 2\left(Q_{\widetilde{x}^{1/2},u}\widetilde{s}\right)' + 2Q_{{}^t x^{1/2},u}\dot{\widetilde{s}} \\ &= 2\left(\mu u^2 + \mu u' + u \circ \dot{\widetilde{s}}\right) + 2u \circ \dot{\widetilde{s}} \\ &= 4u \circ \dot{\widetilde{s}} = 2\dot{\widetilde{x}} \circ \dot{\widetilde{s}} = 2\mu\, Q_{x^{-1/2}}\dot{x} \circ Q_{s^{-1/2}}\dot{s}. \end{aligned}$$

□

Lemma 5.10. *On the central path, $\|W_g''\|$ is invariant under any $g \in G(\Omega)$.*

Proof. Since

$$\begin{aligned} (W_g \circ W_g)' &= 2W_g' \circ W_g, \\ (W_g \circ W_g)'' &= 2\left(W_g'' \circ W_g + (W_g')^2\right), \\ (W_g \circ W_g)^{(3)} &= 2\left(W_g^{(3)} \circ W_g + 3W_g'' \circ W_g'\right), \\ (W_g \circ W_g)^{(4)} &= 2\left(W_g^{(4)} \circ W_g + 4W_g^{(3)} \circ W_g' + 3(W_g'')^2\right), \end{aligned}$$

and $W_g = \mu e$ and $W_g' = e$, we have

$$\operatorname{Tr}((W_g \circ W_g)^{(4)}) = 2\left(\mu \operatorname{Tr}(W_g^{(4)}) + 4\operatorname{Tr}(W_g^{(3)}) + 3\left\|W_g''\right\|^2\right).$$

Now, it is known that $\operatorname{Tr}(W_g) = \mu$ and $\operatorname{Tr}(W_g \circ W_g)$ are invariant for any $g \in G(\Omega)$ (see, for instance, Proposition 4.2 of [12]). Therefore, $\operatorname{Tr}((W_g \circ W_g)^{(4)})$, $\operatorname{Tr}(W_g^{(4)})$, $\operatorname{Tr}(W_g^{(3)})$ are invariant under $g \in G(\Omega)$, so is $\left\|W_g''\right\|^2$. □

Combining Theorem 5.8, Lemma 5.9 and 5.10, we obtain the final theorem:

Theorem 5.11. *Let $\beta \in (0, 1/2]$. For given $w^0 \in \mathcal{N}(\beta)$ and $0 < \nu_f < \mu(w^0)$ denote by $\#_{PD}(\mu(w^0), \nu_f, \beta)$ the number of iterations of the MTY-PC algorithm in MZ family of directions with $\beta \in (0, 1/2]$ needed to reduce the duality gap from $\nu_i := \mu(w^0)$ to ν_f. Then, using the curvature integral:*

$$I_{PD}(\nu_f, \nu_i) = \int_{\nu_f}^{\nu_i} \|Q_{x^{-1/2}}\dot{x} \circ Q_{s^{-1/2}}\dot{s}\|^{1/2}\, \mathrm{d}\nu, \tag{78}$$

we have

$$\lim_{\beta \downarrow 0} \frac{I_{PD}(\nu_f, \nu_i)/\sqrt{\beta}}{\#_{PD}(\nu_i, \nu_f, \beta)} = 1.$$

Remark 5.1. As we discussed in Remark 3.1, it also holds in symmetric cone programs that the curvature of the central path is zero in all the parts of the trajectory if it is zero at some point, i.e.,

$$h_{PD}(\nu) = \|Q_{x^{-1/2}}\dot{x} \circ Q_{s^{-1/2}}\dot{s}\| = 0 \text{ for } \forall \nu > 0 \quad \text{if } h_{PD}(\nu_0) = 0 \text{ at some } \nu_0 > 0.$$

So we restrict our discussion in Theorem 5.5, 5.8 and 5.11 in case the curvature is larger than zero. The proof is similar to Lemma 2.1 of [10] and we will show as follows.

Differentiating the central path equation (55) w.r.t. $\nu > 0$, we have

$$\mathcal{A}^*\dot{y} + \dot{s} = 0,\ \mathcal{A}\dot{x} = 0,\ \dot{x} \circ s + x \circ \dot{s} = e.$$

Assume $h_{PD}(\nu_0) = 0$ for some $\nu_0 > 0$ and define $\widetilde{w}(\nu) = w(\nu_0) + (\nu - \nu_0)\dot{w}(\nu_0)$ for $\nu > 0$. Since $Q_{x^{-1/2}}\dot{x} \circ Q_{s^{-1/2}}\dot{s} = 0$ if and only if $\dot{x} \circ \dot{s} = 0$ from Lemma 5.4 and hence $\widetilde{w}(\nu)$ satisfies the central path equation (55). Since the solution of the equation (55) is unique and hence $w(\nu) = \widetilde{w}(\nu)$ for all $\nu > 0$, we have $\dot{w}(\nu) = \dot{w}(\nu_0)$ for all $\nu > 0$, which implies that $h_{PD}(\nu) = 0$ for all $\nu > 0$.

In SDP case, we obtain same result since it is a special class of symmetric cone programs.

The following two lemmas state the evaluations of the order of $\|w - w_c\|$ and $\|\Delta w - \Delta w_\nu\|$. Since proofs are essentially identical to SDP cases, we omit the proofs.

