Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 5, No. 2, pp. 311–335, May 2009

HOMOGENEOUS ALGORITHMS FOR MONOTONE COMPLEMENTARITY PROBLEMS OVER SYMMETRIC CONES

Akiko Yoshise*

Abstract: In [24], the author proposed a homogeneous model for standard monotone nonlinear complementarity problems over symmetric cones and show that the following properties hold: (a) There is a path that is bounded and has a trivial starting point without any regularity assumption concerning the existence of feasible or strictly feasible solutions. (b) Any accumulation point of the path is a solution of the homogeneous model. (c) If the original problem is solvable, then every accumulation point of the path gives us a finite solution. (d) If the original problem is strongly infeasible, then, under the assumption of Lipschitz continuity, any accumulation point of the path gives us a finite certificate proving infeasibility. In this paper, we propose a class of algorithms for numerically tracing the path in (a) above. Let r be the rank of the intended Euclidian Jordan algebra. By introducing a parameter $\theta \geq 0$ for quantifying a scaled Lipschitz property of a function, we obtain the following results: (e1) The (infeasible) NT method takes $O(\sqrt{r}(1+\sqrt{r}\ \theta)\log\epsilon^{-1})$ iterations for the short-step, and $O(r(1+\sqrt{r}\ \theta)\log\epsilon^{-1})$ iterations for the semi-long- and long-step variants. (e2) The (infeasible) xy method or yx method takes $O(\sqrt{r}(1+\sqrt{r}\ \theta)\log\epsilon^{-1})$ iterations for the short-step, $O(r(1+\sqrt{r}\ \theta)\log\epsilon^{-1})$ iterations for the semi-long-step, and $O(r^{1.5}(1+\sqrt{r}\ \theta)\log\epsilon^{-1})$ iterations for the long-step variant. If the original complementarity problem is linear then $\theta = 0$ and the above results achieve the best iteration-complexity bounds known so far for linear or convex quadratic optimization problems over symmetric cones.

Key words: *complementarity problem, symmetric cone, homogeneous algorithm, interior point method, complexity analysis*

Mathematics Subject Classification: *90C22, 90C25, 90C33, 65K05, 46N10*

---

1 Introduction

In 1999, Andersen and Ye [3] provided a homogeneous model for solving monotone complementarity problems by generalizing the homogeneous self-dual algorithm for linear programming. Their sophisticated model has the following desirable properties:

(a) The homogeneous model has a bounded path with a trivial starting point without any regularity assumption concerning the existence of feasible or strictly feasible solutions.

(b) Any accumulation point of the path is a solution of the homogeneous model.

*Research supported in part by Grants-in-Aid for Scientific Research ((C)1856052, (C)17560050, (B)1831001, Exploratory Research 18651076) of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

(c) If the original problem is solvable, then every accumulation point of the path gives us a finite solution.

(d) If the original problem is strongly infeasible, then, under the assumption of Lipschitz continuity, any accumulation point of the path gives us a finite certificate proving infeasibility.

(e) There exists an algorithm which solves the homogeneous model in $\mathcal{O}(\sqrt{n}\log \epsilon^{-1})$ number of iterations whenever the original problem is linear.

In [24], the author extended the model to problems over symmetric cones in Euclidean Jordan Algebras and showed that the proposed model has the properties (a)–(d) above. This paper addresses to extending the last property (e) to symmetric cone cases. We propose a class of polynomial-time algorithms for numerically tracing the path in (a) above and to derive their iteration-complexity bounds corresponding to (e).

Let $(V, \circ)$ be a Euclidian Jordan algebra with identity e, i.e., V is a finite dimensional vector space and the bilinear product $x \circ y$ satisfies for all $x, y \in V$,

(i) $x \circ y = y \circ x$,

(ii) $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$ where $x^2 = x \circ x$,

(iii) $x^2 + y^2 = 0 \implies [x = 0,\ y = 0]$, and

(iv) $x \circ e = e \circ x = x$.

Here, the statement (iii) can be replaced equivalently by

(iii') there exists an inner product such that $\langle x \circ y, z\rangle = \langle y, x \circ z\rangle$

and specially, we can set

$$\langle x, y\rangle = \mathrm{tr}\,(x \circ y)$$

where $\mathrm{tr}\,(x \circ y)$ denotes the first coefficient of minimal polynomial of $x \circ y$, which is positive definite under (i)–(iii) (cf. [5]). The inner product induces a unitarily invariant norm

$$\|x\|_{\mathrm{F}} = \sqrt{\mathrm{tr}\,(x \circ x)}$$

for $x \in V$.

Let r be the rank of V, and let $\lambda_i(x)$ $(i = 1, 2, \ldots, r)$ be the eigenvalues of $x \in V$. Then

$$\|x\|_{\mathrm{F}} = \sqrt{\sum_{i=1}^{r} \lambda_i(x)^2}$$

holds and

$$\|x\|_2 = \max\{|\lambda_i(x)|\ (i = 1, 2, \ldots, r)\}$$

determines another unitarily invariant norm of $x \in V$. Note that $\|e\|_{\mathrm{F}} = \sqrt{r}$ and $\|e\|_2 = 1$. We denote by K the symmetric cone of V which is a self-dual closed convex cone such that for any two elements $x \in \mathrm{int}K$ and $y \in \mathrm{int}K$, there exists an invertible linear map $\Gamma : V \to V$ satisfying $\Gamma(K) = K$ and $\Gamma(x) = y$. In a Euclidean Jordan algebra, it is known that a cone in V is symmetric if and only if it is the cone of squares of V given by $K = \{x \circ x :\ x \in V\}$.

Since $\circ$ is a bilinear map, for any $x \in V$, we can define a linear operator $L(x)$ satisfying $L(x)y = x \circ y$ for all $y \in V$. For $x, y \in V$, let

$$Q_{x,y} := L(x)L(y) + L(y)L(x) - L(x \circ y), \quad Q_x := Q_{x,x} = 2L^2(x) - L(x^2)$$

where Q_x is called the quadratic representation of x. Since the statement (iii') implies that both of $L(x)$ and $L(x^2)$ are self-adjoint, Q_x is also self-adjoint. For each $x \in \mathrm{int}K$, x is invertible and $Q_{x^{-1}} = Q_x^{-1}$ holds (cf. [5]).

There is extensive literature on the analysis of optimization problems over symmetric cones. Some reasons for this may be

- symmetric cones are convenient tools for investigating the theoretical aspects of interior point algorithms for nonnegative orthants, second-order cones and positive semidefinite cones in a unified manner (cf. [6, 7, 20, 22, 24]),
- self-scaled cones introduced in [18] are closely related to symmetric cones, more precisely the same as symmetric cones (cf. [12, 13, 21]),
- more algebraic approach to optimization problems or complementarity problems becomes possible by considering symmetric cones (cf. [5, 8, 10, 11, 23]).

Among others, Schmieta and Alizadeh [22] established an indispensable basis for developing primal-dual interior point algorithms for solving linear programs over symmetric cones. Many results in this paper depend on their fundamental work.

Consider the following standard monotone nonlinear complementarity problem over the symmetric cone K of V.

$$\begin{array}{lll} \text{(SCP)} & \text{Find} & (x, y) \in K \times K \\ & \text{s.t.} & F(x, y) := y - \psi(x) = 0, \ x \circ y = 0 \end{array}$$

where $\psi : K \to V$ is a differentiable monotone function on K satisfying

$$\langle \psi(x) - \psi(x'), x - x' \rangle \geq 0 \ \text{ for all } x, x' \in K.$$

In [24], the author proposed a homogeneous model HCP for the SCP:

$$\begin{array}{lll} \text{(HCP)} & \text{Find} & (x, \tau, y, \kappa) \in (K \times \Re_{++}) \times (K \times \Re_{+}) \\ & \text{s.t.} & F_{\mathrm{H}}(x, \tau, y, \kappa) = 0, \ (x, \tau) \circ_{\mathrm{H}} (y, \kappa) = 0 \end{array}$$

where

$$\begin{aligned} & \Re_{+} := \{\tau \in \Re : \ \tau \geq 0\}, \quad \Re_{++} := \{\tau \in \Re : \ \tau > 0\}, \\ & V_{\mathrm{H}} := V \times \Re, \quad K_{\mathrm{H}} := K \times \Re_{+}, \quad x_{\mathrm{H}} := (x, \tau) \in V_{\mathrm{H}}, \quad y_{\mathrm{H}} := (y, \kappa) \in V_{\mathrm{H}}, \\ & \psi_{\mathrm{H}}(x_{\mathrm{H}}) = \psi_{\mathrm{H}}(x, \tau) := \begin{pmatrix} \tau\psi(x/\tau) \\ -\langle \psi(x/\tau), x \rangle \end{pmatrix}, \\ & F_{\mathrm{H}}(x_{\mathrm{H}}, y_{\mathrm{H}}) = y_{\mathrm{H}} - \psi_{\mathrm{H}}(x_{\mathrm{H}}), \\ & x_{\mathrm{H}} \circ_{\mathrm{H}} y_{\mathrm{H}} := \begin{pmatrix} x \circ y \\ \tau\kappa \end{pmatrix}. \end{aligned} \tag{1.1}$$

We also define

$$\langle x_{\mathrm{H}}, y_{\mathrm{H}} \rangle_{\mathrm{H}} := \mathrm{tr}\,(x_{\mathrm{H}} \circ y_{\mathrm{H}}) = \mathrm{tr}\,(x \circ y) + \tau\kappa = \langle x, y \rangle + \tau\kappa. \tag{1.2}$$

The set K_{H} is a Cartesian product of two symmetric cones K and $\Re_+$ and is the symmetric cone of V_{H} given by

$$K_{\mathrm{H}} = \left\{ x_{\mathrm{H}}^2 = \begin{pmatrix} x^2 \\ \tau^2 \end{pmatrix} : \ x_{\mathrm{H}} \in V_{\mathrm{H}} \right\}.$$

We can see that the monotonicity of ψ on K implies the monotonicity of ψ_{H} on $\mathrm{int}K_{\mathrm{H}}$ (Proposition 5.3 of [24]). However, we should handle the functions ψ_{H} and F_{H} more carefully since they are not necessarily defined on the boundary of their domains. We introduce the following definitions of *asymptotic* feasibility and infeasibility.

- The SCP is *asymptotically feasible* if and only if there exists a bounded sequence $\{(x^{(k)}, y^{(k)})\} \subseteq \mathrm{int}K \times \mathrm{int}K$ such that
$$\lim_{k\to\infty} F(x^{(k)}, y^{(k)}) = 0.$$
- The SCP is *asymptotically solvable* if and only if there exists a bounded sequence $\{(x^{(k)}, y^{(k)})\} \subseteq \mathrm{int}K \times \mathrm{int}K$ such that
$$\lim_{k\to\infty} F(x^{(k)}, y^{(k)}) = 0 \text{ and } \lim_{k\to\infty} x^{(k)} \circ y^{(k)} = 0.$$
- The SCP is *infeasible* if and only if there is no feasible point $(x, y) \in K \times K$ satisfying $F(x, y) = 0$.
- The SCP is *strongly infeasible* if and only if there is no sequence $\{(x^{(k)}, y^{(k)})\} \subseteq \mathrm{int}K \times \mathrm{int}K$ such that $\lim_{k\to\infty} F(x^{(k)}, y^{(k)}) = 0$.

The following results have been shown in [24].

Theorem 1.1 (Theorem 5.4 and 5.5 of [24]). *Define*

$$h_{\mathrm{H}}^{(0)} := \begin{pmatrix} p_{\mathrm{H}}^{(0)} \\ f_{\mathrm{H}}^{(0)} \end{pmatrix} := \begin{pmatrix} x_{\mathrm{H}}^{(0)} \circ_{\mathrm{H}} y_{\mathrm{H}}^{(0)} \\ F_{\mathrm{H}}(x_{\mathrm{H}}^{(0)}, y_{\mathrm{H}}^{(0)}) \end{pmatrix} = \begin{pmatrix} e_{\mathrm{H}} \\ y_{\mathrm{H}}^{(0)} - \psi_{\mathrm{H}}(x_{\mathrm{H}}^{(0)}) \end{pmatrix}$$

where $(x_{\mathrm{H}}^{(0)}, y_{\mathrm{H}}^{(0)}) = (e_{\mathrm{H}}, e_{\mathrm{H}})$, $e_{\mathrm{H}} = (e, 1) \in \mathrm{int}K_{\mathrm{H}}$ *is the identity element in* V_{H} *satisfying*

$$\mathrm{tr}\,(e_{\mathrm{H}}) = \mathrm{rank}\,(V_{\mathrm{H}}) = r + 1.$$

(i) *Any asymptotically feasible solution* $(\hat{x}_{\mathrm{H}}, \hat{y}_{\mathrm{H}})$ *of the HCP is an asymptotic solution, i.e., there exists a bounded sequence* $\{(\hat{x}_{\mathrm{H}}^{(k)}, \hat{y}_{\mathrm{H}}^{(k)})\} \subseteq \mathrm{int}K_{\mathrm{H}} \times \mathrm{int}K_{\mathrm{H}}$ *such that*

$$\lim_{k\to\infty} (\hat{x}_{\mathrm{H}}^{(k)}, \hat{y}_{\mathrm{H}}^{(k)}) = (\hat{x}_{\mathrm{H}}, \hat{y}_{\mathrm{H}}), \ \lim_{k\to\infty} F_{\mathrm{H}}(\hat{x}_{\mathrm{H}}^{(k)}, \hat{y}_{\mathrm{H}}^{(k)}) = 0 \ \textit{and} \ \lim_{k\to\infty} \hat{x}_{\mathrm{H}}^{(k)} \circ_{\mathrm{H}} \hat{y}_{\mathrm{H}}^{(k)} = 0.$$

(ii) *The HCP is not feasible, but asymptotically feasible.*

(iii) *The SCP has a solution if and only if the HCP has an asymptotic solution* $(x_{\mathrm{H}}^*, y_{\mathrm{H}}^*) = (x^*, \tau^*, y^*, \kappa^*)$ *with* $\tau^* > 0$. *In this case,* $(x^*/\tau^*, y^*/\tau^*)$ *is a solution of the SCP.*

(iv) *Suppose that* ψ *satisfies the Lipschitz condition on* $\mathrm{int}K$, *i.e., there exists a constant* $\gamma \geq 0$ *such that*

$$\|\psi(x + h) - \psi(x)\|_{\mathrm{F}} \leq \gamma \|h\|_{\mathrm{F}} \ \ \textit{for any} \ \ x \in \mathrm{int}K \ \textit{and} \ x + h \in \mathrm{int}K.$$

If the SCP is strongly infeasible then the HCP has an asymptotic solution $(x^*, \tau^*, y^*, \kappa^*)$ *with* $\kappa^* > 0$. *Conversely, if the HCP has an asymptotic solution* $(x^*, \tau^*, y^*, \kappa^*)$ *with* $\kappa^* > 0$ *then the SCP is infeasible. In the latter case,* $(x^*/\kappa^*, y^*/\kappa^*)$ *is a certificate to prove infeasibility of the SCP.*

(v) *Define*

$$H_{\rm H}(x_{\rm H}, y_{\rm H}) := \begin{pmatrix} x_{\rm H} \circ_{\rm H} y_{\rm H} \\ y_{\rm H} - \psi_{\rm H}(x_{\rm H}) \end{pmatrix}.$$

The set

$$P := \{(x_{\rm H}, y_{\rm H}) \in {\rm int}K_{\rm H} \times {\rm int}K_{\rm H} : \ H_{\rm H}(x_{\rm H}(t), y_{\rm H}(t)) = th_{\rm H}^{(0)}, \ t \in (0,1]\}$$

forms a bounded path in ${\rm int}K_{\rm H} \times {\rm int}K_{\rm H}$. *Any accumulation point* $(x_{\rm H}(0), y_{\rm H}(0))$ *is an asymptotic solution of the HCP.*

(vi) *If the HCP has an asymptotic solution* $(x_{\rm H}^*, y_{\rm H}^*) = (x^*, \tau^*, y^*, \kappa^*)$ *with* $\tau^* > 0$ *(* $\kappa^* > 0$, *respectively), then any accumulation point* $(x_{\rm H}(0), y_{\rm H}(0)) = (x(0), \tau(0), y(0), \kappa(0))$ *of the bounded path* P *satisfies* $\tau(0) > 0$ *(* $\kappa(0) > 0$, *respectively).*

The above theorem ensures that the properties (a)–(d) of the homogeneous model for $K = \Re^n_+$ in [3] can be extended to the case of symmetric cones.

In this paper, we propose a class of algorithms which trace the path in (v) of Theorem 1.1. We give an outline of our homogeneous algorithm. See Section 7 for a complete description of the algorithm.

We start the algorithm with the following infeasible initial point.

$$(x_{\rm H}^0, y_{\rm H}^0) := (e_{\rm H}, e_{\rm H}), \quad s_{\rm H}^0 := y_{\rm H}^0 - \psi_{\rm H}(x_{\rm H}), \quad \mu_{\rm H}^0 := \langle x_{\rm H}^0, y_{\rm H}^0 \rangle_{\rm H}/(r+1).$$

At each iteration $(x_{\rm H}, y_{\rm H})$, we consider the following system.

$$H_{\rm H}(x_{\rm H}, y_{\rm H}) = \begin{pmatrix} x_{\rm H} \circ_{\rm H} y_{\rm H} \\ y_{\rm H} - \psi_{\rm H}(x_{\rm H}) \end{pmatrix} = \begin{pmatrix} \mu_{\rm H} e_{\rm H} \\ 0 \end{pmatrix}. \tag{1.3}$$

Applying Newton's method to the system leads us to the linear system

$$\begin{cases} \Delta x_{\rm H} \circ_{\rm H} y_{\rm H} + x_{\rm H} \circ_{\rm H} \Delta y_{\rm H} = \gamma \mu_{\rm H} e - x_{\rm H} \circ_{\rm H} y_{\rm H}, \\ \Delta y_{\rm H} - D\psi(x_{\rm H}) \Delta x_{\rm H} = -\eta s_{\rm H}, \end{cases} \tag{1.4}$$

where

$$(\Delta x_{\rm H}, \ \Delta y_{\rm H}) \in V_{\rm H} \times V_{\rm H}, \ s_{\rm H} := y_{\rm H} - \psi_{\rm H}(x_{\rm H}), \ \mu_{\rm H} = \langle x_{\rm H}, y_{\rm H} \rangle_{\rm H}/(r+1), \tag{1.5}$$

and $\eta, \gamma \in [0,1]$ are parameters for regulating the feasibility and the complementarity, respectively. The direction obtained by the above system is so called $xy + yx$ direction [1]. Here, we consider the commutative class of search directions, which is a subclass of Monteiro and Zhang family (cf. [16,17]). For $(x_{\rm H}, y_{\rm H}) \in {\rm int}K_{\rm H} \times {\rm int}K_{\rm H}$, define

$$\mathcal{P}(x_{\rm H}, y_{\rm H}) := \{p \in {\rm int}K_{\rm H} \mid Q_p x_{\rm H} \text{ and } Q_{p^{-1}} y_{\rm H} \text{ operator commute.}\}$$

and

$$\tilde{x}_{\rm H} := Q_p x_{\rm H}, \quad \tilde{y}_{\rm H} := Q_{p^{-1}} y_{\rm H} = Q_p^{-1} y_{\rm H}. \tag{1.6}$$

The commutative class of search directions $(\Delta x_{\rm H}, \Delta y_{\rm H})$ are given by

$$(\Delta x_{\rm H}, \Delta y_{\rm H}) := (Q_p^{-1} \widetilde{\Delta x}_{\rm H}, Q_p \widetilde{\Delta y}_{\rm H}) \tag{1.7}$$

where $(\widetilde{\Delta x}_{\rm H}, \widetilde{\Delta y}_{\rm H})$ is the solution of the scaled Newton system

$$\begin{cases} \widetilde{\Delta x}_{\rm H} \circ_{\rm H} \tilde{y}_{\rm H} + \tilde{x}_{\rm H} \circ_{\rm H} \widetilde{\Delta y}_{\rm H} = \gamma \mu_{\rm H} e - \tilde{x}_{\rm H} \circ_{\rm H} \tilde{y}_{\rm H}, \\ \widetilde{\Delta y}_{\rm H} - D\widetilde{\psi}_{\rm H}(\tilde{x}_{\rm H}) \widetilde{\Delta x}_{\rm H} = -\eta \tilde{s}_{\rm H} \end{cases} \tag{1.8}$$

where

$$\mu_{\rm H} = \langle \tilde{x}_{\rm H}, \tilde{y}_{\rm H}\rangle_{\rm H}/(r+1),\ \tilde{s}_{\rm H} := \tilde{y}_{\rm H} - \widetilde{\psi}_{\rm H}(\tilde{x}_{\rm H}),\ \widetilde{\psi}_{\rm H}(\tilde{x}_{\rm H}) := Q_p^{-1} \bullet \psi_{\rm H} \bullet Q_p^{-1}(\tilde{x}_{\rm H}) = Q_p^{-1}\psi_{\rm H}(x_{\rm H}) \tag{1.9}$$

for some $\gamma, \eta \in [0,1]$ and $p \in \mathcal{P}(x_{\rm H}, y_{\rm H})$. Here, $\phi_1 \bullet \phi_2$ denotes the composite function of ϕ_1 and ϕ_2. If we choose $p = y_{\rm H}^{1/2}$ ($p = x_{\rm H}^{-1/2}$), then $\tilde{y}_{\rm H} = e_{\rm H}$ ($\tilde{x}_{\rm H} = e_{\rm H}$) and we see that $p \in \mathcal{P}(x_{\rm H}, y_{\rm H})$. We call the method xy-method (yx-method). If we choose $p \in \mathcal{P}(x_{\rm H}, y_{\rm H})$ so that $\tilde{x}_{\rm H} = \tilde{y}_{\rm H}$, then the method is called Nesterov-Todd (NT) method.

The next iterate is determined by moving along the following one-dimensional curve $(x_{\rm H}(\alpha), y_{\rm H}(\alpha))$:

$$\left\{ \begin{array}{rcl} x_{\rm H}(\alpha) & := & x_{\rm H} + \alpha \Delta x_{\rm H}, \\ y_{\rm H}(\alpha) & := & y_{\rm H} + \alpha \Delta y_{\rm H} + \psi_{\rm H}(x_{\rm H}(\alpha)) - \psi_{\rm H}(x_{\rm H}) - \alpha D\psi_{\rm H}(x_{\rm H})\Delta x_{\rm H} \\ & = & \psi_{\rm H}(x_{\rm H}(\alpha)) + (1-\alpha\eta)s_{\rm H} \end{array} \right. \tag{1.10}$$

where the last equation follows from the first equation of (1.4) and the definition (1.5) of $s_{\rm H}$. Note that the curve search technique was first introduced by Monteiro and Adler [15], and then used in many literatures (cf. [14,19,25]) for solving nonlinear programs over nonnegative orthants.