Lemma 5.12. *Let $w \in \mathcal{N}\left(\beta^2\right)$ be a feasible point with the duality gap being ν and let w_ν be the point on the central path with the same duality gap. Then, asymptotically, the following relation holds:*

$$\|w - w_\nu\| = \mathcal{O}\left(\beta^2\right).$$

Lemma 5.13. *Let $(x, y, s) \in \mathcal{N}\left(\beta^2\right)$ be a feasible point whose duality gap is ν. Let $(\Delta x, \Delta y, \Delta s)$ be one instance of MZ family of directions at the predictor step, namely, the solution to the system of equations:*

$$\begin{array}{rcl} gx \circ g^{-*}\Delta s + g\Delta x \circ g^{-*}s &=& -(gs) \circ (g^{-*}s), \\ \sum_i \Delta y_i a_i + \Delta s &=& 0, \\ \langle a_i, \Delta x\rangle &=& 0, \quad i = 1, \ldots, m. \end{array} \tag{79}$$

In particular, let $(\Delta x_\nu, \Delta y_\nu, \Delta s_\nu)$ be the solution of (79) with (x_ν, y_ν, s_ν) being on the central path whose duality gap is ν. Then it holds that

$$\|(\Delta x, \Delta y, \Delta s) - (\Delta x_\nu, \Delta y_\nu, \Delta s_\nu)\| = \mathcal{O}\left(\beta^2\right).$$

6 Conclusions

In this paper, we studied iteration complexities of MTY-PC algorithms in SDP and symmetric cone programs via curvature integrals. We essentially extended the result in Monteiro-Tsuchiya [10] in LP case, namely, the curvature integral in LP:

$$h_{PD}(\nu) := \frac{\|\dot{x} \circ \dot{s}\|}{\nu}, \quad I_{PD}(\nu_f, \nu_i) = \int_{\nu_f}^{\nu_i} h_{PD}^{1/2}(\nu)\mathrm{d}\nu \tag{80}$$

is approximated as

$$I_{PD}(\nu_f, \nu_i) \approx \sqrt{\beta} \times \#_{PD}(\nu_i, \nu_f, \beta). \tag{81}$$

More specifically, using the following square of integrands

$$\begin{array}{rcl} h_{PD}(\nu) &:=& \left\|X^{-1/2}\dot{X}X^{-1/2} \circ S^{-1/2}\dot{S}S^{-1/2}\right\|_F \quad \text{(SDP case)}, \\ h_{PD}(\nu) &:=& \|Q_{x^{-1/2}}\dot{x} \circ Q_{s^{-1/2}}\dot{s}\| \quad \text{(symmetric cone programs case)}, \end{array}$$

we obtained the same asymptotic results as (81). Additionally, we conducted numerical experiments in practically large SDP problems from SDPLIB, which indicates that our results serve as a useful analytical tool for various SDP problems.

Acknowledgement

The authors are grateful to Professor Renato D. C. Monteiro of the School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, for his insightful discussion about the curvature integral I_{PD}.

References

[1] B. Borchers. SDPLIB 1.2, library of semidefinite programming test problems. *Optim. Methods Softw.*, 11/12(1-4):683–690, 1999. Interior point methods.

[2] J. W. Demmel. *Applied numerical linear algebra.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[3] J. Faraut and A. Korányi. *Analysis on symmetric cones.* Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 1994. Oxford Science Publications.

[4] L. Faybusovich. Linear systems in Jordan algebras and primal-dual interior-point algorithms. *J. Comput. Appl. Math.*, 86(1):149–175, 1997. Special issue dedicated to William B. Gragg (Monterey, CA, 1996).

[5] S. Kakihara, A. Ohara, and T. Tsuchiya. A geometric relationship of iteration complexities between primal algorithms and primal-dual algorithms in SDP and symmetric cone programs. *Optimization Online*, 2011.

[6] N. Karmarkar. Riemannian geometry underlying interior-point methods for linear programming. In *Mathematical developments arising from linear programming (Brunswick, ME, 1988)*, volume 114 of *Contemp. Math.*, pages 51–75. Amer. Math. Soc., Providence, RI, 1990.

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

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

[9] R. D. C. Monteiro and T. Tsuchiya. Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming. *SIAM J. Optim.*, 9(3):551–577 (electronic), 1999.

[10] R. D. C. Monteiro and T. Tsuchiya. A strong bound on the integral of the central path curvature and its relationship with the iteration-complexity of primal-dual path-following LP algorithms. *Math. Program.*, 115(1, Ser. A):105–149, 2008.

[11] R. D. C. Monteiro and Y. Zhang. A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinite programming. *Math. Program.*, 81(3, Ser. A):281–299, 1998.

[12] M. Muramatsu. On a commutative class of search directions for linear programming over symmetric cones. *J. Optim. Theory Appl.*, 112(3):595–625, 2002.

[13] A. Ohara and T. Tsuchiya. An information geometric approach to polynomial-time interior-point algorithms: Complexity bound via curvature integral. *Research Memorandum The Institute of Statistical Mathematics*, 1055, 2007.

[14] S. H. Schmieta and F. Alizadeh. Extension of primal-dual interior point algorithms to symmetric cones. *Math. Program.*, 96(3, Ser. A):409–438, 2003.

[15] G. Sonnevend, J. Stoer, and G. Zhao. On the complexity of following the central path of linear programs by linear extrapolation. II. *Math. Program.*, 52(3, Ser. B):527–553 (1992), 1991. Interior point methods for linear programming: theory and practice (Scheveningen, 1990).

[16] M. J. Todd, K. C. Toh, and R. H. Tütüncü. On the Nesterov-Todd direction in semidefinite programming. *SIAM J. Optim.*, 8(3):769–796 (electronic), 1998.

[17] G. Zhao and J. Stoer. Estimating the complexity of a class of path-following methods for solving linear programs by curvature integrals. *Appl. Math. Optim.*, 27(1):85–103, 1993.