We consider the following three types of neighborhood:

$$\begin{array}{l} \mathcal{N}_{\rm F}(\beta) := \{(x_{\rm H}, y_{\rm H}) \in K_{\rm H} \times K_{\rm H} \mid d_{\rm F}(x_{\rm H}, y_{\rm H}) \le \beta\mu_{\rm H}\}, \\ \mathcal{N}_2(\beta) := \{(x_{\rm H}, y_{\rm H}) \in K_{\rm H} \times K_{\rm H} \mid d_2(x_{\rm H}, y_{\rm H}) \le \beta\mu_{\rm H}\}, \\ \mathcal{N}_{-\infty}(\beta) := \{(x_{\rm H}, y_{\rm H}) \in K_{\rm H} \times K_{\rm H} \mid d_{-\infty}(x_{\rm H}, y_{\rm H}) \le \beta\mu_{\rm H}\} \end{array} \tag{1.11}$$

where $\beta \in (0,1)$, $w_{\rm H} = Q_{x_{\rm H}^{1/2}} y_{\rm H}$ and

$$\begin{array}{rcl} d_{\rm F}(x_{\rm H}, y_{\rm H}): & = & \|Q_{x_{\rm H}^{1/2}} y_{\rm H} - \mu_{\rm H} e_{\rm H}\|_{\rm F} = \sqrt{\displaystyle\sum_{i=1}^{r+1} (\lambda_i(w_{\rm H}) - \mu_{\rm H})^2}, \\ d_2(x_{\rm H}, y_{\rm H}): & = & \|Q_{x_{\rm H}^{1/2}} y_{\rm H} - \mu_{\rm H} e_{\rm H}\|_2 \\ & = & \max\{|\lambda_i(w_{\rm H}) - \mu_{\rm H}| \ \ (i = 1, \ldots, r+1)\} \\ & = & \max\{\lambda_{\max}(w_{\rm H}) - \mu_{\rm H},\ \mu_{\rm H} - \lambda_{\min}(w_{\rm H})\}, \\ d_{-\infty}(x_{\rm H}, y_{\rm H}): & = & \mu_{\rm H} - \lambda_{\min}(w_{\rm H}). \end{array}$$

Since the inclusive relation $\mathcal{N}_{\rm F}(\beta) \subseteq \mathcal{N}_2(\beta) \subseteq \mathcal{N}_{-\infty}(\beta)$ holds for any $\beta \in (0,1)$ (cf. Proposition 29 of [22]), we call the algorithms using $\mathcal{N}_{\rm F}(\beta)$, $\mathcal{N}_2(\beta)$ and $\mathcal{N}_{-\infty}(\beta)$ the *short-step* algorithm, the *semi-long-step* algorithm and the *long-step* algorithm, respectively.

By exploring the behavior of the curve (1.10), we derive complexity bounds of six path-following algorithms, the combinations of two types of search directions and three types of neighborhoods. For this purpose, we introduce a parameter $\theta \ge 0$ for quantifying a scaled Lipschitz property of the function ψ. We impose the following assumption on ψ, which can be considered as an extension of scaled Lipschitz properties (cf. [3, 14, 15, 19, 25]):

Assumption 1.2. *There exists a $\theta \ge 0$ such that*

$$\left\| \tilde{z}(\alpha) \circ \left(\widetilde{\psi}(\tilde{z}(\alpha)) - \widetilde{\psi}(\tilde{z}) - \alpha D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z} \right) \right\|_{\rm F} \le \alpha^2 \theta \langle \widetilde{\Delta z}, D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z} \rangle$$

for all $z \in {\rm int}K$, $\Delta z \in V$, $p \in \mathcal{P}(x,y)$ and $\alpha \in [0,1]$ such that $z(\alpha) \in {\rm int}K$, where

$$\tilde{z}(\alpha) = Q_p(z + \alpha\Delta z),\ \widetilde{\psi}(\tilde{z}) = Q_p^{-1} \bullet \psi \bullet Q_p^{-1}(\tilde{z}) = Q_p^{-1}\psi(z).$$

Here, $\phi_1 \bullet \phi_2$ denotes the composite function of ϕ_1 and ϕ_2.

Obviously, if ψ is affine then ψ satisfies the assumption with $\theta = 0$. It should be noted that the assumption is not introduced to intend an empty generalization of affine functions, while the assumption has high affinity with the homogeneous function defined in (1.1). Let us consider a simple example, $\phi : \Re \to \Re$, $\phi(x) = x$. Then the induced homogeneous function $\psi_{\rm H} : \Re \times \Re_{++} \to \Re^2$ is given by $\psi_{\rm H}(x, \tau) = (x, -x^2/\tau)$. The function $\psi_{\rm H}$ is no longer linear, but we can see that $\psi_{\rm H}$ is monotone (cf. (iv) of Theorem 2.2) and satisfies Assumption 1.2 with $\theta = 1$ (cf. Theorem 5.1). On the other hand, unfortunately, the function $\psi(x_1, x_2) = (x_1, -x_1^2/x_2)$ does not satisfy the Lipschitz condition on $\mathrm{int}K := \Re^2_{++}$ as imposed in (iv) of Theorem 1.1. A further issue may be to find a monotone function ψ which is not affine but satisfies both of Assumption 1.2 and the Lipschitz condition.

Under the assumption, we obtain the following results instead of (e) above, by analogous discussions as in [22] and in [20]:

(e1) The (infeasible) NT method takes $O(\sqrt{r}(1+\sqrt{r}\,\theta)\log\epsilon^{-1})$ iterations for the short-step, and $O(r(1+\sqrt{r}\,\theta)\log\epsilon^{-1})$ iterations for the semi-long- and long-step variants.

(e2) The (infeasible) xy method or yx method takes $O(\sqrt{r}(1+\sqrt{r}\,\theta)\log\epsilon^{-1})$ iterations for the short-step, $O(r(1+\sqrt{r}\,\theta)\log\epsilon^{-1})$ iterations for the semi-long-step, and $O(r^{1.5}(1+\sqrt{r}\,\theta)\log\epsilon^{-1})$ iterations for the long-step variant.

Since $\theta = 0$ for any affine function ψ, the above results achieve the best complexity bounds for linear or convex quadratic optimization problems over symmetric cones.

The paper is organized as follows.

In Section 2, we first observe some basic properties of the homogeneous function $\psi_{\rm H}$ defined by (1.1). In Section 3, we discuss the existence of the scaled Newton direction for the HCP and explore the behavior of the search curve (1.10). More precise analyses of the curve are carried out in Sections 4 and 5. The number of iterations of our algorithms depends on the range of step sizes for which the next iterate stays in the neighborhoods (1.11). We determine the range in Section 6 using the results in Sections 4 and 5. After providing a detailed description of the algorithms, we give the iteration-complexity bounds of our algorithms in Section 7. Some conclusions are drawn in Section 8.

2 Properties of the Function $\psi_{\rm H}$

In this section, we provide some key properties of the function $\psi_{\rm H}$ defined by (1.1), most of which are obtained by simple calculations.

First, we introduce a well-know result of differentiable monotone functions, which is helpful in exploring the monotonicity of $\psi_{\rm H}$.

Lemma 2.1 (Proposition 2.3.2 (a) of [4]). *If ψ is monotone on* $\mathrm{int}K$ *then the Jacobian $D\psi(x)$ is positive semidefinite for any $x \in \mathrm{int}K$ with respect to $\langle\cdot,\cdot\rangle$.*

By a simple calculation, the Jacobian $D\psi_{\rm H}(x_{\rm H})$ of the function $\psi_{\rm H}$ at $x_{\rm H}$ is given by

$$D\psi_{\rm H}(x_{\rm H}) = \begin{pmatrix} D\psi(x/\tau) & \psi(x/\tau) - D\psi(x/\tau)(x/\tau) \\ -\psi(x/\tau)^T - [D\psi(x/\tau)^*(x/\tau)]^T & \langle D\psi(x/\tau)(x/\tau), (x/\tau)\rangle \end{pmatrix} \tag{2.1}$$

where A^* is the adjoint of a linear operator A, i.e.,

$$\langle Ax, y\rangle = \langle x, A^*y\rangle, \quad \forall\ x, y \in V$$

and a^T is a linear operator such that

$$a^T x = \langle a, x\rangle, \quad \forall\, x \in V.$$

By (1.2) and the definitions of A^* and a^T above, we see that

$$\left\langle \begin{pmatrix} y \\ \kappa \end{pmatrix}, \begin{pmatrix} A & b \\ c^T & d \end{pmatrix}\begin{pmatrix} x \\ \tau \end{pmatrix} \right\rangle_{\text{H}} = \left\langle \begin{pmatrix} A^* & c \\ b^T & d \end{pmatrix}\begin{pmatrix} y \\ \kappa \end{pmatrix}, \begin{pmatrix} x \\ \tau \end{pmatrix} \right\rangle_{\text{H}}.$$

This yields

$$D\psi_{\text{H}}(x_{\text{H}})^* = \begin{pmatrix} D\psi(x/\tau)^* & -\psi(x/\tau) - D\psi(x/\tau)^*(x/\tau) \\ \psi(x/\tau)^T - [D\psi(x/\tau)^*(x/\tau)]^T & \langle D\psi(x/\tau)(x/\tau), (x/\tau)\rangle \end{pmatrix}. \tag{2.2}$$

The function ψ_{H} and its Jacobian $D\psi_{\text{H}}$ have the following properties.

Lemma 2.2. **(i)** $\langle x_{\text{H}}, \psi_{\text{H}}(x_{\text{H}})\rangle_{\text{H}} = 0.$

(ii) $D\psi_{\text{H}}(x_{\text{H}})^* x_{\text{H}} = -\psi_{\text{H}}(x_{\text{H}}).$

(iii)

$$D\psi_{\text{H}}(x_{\text{H}})\Delta x_{\text{H}} \quad = \quad \begin{pmatrix} D\psi(x/\tau)\Delta x + [\psi(x/\tau) - D\psi(x/\tau)(x/\tau)]\,\Delta\tau \\ -\langle \psi(x/\tau), \Delta x\rangle - \langle x/\tau, D\psi(z)(\Delta x - (x/\tau)\Delta\tau)\rangle \end{pmatrix}.$$

(iv)

$$\langle \Delta x_{\text{H}}, D\psi_{\text{H}}(x_{\text{H}})\Delta x_{\text{H}}\rangle_{\text{H}} = \langle \Delta x - (x/\tau)\Delta\tau, D\psi(x/\tau)(\Delta x - (x/\tau)\Delta\tau)\rangle \geq 0,$$

i.e., $D\psi_{\text{H}}(x_{\text{H}})$ *is positive semidefinite on* $\text{int}K_{\text{H}}$ *with respect to* $\langle\cdot,\cdot\rangle_{\text{H}}$*. Therefore,* ψ_{H} *is monotone on* $\text{int}K_{\text{H}}$ *with respect to* $\langle\cdot,\cdot\rangle_{\text{H}}$.

Proof. (i): The proof is straightforward.
(ii), (iii), (iv): Using (2.1) and (2.2), we can calculate them as follows:

$$\begin{aligned} D\psi_{\text{H}}(x_{\text{H}})^* x_{\text{H}} &= \begin{pmatrix} D\psi(x/\tau)^* & -\psi(x/\tau) - D\psi(x/\tau)^*(x/\tau) \\ \psi(x/\tau)^T - [D\psi(x/\tau)^*(x/\tau)]^T & \langle D\psi(x/\tau)(x/\tau), (x/\tau)\rangle \end{pmatrix}\begin{pmatrix} x \\ \tau \end{pmatrix} \\ &= \begin{pmatrix} D\psi(x/\tau)^* x - \tau\psi(x/\tau) - \tau D\psi(x/\tau)^*(x/\tau) \\ \langle \psi(x/\tau), x\rangle - \langle D\psi(x/\tau)(x/\tau), x\rangle + \tau\langle D\psi(x/\tau)(x/\tau), (x/\tau)\rangle \end{pmatrix} \\ &= \begin{pmatrix} -\tau\psi(x/\tau) \\ \langle \psi(x/\tau), x\rangle \end{pmatrix} \\ &= -\psi_{\text{H}}(x_{\text{H}}), \end{aligned}$$

$$\begin{aligned} D\psi_{\text{H}}(x_{\text{H}})\Delta x_{\text{H}} &= \begin{pmatrix} D\psi(x/\tau) & \psi(x/\tau) - D\psi(x/\tau)(x/\tau) \\ -\psi(x/\tau)^T - [D\psi(x/\tau)^*(x/\tau)]^T & \langle D\psi(x/\tau)(x/\tau), (x/\tau)\rangle \end{pmatrix}\begin{pmatrix} \Delta x \\ \Delta\tau \end{pmatrix} \\ &= \begin{pmatrix} D\psi(x/\tau)\Delta x + \Delta\tau\,[\psi(x/\tau) - D\psi(x/\tau)(x/\tau)] \\ -\langle \psi(x/\tau), \Delta x\rangle - \langle D\psi(x/\tau)^*(x/\tau), \Delta x\rangle + \Delta\tau\langle D\psi(x/\tau)(x/\tau), (x/\tau)\rangle \end{pmatrix} \\ &= \begin{pmatrix} D\psi(x/\tau)(\Delta x - (x/\tau)\Delta\tau) + \Delta\tau\psi(x/\tau) \\ -\langle \psi(x/\tau), \Delta x\rangle - \langle (x/\tau), D\psi(x/\tau)\Delta x\rangle + \Delta\tau\langle (x/\tau), D\psi(x/\tau)(x/\tau)\rangle \end{pmatrix} \\ &= \begin{pmatrix} D\psi(x/\tau)(\Delta x - (x/\tau)\Delta\tau) + \Delta\tau\psi(x/\tau) \\ -\langle \psi(x/\tau), \Delta x\rangle - \langle (x/\tau), D\psi(x/\tau)(\Delta x - (x/\tau)\Delta\tau)\rangle \end{pmatrix} \qquad (2.3) \\ &= \begin{pmatrix} D\psi(x/\tau)\Delta x + [\psi(x/\tau) - D\psi(x/\tau)(x/\tau)]\,\Delta\tau \\ -\langle \psi(x/\tau), \Delta x\rangle - \langle x/\tau, D\psi(z)(\Delta x - (x/\tau)\Delta\tau)\rangle \end{pmatrix}, \end{aligned}$$

$$\begin{aligned}
&\langle \Delta x_{\mathrm{H}}, D\psi_{\mathrm{H}}(x_{\mathrm{H}})\Delta x_{\mathrm{H}}\rangle_{\mathrm{H}} \\
&= \left\langle \begin{pmatrix} \Delta x \\ \Delta\tau \end{pmatrix}, \begin{pmatrix} D\psi(x/\tau)(\Delta x - (x/\tau)\Delta\tau) + \Delta\tau\psi(x/\tau) \\ -\langle\psi(x/\tau), \Delta x\rangle - \langle (x/\tau), D\psi(x/\tau)(\Delta x - (x/\tau)\Delta\tau)\rangle \end{pmatrix} \right\rangle_{\mathrm{H}} \\
&\qquad \text{(by (2.3))} \\
&= \langle \Delta x, D\psi(x/\tau)\,(\Delta x - (x/\tau)\Delta\tau) + \Delta\tau\psi(x/\tau)\rangle \\
&\qquad +\Delta\tau\,[-\langle\psi(x/\tau), \Delta x\rangle - \langle (x/\tau), D\psi(x/\tau)(\Delta x - (x/\tau)\Delta\tau)\rangle] \\
&= \langle \Delta x, D\psi(x/\tau)\,(\Delta x - (x/\tau)\Delta\tau)\rangle + \Delta\tau\langle\Delta x, \psi(x/\tau)\rangle \\
&\qquad -\Delta\tau\langle\psi(x/\tau), \Delta x\rangle - \Delta\tau\langle (x/\tau), D\psi(x/\tau)\,(\Delta x - (x/\tau)\Delta\tau)\rangle \\
&= \langle \Delta x - (x/\tau)\Delta\tau, D\psi(x/\tau)\,(\Delta x - (x/\tau)\Delta\tau)\rangle \\
&\geq 0
\end{aligned}$$

where the last inequality follows from the monotonicity of ψ and Lemma 2.1. □

3 Scaled Newton Directions and Search Curves

In this section, we show the existence of the scaled Newton direction for the HCP and explore the behavior of the search curve (1.10). For simplicity, we use the symbols $x, y, \psi, \langle\cdot,\cdot\rangle$ to denote $x_{\mathrm{H}}, y_{\mathrm{H}}, \psi_{\mathrm{H}}, \langle\cdot,\cdot\rangle_{\mathrm{H}}$ throughout this section.

The following lemma shows the invariance of the complementarity under the scaling Q_p for $p \in \mathcal{P}(x,y)$.

Lemma 3.1 (Lemma 28 of [22]). *Suppose that $p \in V$ is invertible. Then for any $x, y \in K$, $x \circ y = \mu e$ if and only if $Q_p x \circ Q_p^{-1} y = \mu e$.*

The following lemma shows that the properties in Lemma 2.2 are also invariant under the scaling Q_p for $p \in \mathcal{P}(x,y)$.

Lemma 3.2. *Suppose that $p \in \mathrm{int}K$*

(i) *$\widetilde{\psi}$ is monotone on* $\mathrm{int}K$.

(ii) $\langle \tilde{x}, \widetilde{\psi}(\tilde{x})\rangle = 0$.

(iii) $D\widetilde{\psi}(\tilde{x})^*\tilde{x} = -\widetilde{\psi}(\tilde{x})$.

(iv) $\langle \widetilde{\Delta x}, D\widetilde{\psi}(x)\widetilde{\Delta x}\rangle = \langle \Delta x, D\psi(x)\Delta x\rangle$.

Proof. (i): For any $\tilde{x}, \tilde{x}' \in \mathrm{int}K$, by the self-adjointness of Q_p^{-1}, we can see that

$$\begin{aligned}
\langle \tilde{x} - \tilde{x}', \widetilde{\psi}(\tilde{x}) - \widetilde{\psi}(\tilde{x}')\rangle &= \langle \tilde{x} - \tilde{x}', Q_p^{-1} \bullet \psi \bullet Q_p^{-1}(\tilde{x}) - Q_p^{-1} \bullet \psi \bullet Q_p^{-1}(\tilde{x}')\rangle \\
&= \langle Q_p^{-1}\tilde{x} - Q_p^{-1}\tilde{x}', \psi(Q_p^{-1}\tilde{x}) - \psi(Q_p^{-1}\tilde{x}')\rangle \\
&\geq 0
\end{aligned}$$

where the last inequality follows from $Q_p^{-1}\tilde{x}, Q_p^{-1}\tilde{x}' \in \mathrm{int}K$ (cf. Proposition 18 of [22]) and the monotonicity of ψ on $\mathrm{int}K$.
(ii), (iii), (iv): The following equations follow from Lemma 2.2, the definitions (1.6) – (1.9), and the self-adjointness of Q_p and Q_p^{-1}:

$$\begin{aligned}
\langle \tilde{x}, \widetilde{\psi}(\tilde{x})\rangle &= \langle Q_p x, Q_p^{-1}\psi(x)\rangle \\
&= \langle x, \psi(x)\rangle \\
&= 0 \qquad \text{(by (i) of Lemma 2.2)},
\end{aligned}$$

$$
\begin{aligned}
D\widetilde{\psi}(\tilde{x})^*\tilde{x} &= \left[D\left(Q_p^{-1}\psi(Q_p^{-1}\tilde{x})\right)\right]^*\tilde{x} \\
&= \left[Q_p^{-1}D\left(\psi(Q_p^{-1}\tilde{x})\right)\right]^*\tilde{x} \\
&= \left[Q_p^{-1}D\psi(Q_p^{-1}\tilde{x})Q_p^{-1}\right]^*\tilde{x} \\
&= Q_p^{-1}D\psi(Q_p^{-1}\tilde{x})^*Q_p^{-1}\tilde{x} \\
&= Q_p^{-1}D\psi(x)^*x \\
&= Q_p^{-1}(-\psi(x)) \qquad \text{(by (ii) of Lemma 2.2)} \\
&= -\widetilde{\psi}(\tilde{x}),
\end{aligned}
$$

$$
\begin{aligned}
\langle\widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle &= \langle Q_p\Delta x, Q_p^{-1}D\psi(Q_p^{-1}\tilde{x})Q_p^{-1}Q_p\Delta x\rangle \\
&= \langle\Delta x, D\psi(x)\Delta x\rangle.
\end{aligned}
$$

□

Now we show the existence of the scaled Newton direction for the system (1.3). For a $p \in \mathcal{P}(x,y)$, the scaled system of (1.3) is given by (1.8).

Lemma 3.3. *The system (1.8) has a unique solution* $(\widetilde{\Delta x}, \widetilde{\Delta y}) := (Q_p\Delta x, Q_p^{-1}\Delta y)$

Proof. The system (1.8) is equivalently represented by

$$
\begin{cases}
L(\tilde{y})\widetilde{\Delta x} + L(\tilde{x})\widetilde{\Delta y} = \gamma\mu e - \tilde{x}\circ\tilde{y}, \\
\widetilde{\Delta y} - D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x} = -\eta\tilde{s}.
\end{cases}
$$

Since the above system consists of $2(n+1)$ linear equations, the system has a unique solution if and only if

$$
\begin{cases}
L(\tilde{y})\widetilde{\Delta x} + L(\tilde{x})\widetilde{\Delta y} = 0, \\
\widetilde{\Delta y} - D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x} = 0
\end{cases}
\quad\Longrightarrow\quad (\widetilde{\Delta x}, \widetilde{\Delta y}) = (0,0) \tag{3.1}
$$

holds. As we have seen in Lemma 2.2, $\widetilde{\psi}$ is monotone on $\mathrm{int}K$. Therefore, $\widetilde{\Delta y} - D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x} = 0$ implies that $\langle\widetilde{\Delta x}, \widetilde{\Delta y}\rangle \geq 0$. Suppose that $(\widetilde{\Delta x}, \widetilde{\Delta y})$ satisfies the left-hand side of (3.1) but $(\widetilde{\Delta x}, \widetilde{\Delta y}) \neq (0,0)$. Obviously, we have $\widetilde{\Delta x} \neq 0$. Since $\tilde{x} \in \mathrm{int}K$ and $\tilde{y} \in \mathrm{int}K$ operator commute, there exists a Jordan frame $\{c_1, \cdots, c_{r+1}\}$, $\tilde{x}$ and $\tilde{y}$ are given by

$$
\tilde{x} = \sum_{i=1}^{r+1}\lambda_i c_i, \quad \tilde{y} = \sum_{i=1}^{r+1}\mu_i c_i
$$

for any $\lambda_i > 0, \mu_i > 0$ $(i = 1, \ldots, r+1)$ (cf. Theorem 27 of [22]). Note that $\tilde{x}^{-1} = \sum_{i=1}^{r+1}\frac{1}{\lambda_i}c_i$ which implies that $\tilde{x}^{-1}$ and $\tilde{y}$ operator commute, too. This yields

$$
L(\tilde{x}^{-1})L(\tilde{y}) = L(\tilde{y})L(\tilde{x}^{-1})
$$

and, since $L(\tilde{x}^{-1})$ and $L(\tilde{y})$ are positive definite, $L(\tilde{x}^{-1})L(\tilde{y})$ is also positive definite. Therefore, the implication

$$
\begin{aligned}
L(\tilde{y})\widetilde{\Delta x} + L(\tilde{x})\widetilde{\Delta y} = 0 \quad &\Rightarrow \quad L(\tilde{x})^{-1}L(\tilde{y})\widetilde{\Delta x} + \widetilde{\Delta y} = 0 \\
&\Rightarrow \quad L(\tilde{x}^{-1})L(\tilde{y})\widetilde{\Delta x} + \widetilde{\Delta y} = 0 \\
&\Rightarrow \quad \langle\widetilde{\Delta x}, L(\tilde{x}^{-1})L(\tilde{y})\widetilde{\Delta x}\rangle + \langle\widetilde{\Delta x}, \widetilde{\Delta y}\rangle = 0
\end{aligned}
$$

holds which contradicts to the facts

$$\langle \widetilde{\Delta x}, L(\tilde{x}^{-1})L(\tilde{y})\widetilde{\Delta x}\rangle > 0, \quad \langle \widetilde{\Delta x}, \widetilde{\Delta y}\rangle \geq 0.$$

Consequently, we have $\widetilde{\Delta x} = 0$ and $\widetilde{\Delta y} = 0$. □

The following lemma shows that $\widetilde{\Delta y}$ and $D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}$ are very similar in the sense that $\langle \widetilde{\Delta x}, \widetilde{\Delta y}\rangle = \langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle$ holds if we set $1 - \eta - \gamma = 0$.

Lemma 3.4. *$\langle \widetilde{\Delta x}, \widetilde{\Delta y}\rangle = \langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle + \eta(1 - \eta - \gamma)(r + 1)\mu$.*

Proof. It follows from the second equation in (1.8) that

$$\langle \widetilde{\Delta x}, \widetilde{\Delta y}\rangle - \langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle = -\eta\langle \widetilde{\Delta x}, \tilde{s}\rangle = -\eta\langle \widetilde{\Delta x}, \tilde{y} - \widetilde{\psi}(\tilde{x})\rangle, \tag{3.2}$$

and

$$\langle \tilde{x}, \widetilde{\Delta y}\rangle - \langle \tilde{x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle = -\eta\langle \tilde{x}, \tilde{s}\rangle = -\eta\langle \tilde{x}, \tilde{y} - \widetilde{\psi}(\tilde{x})\rangle = -\eta\langle \tilde{x}, \tilde{y}\rangle = -\eta(r + 1)\mu, \tag{3.3}$$

where the second last equation follows from (ii) of Lemma 3.2. By (iii) of Lemma 3.2, the left-hand side above becomes

$$\langle \tilde{x}, \widetilde{\Delta y}\rangle - \langle \tilde{x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle = \langle \tilde{x}, \widetilde{\Delta y}\rangle - \langle D\widetilde{\psi}(\tilde{x})^*\tilde{x}, \widetilde{\Delta x}\rangle = \langle \tilde{x}, \widetilde{\Delta y}\rangle + \langle \widetilde{\psi}(\tilde{x}), \widetilde{\Delta x}\rangle, \tag{3.4}$$

and by (3.3) and (3.4), we have

$$\langle \widetilde{\Delta x}, \widetilde{\psi}(\tilde{x})\rangle = -\eta(r + 1)\mu - \langle \tilde{x}, \widetilde{\Delta y}\rangle. \tag{3.5}$$

Combining (3.2) and (3.5), this yields

$$\begin{aligned}
\langle \widetilde{\Delta x}, \widetilde{\Delta y}\rangle &= \langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle - \eta\langle \widetilde{\Delta x}, \tilde{y} - \widetilde{\psi}(\tilde{x})\rangle \\
&= \langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle - \eta\langle \widetilde{\Delta x}, \tilde{y}\rangle + \eta\langle \widetilde{\Delta x}, \widetilde{\psi}(\tilde{x})\rangle \\
&= \langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle - \eta\langle \widetilde{\Delta x}, \tilde{y}\rangle + \eta\left\{-\eta(r + 1)\mu - \langle \tilde{x}, \widetilde{\Delta y}\rangle\right\} \\
&= \langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle - \eta\left\{\langle \widetilde{\Delta x}, \tilde{y}\rangle + \langle \tilde{x}, \widetilde{\Delta y}\rangle + \eta(r + 1)\mu\right\}.
\end{aligned} \tag{3.6}$$

By using the first equation in (1.8) and the fact $\langle \tilde{x}, \tilde{y}\rangle = \langle x, y\rangle = (r + 1)\mu$, we have

$$\begin{aligned}
\langle \widetilde{\Delta x}, \tilde{y}\rangle + \langle \tilde{x}, \widetilde{\Delta y}\rangle &= \langle e, \widetilde{\Delta x} \circ \tilde{y}\rangle + \langle e, \tilde{x} \circ \widetilde{\Delta y}\rangle \\
&= \langle e, \widetilde{\Delta x} \circ \tilde{y} + \tilde{x} \circ \widetilde{\Delta y}\rangle \\
&= \langle e, \gamma\mu e - \tilde{x} \circ \tilde{y}\rangle \\
&= \gamma\mu(r + 1) - (r + 1)\mu \\
&= (\gamma - 1)\mu(r + 1).
\end{aligned} \tag{3.7}$$

Substituting (3.7) into (3.6) leads to

$$\begin{aligned}
\langle \widetilde{\Delta x}, \widetilde{\Delta y}\rangle &= \langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle - \eta\{(\gamma - 1)\mu(r + 1) + \eta(r + 1)\mu\} \\
&= \langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle + \eta\mu(r + 1)(1 - \eta - \gamma).
\end{aligned}$$

□

In our algorithms, the next iterate is determined by moving along the curve $(x(\alpha), y(\alpha))$ defined by (1.10). Define the scaled curve $(\tilde{x}_{\mathrm{H}}(\alpha), \tilde{y}_{\mathrm{H}}(\alpha))$ as follows:

$$\left\{\begin{array}{lcl} \tilde{x}(\alpha) & := & \tilde{x} + \alpha\widetilde{\Delta x}, \\ \tilde{y}(\alpha) & := & \tilde{y} + \alpha\widetilde{\Delta y} + \widetilde{\psi}(\tilde{x}(\alpha)) - \widetilde{\psi}(\tilde{x}) - \alpha D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x} \\ & = & \widetilde{\psi}(\tilde{x}(\alpha)) + (1-\alpha\eta)\tilde{s}. \end{array}\right. \tag{3.8}$$

The following lemma shows that if we set $1-\eta-\gamma=0$ then the inner product $\langle \tilde{x}(\alpha), \tilde{y}(\alpha)\rangle$ is reducing linearly with $\alpha > 0$.

Lemma 3.5. *Define*

$$\tilde{s}(\alpha) := \tilde{y}(\alpha) - \widetilde{\psi}(\tilde{x}(\alpha)).$$

(i) $\tilde{s}(\alpha) = (1-\alpha\eta)\tilde{s}$

(ii) $\langle \tilde{x}(\alpha), \tilde{y}(\alpha)\rangle = \{1-\alpha(1-\gamma)\}\langle \tilde{x}, \tilde{y}\rangle + \alpha^2\eta(1-\eta-\gamma)(r+1)\mu$

Proof. (i): It is straightforward from (3.8).
(ii): It follows from Lemma 3.2 and 3.4 that

$$\begin{array}{ll} \multicolumn{2}{l}{\langle \tilde{x}(\alpha), \tilde{y}(\alpha)\rangle} \\ = & \langle \tilde{x}(\alpha), \tilde{y} + \alpha\widetilde{\Delta y} + \widetilde{\psi}(\tilde{x}(\alpha)) - \widetilde{\psi}(\tilde{x}) - \alpha D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle \\ = & \langle \tilde{x}(\alpha), \tilde{y} + \alpha\widetilde{\Delta y}\rangle + \langle \tilde{x}(\alpha), \widetilde{\psi}(\tilde{x}(\alpha))\rangle + \langle \tilde{x}(\alpha), -\widetilde{\psi}(\tilde{x}) - \alpha D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle \\ = & \langle \tilde{x}(\alpha), \tilde{y} + \alpha\widetilde{\Delta y}\rangle + 0 - \langle \tilde{x}(\alpha), \widetilde{\psi}(\tilde{x}) + \alpha D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle \\ & \qquad \text{(by (ii) of Lemma 3.2)} \\ = & \langle \tilde{x}(\alpha), \tilde{y} + \alpha\widetilde{\Delta y}\rangle - \langle \tilde{x} + \alpha\widetilde{\Delta x}, \widetilde{\psi}(\tilde{x}) + \alpha D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle \\ = & \langle \tilde{x}(\alpha), \tilde{y} + \alpha\widetilde{\Delta y}\rangle - \langle \tilde{x}, \widetilde{\psi}(\tilde{x})\rangle - \alpha\langle D\widetilde{\psi}(\tilde{x})^*\tilde{x}, \widetilde{\Delta x}\rangle - \alpha\langle \widetilde{\Delta x}, \widetilde{\psi}(\tilde{x})\rangle - \alpha^2\langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle \\ = & \langle \tilde{x}(\alpha), \tilde{y} + \alpha\widetilde{\Delta y}\rangle - 0 + \alpha\langle \widetilde{\psi}(\tilde{x}), \widetilde{\Delta x}\rangle - \alpha\langle \widetilde{\Delta x}, \widetilde{\psi}(\tilde{x})\rangle - \alpha^2\langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle \\ & \qquad \text{(by (ii) and (iii) of Lemma 3.2)} \\ = & \langle \tilde{x}(\alpha), \tilde{y} + \alpha\widetilde{\Delta y}\rangle - \alpha^2\langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle \\ = & \langle \tilde{x} + \alpha\widetilde{\Delta x}, \tilde{y} + \alpha\widetilde{\Delta y}\rangle - \alpha^2\langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle \\ = & \langle \tilde{x}, \tilde{y}\rangle + \alpha\left(\langle \widetilde{\Delta x}, \tilde{y}\rangle + \langle \tilde{x}, \widetilde{\Delta y}\rangle\right) + \alpha^2\left(\langle \widetilde{\Delta x}, \widetilde{\Delta y}\rangle - \langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle\right) \\ = & \langle \tilde{x}, \tilde{y}\rangle + \alpha\langle e, \gamma\mu e - \tilde{x}\circ\tilde{y}\rangle + \alpha^2\eta(1-\eta-\gamma)(r+1)\mu \\ & \qquad \text{(by the first equation of (1.8) and Lemma 3.4)} \\ = & (r+1)\mu - \alpha(1-\gamma)(r+1)\mu + \alpha^2\eta(1-\eta-\gamma)(r+1)\mu \\ & \qquad \text{(by the fact } \langle \tilde{x}, \tilde{y}\rangle = \langle x, y\rangle = (r+1)\mu) \\ = & \{1-\alpha(1-\gamma)\}(r+1)\mu + \alpha^2\eta(1-\eta-\gamma)(r+1)\mu. \end{array}$$

□

We conclude this section by observing $\tilde{x}(\alpha)\circ\tilde{y}(\alpha)$.

Lemma 3.6.

$$\tilde{x}(\alpha)\circ\tilde{y}(\alpha) = (1-\alpha)\tilde{x}\circ\tilde{y} + \gamma\alpha\mu e + \alpha^2\widetilde{\Delta x}\circ\widetilde{\Delta y} + \tilde{x}(\alpha)\circ\tilde{d}(\alpha)$$

where

$$\begin{aligned}\tilde{d}(\alpha) &:= \widetilde{\psi}(\tilde{x}(\alpha)) - \widetilde{\psi}(\tilde{x}) - \alpha D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x} \\ &= Q_{p^{-1}}\left[\psi(x(\alpha)) - \psi(x) - \alpha D\psi(z)\Delta x\right]. \end{aligned} \tag{3.9}$$

Proof. By the definitions (3.8) and (3.9), we have

$$\begin{aligned}\tilde{x}(\alpha)\circ\tilde{y}(\alpha) &= (\tilde{x}+\alpha\widetilde{\Delta x})\circ(\tilde{y}+\alpha\widetilde{\Delta y}) + \tilde{x}(\alpha)\circ\tilde{d}(\alpha) \\ &= \tilde{x}\circ\tilde{y} + \alpha(\widetilde{\Delta x}\circ\tilde{y} + \tilde{x}\circ\widetilde{\Delta y}) + \alpha^2\widetilde{\Delta x}\circ\widetilde{\Delta y} + \tilde{x}(\alpha)\circ\tilde{d}(\alpha) \\ &= \tilde{x}\circ\tilde{y} + \alpha(\gamma\mu e - \tilde{x}\circ\tilde{y}) + \alpha^2\widetilde{\Delta x}\circ\widetilde{\Delta y} + \tilde{x}(\alpha)\circ\tilde{d}(\alpha) \\ &\qquad \text{(by the first equation of (1.8))} \\ &= (1-\alpha)\tilde{x}\circ\tilde{y} + \gamma\alpha\mu e + \alpha^2\widetilde{\Delta x}\circ\widetilde{\Delta y} + \tilde{x}(\alpha)\circ\tilde{d}(\alpha)\end{aligned}$$

□

The above lemma suggests that we may have to estimate the values of $\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}$ and $\|\tilde{x}(\alpha)\circ\tilde{d}(\alpha)\|_{\mathrm{F}}$ for the further discussion. Those bounds are derived in the succeeding two sections.

4 Upper Bounds of $\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}$

Similarly to the previous section, we use the symbols $x, y, \psi, \langle\cdot,\cdot\rangle$ to denote $x_{\mathrm{H}}, y_{\mathrm{H}}, \psi_{\mathrm{H}}, \langle\cdot,\cdot\rangle_{\mathrm{H}}$ throughout this section. The aim of this section is to derive upper bounds of $\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}$. The first step is given by the following lemma proposed in [22].

Lemma 4.1 (Lemma 33 of [22]). *Let $u, v \in V$ and let G be positive definite and self-adjoint linear operator with respect to $\langle\cdot,\cdot\rangle$. Then*

$$\|u\|_{\mathrm{F}}\|v\|_{\mathrm{F}} \le \frac{1}{2}\sqrt{K_G}\left(\|G^{1/2}u\|_{\mathrm{F}}^2 + \|G^{1/2}v\|_{\mathrm{F}}^2\right)$$

where K_G is the condition number of G defined by

$$K_G = \frac{\lambda_{\max}(G)}{\lambda_{\min}(G)}. \tag{4.1}$$

For the scaled iterate $(\tilde{x}, \tilde{y})$, the first equation of (1.8) is equivalently represented by

$$L(\tilde{y})\widetilde{\Delta x} + L(\tilde{x})\widetilde{\Delta y} = \gamma\mu e - L(\tilde{y})L(\tilde{x}). \tag{4.2}$$

Let $G = L(\tilde{y})^{-1}L(\tilde{x})$. Since $\tilde{x} \in \mathrm{int}K$ and $\tilde{y} \in \mathrm{int}K$ operator commute, we see that G is positive definite and self-adjoint linear operator with respect to $\langle\cdot,\cdot\rangle$. Multiplying (4.2) by $(L(\tilde{x})L(\tilde{y}))^{-1/2}$, we have

$$G^{1/2}\widetilde{\Delta x} + G^{1/2}\widetilde{\Delta y} = \gamma\mu\left(L(\tilde{x})L(\tilde{y})\right)^{-1/2}e - G^{1/2}\tilde{y}.$$

Using Lemma 4.1, we obtain the following result.

Lemma 4.2. *Suppose that* $(\widetilde{\Delta x}, \widetilde{\Delta y})$ *satisfies (1.8). Let us define*

$$G := L(\tilde{y})^{-1}L(\tilde{x}), \quad \tilde{h} := \gamma\mu(L(\tilde{x})L(\tilde{y}))^{-1/2}e - G^{1/2}\tilde{y}.$$

Then

$$\|G^{1/2}\widetilde{\Delta y}\|_{\mathrm{F}} + \|G^{1/2}\widetilde{\Delta x}\|_{\mathrm{F}} \le \|\tilde{h}\|_{\mathrm{F}}^2 - 2\eta(1-\eta-\gamma)(r+1)\mu.$$

Proof. Since G is self-adjoint, we see that

$$\begin{aligned}\|G^{1/2}\widetilde{\Delta y} + G^{1/2}\widetilde{\Delta x}\|_{\mathrm{F}}^2 &= \|G^{1/2}\widetilde{\Delta y}\|_{\mathrm{F}}^2 + \|G^{1/2}\widetilde{\Delta x}\|_{\mathrm{F}}^2 + 2\langle G^{1/2}\widetilde{\Delta y}, G^{-1/2}\widetilde{\Delta x}\rangle \\ &= \|G^{1/2}\widetilde{\Delta y}\|_{\mathrm{F}}^2 + \|G^{1/2}\widetilde{\Delta x}\|_{\mathrm{F}}^2 + 2\langle \widetilde{\Delta y}, \widetilde{\Delta x}\rangle.\end{aligned}$$

As we have seen above, by (1.8),

$$G^{1/2}\widetilde{\Delta y} + G^{1/2}\widetilde{\Delta x} = \gamma\mu(L(\tilde{x})L(\tilde{y}))^{-1/2}e - G^{1/2}\tilde{y} = \tilde{h}.$$

By Lemmas 3.4 and 3.2, we obtain the inequality

$$\begin{aligned}\|G^{1/2}\widetilde{\Delta y}\|_{\mathrm{F}} + \|G^{1/2}\widetilde{\Delta x}\|_{\mathrm{F}} &= \|G^{1/2}\widetilde{\Delta y} + G^{1/2}\widetilde{\Delta x}\|_{\mathrm{F}}^2 - 2\langle \widetilde{\Delta y}, \widetilde{\Delta x}\rangle \\ &= \|\tilde{h}\|_{\mathrm{F}}^2 - 2\langle \widetilde{\Delta y}, \widetilde{\Delta x}\rangle \\ &= \|\tilde{h}\|_{\mathrm{F}}^2 - 2\{\langle \widetilde{\Delta x}, D\widetilde{\psi}(\tilde{x})\widetilde{\Delta x}\rangle + \eta(1-\eta-\gamma)(r+1)\mu\} \\ &\qquad \text{(by Lemma 3.4)} \\ &\le \|\tilde{h}\|_{\mathrm{F}}^2 - 2\eta(1-\eta-\gamma)(r+1)\mu\} \\ &\qquad \text{(by (iv) of Lemma 3.2 and (iv) of Lemma 2.2).}\end{aligned}$$

□

Now we estimate the value of $\|\tilde{h}\|_{\mathrm{F}}$. We introduce the following three lemmas which are shown in [22]. Note that if we choose $p = y^{1/2}$ ($p = x^{-1/2}$) then we call the method xy-method (yx-method), and if we choose $p \in \mathcal{P}(x, y)$ so that $\tilde{x} = \tilde{y}$ then we call the method Nesterov-Todd (NT) method. In the latter case, the choice of p is unique and given by

$$p = \left[Q_{x^{1/2}}(Q_{x^{1/2}}y)^{-1/2}\right]^{-1/2} = \left[Q_{y^{-1/2}}(Q_{y^{-1/2}}x)^{1/2}\right]^{-1/2}.$$

Lemma 4.3 (Lemmas 34 (with its proof) and 35 of [22]). *Define* $G = L(\tilde{y})^{-1}L(\tilde{x})$ *and* $\tilde{w} = Q_{\tilde{x}^{1/2}}\tilde{y}$.

$$\|\tilde{h}\|_{\mathrm{F}}^2 = \sum_{i=1}^{r+1} \frac{(\gamma\mu - \lambda_i(\tilde{w}))^2}{\lambda_i(\tilde{w})}.$$

Lemma 4.4 (Lemma 35 of [22]). *Define* $G = L(\tilde{y})^{-1}L(\tilde{x})$ *and* $\tilde{w} = Q_{\tilde{x}^{1/2}}\tilde{y}$.

(i) *If* $(x, y) \in \mathcal{N}_{\mathrm{F}}(\beta)$ *then*

$$\sum_{i=1}^{r+1} \frac{(\gamma\mu - \lambda_i(\tilde{w}))^2}{\lambda_i(\tilde{w})} \le \left(\frac{\beta^2 + (1-\gamma)^2(r+1)}{1-\beta}\right)\mu.$$

(ii) *If* $(x, y) \in \mathcal{N}_2(\beta) \cup \mathcal{N}_{-\infty}(\beta)$ *then*

$$\sum_{i=1}^{r+1} \frac{(\gamma\mu - \lambda_i(\tilde{w}))^2}{\lambda_i(\tilde{w})} \le \left(1 - 2\gamma + \frac{\gamma^2}{1-\beta}\right)\mu(r+1).$$

Lemma 4.5 (Lemma 36 of [22]). *Define* $G = L(\tilde{y})^{-1}L(\tilde{x})$.

(i) *For the NT method, the condition number* K_G *of* G *is always* 1.

(ii) *For the xy and yx methods, we have*

- **(a)** *If* $(x, y) \in \mathcal{N}_{\mathrm{F}}(\beta) \cup \mathcal{N}_2(\beta)$ *then* $K_G \le 2/(1-\beta)$.
- **(b)** *If* $(x, y) \in \mathcal{N}_{-\infty}(\beta)$ *then* $K_G \le (r+1)/(1-\beta)$.

The next lemma is a direct consequence of Lemmas 4.3 and 4.4.

Lemma 4.6. **(i)** *If* $(x, y) \in \mathcal{N}_{\mathrm{F}}(\beta)$ *then*

$$\|\tilde{h}\|_{\mathrm{F}}^2 \le \frac{\beta^2 + (1-\gamma)^2(r+1)}{1-\beta}\mu.$$

(ii) *If* $(x, y) \in \mathcal{N}_2(\beta) \cup \mathcal{N}_{-\infty}(\beta)$ *then*

$$\|\tilde{h}\|_{\mathrm{F}}^2 \le \left(1 - 2\gamma + \frac{\gamma^2}{1-\beta}\right)\mu(r+1).$$

By Lemmas 4.1 and 4.2, we see that

$$\begin{aligned}\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}} &\le \frac{1}{2}\sqrt{K_G}\left(\|G^{1/2}u\|_{\mathrm{F}}^2 + \|G^{1/2}v\|_{\mathrm{F}}^2\right)\\ &\le \frac{1}{2}\sqrt{K_G}\left(\|\tilde{h}\|_{\mathrm{F}}^2 - \eta(1-\eta-\gamma)(r+1)\mu\right).\end{aligned}$$

Combining this with Lemma 4.6, we obtain the next theorem, which is the main result of this section.

Theorem 4.7. *Define* $G = L(\tilde{y})^{-1}L(\tilde{x})$. *Let* K_G *be the condition number of* G *defined by (4.1). Choose* $\eta, \gamma \in (0,1)$ *so that* $1 - \eta - \gamma = 0$.

(i) *If* $(x, y) \in \mathcal{N}_{\mathrm{F}}(\beta)$ *then*

$$\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}} \le \frac{1}{2}\sqrt{K_G}\left(\frac{\beta^2 + (1-\gamma)^2(r+1)}{1-\beta}\right)\mu.$$

(ii) *If* $(x, y) \in \mathcal{N}_2(\beta) \cup \mathcal{N}_{-\infty}(\beta)$ *then*

$$\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}} \le \frac{1}{2}\sqrt{K_G}\left(1 - 2\gamma + \frac{\gamma^2}{1-\beta}\right)\mu(r+1).$$

$\boxed{5}$ Scaled Lipschitz Condition

As we have seen in Lemma 3.6, to estimate the value of $\|\tilde{x}(\alpha) \circ \tilde{y}(\alpha)\|_{\mathrm{F}}$, we have to derive a bound of

$$\|\tilde{x}_{\mathrm{H}}(\alpha) \circ_{\mathrm{H}} \tilde{d}_{\mathrm{H}}(\alpha)\|_{\mathrm{F}} = \left\|(\tilde{x}_{\mathrm{H}} + \alpha\widetilde{\Delta x}_{\mathrm{H}}) \circ_{\mathrm{H}} \left[\widetilde{\psi}_{\mathrm{H}}(\tilde{x}_{\mathrm{H}}(\alpha)) - \widetilde{\psi}_{\mathrm{H}}(\tilde{x}) - \alpha D\widetilde{\psi}_{\mathrm{H}}(\tilde{x}_{\mathrm{H}})\widetilde{\Delta x}_{\mathrm{H}}\right]\right\|_{\mathrm{F}}$$

for $\alpha \in (0, 1]$. In this section, we devote ourselves to proving the following theorem by calculating $\widetilde{\psi}_{\mathrm{H}}(\tilde{x}_{\mathrm{H}}(\alpha)) - \widetilde{\psi}_{\mathrm{H}}(\tilde{x}_{\mathrm{H}}) - \alpha D\widetilde{\psi}_{\mathrm{H}}(\tilde{x}_{\mathrm{H}})\widetilde{\Delta x}_{\mathrm{H}}$ carefully.

Theorem 5.1. *Suppose that $\psi : K \to V$ satisfies Assumption 1.2. Then ψ_{H} satisfies*

$$\begin{aligned}&\left\|(\tilde{x}_{\mathrm{H}} + \alpha\widetilde{\Delta x}_{\mathrm{H}}) \circ_{\mathrm{H}} \left(\widetilde{\psi}_{\mathrm{H}}(\tilde{x}_{\mathrm{H}}(\alpha)) - \widetilde{\psi}_{\mathrm{H}}(\tilde{x}_{\mathrm{H}}) - \alpha D\widetilde{\psi}_{\mathrm{H}}(\tilde{x}_{\mathrm{H}})\widetilde{\Delta x}_{\mathrm{H}}\right)\right\|_{\mathrm{F}} \\ &\qquad\qquad \leq (2\sqrt{r}\,\theta + 1)\alpha^2\langle \widetilde{\Delta x}_{\mathrm{H}}, D\widetilde{\psi}_{\mathrm{H}}(\tilde{x}_{\mathrm{H}})\widetilde{\Delta x}_{\mathrm{H}}\rangle_{\mathrm{H}}\end{aligned}$$

for all $x_{\mathrm{H}} \in \mathrm{int}K_{\mathrm{H}}$, $\Delta x_{\mathrm{H}} \in V$, $p_{\mathrm{H}} \in \mathcal{P}(x_{\mathrm{H}}, y_{\mathrm{H}})$ and $\alpha \in [0, 1]$ such that $x_{\mathrm{H}}(\alpha) \in \mathrm{int}K_{\mathrm{H}}$ where

$$\tilde{x}_{\mathrm{H}}(\alpha) = Q_{p_{\mathrm{H}}}(x_{\mathrm{H}} + \alpha\Delta x_{\mathrm{H}}),\ \widetilde{\psi}_{\mathrm{H}}(\tilde{x}_{\mathrm{H}}) = Q_{p_{\mathrm{H}}}^{-1} \bullet_{\mathrm{H}} \psi_{\mathrm{H}} \bullet_{\mathrm{H}} Q_{p_{\mathrm{H}}}^{-1}(\tilde{x}_{\mathrm{H}}) = Q_{p_{\mathrm{H}}}^{-1}\psi_{\mathrm{H}}(x_{\mathrm{H}})$$

That is, ψ_{H} satisfies Assumption 1.2 with $2\sqrt{r}\,\theta + 1$ instead of θ.

Proof. By the definition (1.9) of $\widetilde{\psi}(\tilde{x})$ and (iii) of Lemma 2.2, we can calculate

$$\begin{aligned}
&\widetilde{\psi}(\tilde{x}_{\mathrm{H}}(\alpha)) - \widetilde{\psi}(\tilde{x}_{\mathrm{H}}) - \alpha D\widetilde{\psi}_{\mathrm{H}}(\tilde{x}_{\mathrm{H}})\widetilde{\Delta x}_{\mathrm{H}} \\
&= \begin{pmatrix} Q_p^{-1} & 0 \\ 0 & q_p^{-1}\end{pmatrix} \times \\
&\quad \begin{pmatrix} \tau(\alpha)\psi\left(\frac{x(\alpha)}{\tau(\alpha)}\right) - \tau\psi\left(\frac{x}{\tau}\right) - \alpha D\psi\left(\frac{x}{\tau}\right)\Delta x \\ \qquad -\alpha\left[\psi\left(\frac{x}{\tau}\right) - D\psi\left(\frac{x}{\tau}\right)\left(\frac{x}{\tau}\right)\right]\Delta\tau \\ -\left\langle\psi\left(\frac{x(\alpha)}{\tau(\alpha)}\right), x(\alpha)\right\rangle + \left\langle\psi\left(\frac{x}{\tau}\right), x\right\rangle \\ \qquad +\alpha\left[\left\langle\psi\left(\frac{x}{\tau}\right), \Delta x\right\rangle + \left\langle\left(\frac{x}{\tau}\right), D\psi\left(\frac{x}{\tau}\right)\left(\Delta x - \left(\frac{x}{\tau}\right)\Delta\tau\right)\right\rangle\right]\end{pmatrix} \\
&= \begin{pmatrix} Q_p^{-1}\left[\tau(\alpha)\psi\left(\frac{x(\alpha)}{\tau(\alpha)}\right) - \tau\psi\left(\frac{x}{\tau}\right) - \alpha D\psi\left(\frac{x}{\tau}\right)\Delta x - \alpha\left[\psi\left(\frac{x}{\tau}\right) - D\psi\left(\frac{x}{\tau}\right)\left(\frac{x}{\tau}\right)\right]\Delta\tau\right] \\ q_p^{-1}\left[-\left\langle\psi\left(\frac{x(\alpha)}{\tau(\alpha)}\right), x(\alpha)\right\rangle + \left\langle\psi\left(\frac{x}{\tau}\right), x\right\rangle \right.\\ \qquad \left. +\alpha\left[\left\langle\psi\left(\frac{x}{\tau}\right), \Delta x\right\rangle + \left\langle\left(\frac{x}{\tau}\right), D\psi\left(\frac{x}{\tau}\right)\left(\Delta x - \left(\frac{x}{\tau}\right)\Delta\tau\right)\right\rangle\right]\right]\end{pmatrix}
\end{aligned} \tag{5.1}$$

Define

$$z := \frac{x}{\tau}, \quad \Delta z := \frac{\Delta x - z\Delta\tau}{\tau + \alpha\Delta\tau}. \tag{5.2}$$

Then

$$\frac{x(\alpha)}{\tau(\alpha)} - \frac{x}{\tau} = \frac{x + \alpha\Delta x}{\tau + \alpha\Delta\tau} - \frac{x}{\tau} = \alpha\frac{\Delta x - z\Delta\tau}{\tau + \alpha\Delta\tau} = \alpha\Delta z$$

holds and we obtain

$$\frac{x(\alpha)}{\tau(\alpha)} = \frac{x + \alpha\Delta x}{\tau + \alpha\Delta\tau} = z + \alpha\Delta z.$$

Using them, the first and the second parts of (5.1) are given by

$$
\begin{aligned}
&Q_p^{-1}\left[\tau(\alpha)\psi\left(\frac{x(\alpha)}{\tau(\alpha)}\right)-\tau\psi\left(\frac{x}{\tau}\right)-\alpha D\psi\left(\frac{x}{\tau}\right)\Delta x-\alpha\left[\psi\left(\frac{x}{\tau}\right)-D\psi\left(\frac{x}{\tau}\right)\left(\frac{x}{\tau}\right)\right]\Delta\tau\right]\\
&= Q_p^{-1}\left[(\tau+\alpha\Delta\tau)\psi(z+\alpha\Delta z)-\tau\psi(z)-\alpha D\psi(z)\Delta x-\alpha\left[\psi(z)-D\psi(z)z\right]\Delta\tau\right]\\
&= Q_p^{-1}\left[(\tau+\alpha\Delta\tau)\psi(z+\alpha\Delta z)-(\tau+\alpha\Delta\tau)\psi(z)-\alpha D\psi(z)(\Delta x-z\Delta\tau)\right]\\
&= Q_p^{-1}\left[(\tau+\alpha\Delta\tau)\left(\psi(z+\alpha\Delta z)-\psi(z)\right)-\alpha D\psi(z)(\Delta x-z\Delta\tau)\right]\\
&= Q_p^{-1}\left[(\tau+\alpha\Delta\tau)\left(\psi(z+\alpha\Delta z)-\psi(z)\right)-\alpha(\tau+\alpha\Delta\tau)D\psi(z)\Delta z\right]\\
&= (\tau+\alpha\Delta\tau)Q_p^{-1}\left[\psi(z+\alpha\Delta z)-\psi(z)-\alpha D\psi(z)\Delta z\right]
\end{aligned}
$$

and

$$
\begin{aligned}
&q_p^{-1}\left[-\left\langle\psi\left(\frac{x(\alpha)}{\tau(\alpha)}\right),x(\alpha)\right\rangle+\left\langle\psi\left(\frac{x}{\tau}\right),x\right\rangle\right.\\
&\qquad\left.+\alpha\left[\left\langle\psi\left(\frac{x}{\tau}\right),\Delta x\right\rangle+\left\langle\left(\frac{x}{\tau}\right),D\psi\left(\frac{x}{\tau}\right)\left(\Delta x-\left(\frac{x}{\tau}\right)\Delta\tau\right)\right\rangle\right]\right]\\
&= q_p^{-1}\left[-\langle\psi(z+\alpha\Delta z),x+\alpha\Delta x\rangle+\langle\psi(z),x\rangle+\alpha\left[\langle\psi(z),\Delta x\rangle+\langle z,D\psi(z)\left(\Delta x-z\Delta\tau\right)\rangle\right]\right]\\
&= q_p^{-1}\left[-\langle\psi(z+\alpha\Delta z),x+\alpha\Delta x\rangle+\langle\psi(z),x\rangle+\alpha\left[\langle\psi(z),\Delta x\rangle+(\tau+\alpha\Delta\tau)\langle z,D\psi(z)\Delta z\rangle\right]\right]\\
&= q_p^{-1}\left[-\langle\psi(z+\alpha\Delta z),x+\alpha\Delta x\rangle+\langle\psi(z),x+\alpha\Delta x\rangle+\alpha(\tau+\alpha\Delta\tau)\langle z,D\psi(z)\Delta z\rangle\right]\\
&= q_p^{-1}\left[-\langle x+\alpha\Delta x,\psi(z+\alpha\Delta z)-\psi(z)\rangle+\alpha(\tau+\alpha\Delta\tau)\langle z,D\psi(z)\Delta z\rangle\right]\\
&= q_p^{-1}\left[-\langle x+\alpha\Delta x,\psi(z+\alpha\Delta z)-\psi(z)\rangle+\alpha\langle x+\alpha\Delta x,D\psi(z)\Delta z\rangle\right.\\
&\qquad\left.-\alpha\langle x+\alpha\Delta x,D\psi(z)\Delta z\rangle+\alpha(\tau+\alpha\Delta\tau)\langle z,D\psi(z)\Delta z\rangle\right]\\
&= q_p^{-1}\left[-\langle x+\alpha\Delta x,\psi(z+\alpha\Delta z)-\psi(z)-\alpha D\psi(z)\Delta z\rangle\right.\\
&\qquad\left.-\alpha(\tau+\alpha\Delta\tau)\langle z+\alpha\Delta z,D\psi(z)\Delta z\rangle+\alpha(\tau+\alpha\Delta\tau)\langle z,D\psi(z)\Delta z\rangle\right]\\
&= q_p^{-1}\left[-\langle x+\alpha\Delta x,\psi(z+\alpha\Delta z)-\psi(z)-\alpha D\psi(z)\Delta z\rangle\right.\\
&\qquad\left.-\alpha^2(\tau+\alpha\Delta\tau)\langle\Delta z,D\psi(z)\Delta z\rangle\right]\\
&= -(\tau+\alpha\Delta\tau)q_p^{-1}\left[\langle z+\alpha\Delta z,\psi(z+\alpha\Delta z)-\psi(z)-\alpha D\psi(z)\Delta z\rangle+\alpha^2\langle\Delta z,D\psi(z)\Delta z\rangle\right].
\end{aligned}
$$

Therefore, we see that

$$
\begin{aligned}
&\tilde{x}_{\mathrm{H}}(\alpha)\circ_{\mathrm{H}}\tilde{d}_{\mathrm{H}}(\alpha)\\
&= (\tilde{x}_{\mathrm{H}}+\alpha\widetilde{\Delta x}_{\mathrm{H}})\circ_{\mathrm{H}}\left[\widetilde{\psi}_{\mathrm{H}}(\tilde{x}_{\mathrm{H}}(\alpha))-\widetilde{\psi}_{\mathrm{H}}(\tilde{x}))-\alpha D\widetilde{\psi}_{\mathrm{H}}(\tilde{x}_{\mathrm{H}})\Delta\tilde{x}_{\mathrm{H}}\right]\\
&= \begin{pmatrix} Q_p(x+\alpha\Delta x)\circ\left[(\tau+\alpha\Delta\tau)Q_p^{-1}\left[\psi(z+\alpha\Delta z)-\psi(z)-\alpha D\psi(z)\Delta z\right]\right]\\ q_p(\tau+\alpha\Delta\tau)\left[-(\tau+\alpha\Delta\tau)q_p^{-1}\left[\langle z+\alpha\Delta z,\psi(z+\alpha\Delta z)\right.\right.\\ \qquad\left.\left.-\psi(z)-\alpha D\psi(z)\Delta z\rangle+\alpha^2\langle\Delta z,D\psi(z)\Delta z\rangle\right]\right]\end{pmatrix}\\
&= \begin{pmatrix}(\tau+\alpha\Delta\tau)^2Q_p(z+\alpha\Delta z)\circ\left[Q_p^{-1}\left[\psi(z+\alpha\Delta z)-\psi(z)-\alpha D\psi(z)\Delta z\right]\right]\\ -(\tau+\alpha\Delta\tau)^2\left[\langle z+\alpha\Delta z,\psi(z+\alpha\Delta z)-\psi(z)-\alpha D\psi(z)\Delta z\rangle+\alpha^2\langle\Delta z,D\psi(z)\Delta z\rangle\right]\end{pmatrix}\\
&= (\tau+\alpha\Delta\tau)^2\begin{pmatrix}Q_p(z+\alpha\Delta z)\circ\left[Q_p^{-1}\left[\psi(z+\alpha\Delta z)-\psi(z)-\alpha D\psi(z)\Delta z\right]\right]\\ -\left[\langle z+\alpha\Delta z,\psi(z+\alpha\Delta z)-\psi(z)-\alpha D\psi(z)\Delta z\rangle+\alpha^2\langle\Delta z,D\psi(z)\Delta z\rangle\right]\end{pmatrix}\\
&= (\tau+\alpha\Delta\tau)^2\begin{pmatrix}(\tilde{z}+\alpha\widetilde{\Delta z})\circ\left[\widetilde{\psi}(\tilde{z}+\alpha\widetilde{\Delta z})-\widetilde{\psi}(\tilde{z})-\alpha D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z}\right]\\ -\left[\langle z+\alpha\Delta z,\psi(z+\alpha\Delta z)-\psi(z)-\alpha D\psi(z)\Delta z\rangle+\alpha^2\langle\Delta z,D\psi(z)\Delta z\rangle\right]\end{pmatrix}
\end{aligned}
$$

and that

$$
\begin{aligned}
&\|\tilde{x}_{\mathrm{H}}(\alpha) \circ_{\mathrm{H}} \tilde{d}_{\mathrm{H}}(\alpha)\|_{\mathrm{F}}^2 \\
&\quad = (\tau+\alpha\Delta\tau)^4 \left[\left\| (\tilde{z}+\alpha\widetilde{\Delta z}) \circ \left[\widetilde{\psi}(\tilde{z}+\alpha\widetilde{\Delta z}) - \widetilde{\psi}(\tilde{z}) - \alpha D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z} \right] \right\|_{\mathrm{F}}^2 \right. \\
&\qquad\quad \left. + \big(\langle z+\alpha\Delta z, \psi(z+\alpha\Delta z) - \psi(z) - \alpha D\psi(z)\Delta z\rangle + \alpha^2\langle \Delta z, D\psi(z)\Delta z\rangle \big)^2 \right] \\
&\quad \leq (\tau+\alpha\Delta\tau)^4 \left[\left\| (\tilde{z}+\alpha\widetilde{\Delta z}) \circ \left[\widetilde{\psi}(\tilde{z}+\alpha\widetilde{\Delta z}) - \widetilde{\psi}(\tilde{z}) - \alpha D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z} \right] \right\|_{\mathrm{F}}^2 \right. \\
&\qquad\quad \left. + \big(|\langle z+\alpha\Delta z, \psi(z+\alpha\Delta z) - \psi(z) - \alpha D\psi(z)\Delta z\rangle| + \left|\alpha^2\langle \Delta z, D\psi(z)\Delta z\rangle\right| \big)^2 \right] \\
&\quad \leq (\tau+\alpha\Delta\tau)^4 \left[\left\| (\tilde{z}+\alpha\widetilde{\Delta z}) \circ \left[\widetilde{\psi}(\tilde{z}+\alpha\widetilde{\Delta z}) - \widetilde{\psi}(\tilde{z}) - \alpha D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z} \right] \right\|_{\mathrm{F}}^2 \right. \qquad (5.3) \\
&\qquad\quad \left. + \big(\sqrt{r}\, \|(z+\alpha\Delta z) \circ [\psi(z+\alpha\Delta z) - \psi(z) - \alpha D\psi(z)\Delta z]\|_{\mathrm{F}} + \left|\alpha^2\langle \Delta z, D\psi(z)\Delta z\rangle\right| \big)^2 \right]
\end{aligned}
$$

where the last inequality follows from the fact

$$|\langle u, v\rangle| = |\langle e, u\circ v\rangle| \leq \|e\|_{\mathrm{F}}\|u\circ v\|_{\mathrm{F}} = \sqrt{r}\|u\circ v\|_{\mathrm{F}}.$$

Since Assumption 1.2 holds, we have

$$
\begin{aligned}
\left\| (\tilde{z}+\alpha\widetilde{\Delta z}) \circ \left[\widetilde{\psi}(\tilde{z}+\alpha\widetilde{\Delta z}) - \widetilde{\psi}(\tilde{z}) - \alpha D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z} \right] \right\|_{\mathrm{F}}^2 &\leq \alpha^4\theta^2\langle \widetilde{\Delta z}, D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z}\rangle^2, \\
\|(z+\alpha\Delta z) \circ [\psi(z+\alpha\Delta z) - \psi(z) - \alpha D\psi(z)\Delta z]\|_{\mathrm{F}} &\leq \alpha^2\theta\langle z, D\psi(z)\Delta z\rangle \\
&= \alpha^2\theta\langle \widetilde{\Delta z}, D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z}\rangle.
\end{aligned}
$$

Therefore, by (5.3), it holds that

$$
\begin{aligned}
&\|\tilde{x}_{\mathrm{H}}(\alpha) \circ_{\mathrm{H}} \tilde{d}_{\mathrm{H}}(\alpha)\|_{\mathrm{F}}^2 \\
&\quad \leq (\tau+\alpha\Delta\tau)^4 \left[\alpha^4\theta^2\langle \widetilde{\Delta z}, D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z}\rangle^2 + \left(\sqrt{r}\alpha^2\theta\langle \widetilde{\Delta z}, D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z}\rangle + \left|\alpha^2\langle \Delta z, D\psi(z)\Delta z\rangle\right| \right)^2 \right] \\
&\quad \leq (\tau+\alpha\Delta\tau)^4 \left[\alpha^4\theta^2\langle \widetilde{\Delta z}, D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z}\rangle^2 + \left((\sqrt{r}\,\theta+1)\alpha^2 \left|\langle \widetilde{\Delta z}, D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z}\rangle\right| \right)^2 \right] \\
&\quad = (\tau+\alpha\Delta\tau)^4 \left[\alpha^4\theta^2\langle \widetilde{\Delta z}, D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z}\rangle^2 + (\sqrt{r}\,\theta+1)^2\alpha^4\langle \widetilde{\Delta z}, D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z}\rangle^2 \right] \\
&\quad \leq (\tau+\alpha\Delta\tau)^4 \left[(2\sqrt{r}\,\theta+1)^2\alpha^4\langle \widetilde{\Delta z}, D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z}\rangle^2 \right]. \qquad (5.4)
\end{aligned}
$$

By (iv) of Lemma 2.2, the definition (5.2) of Δz and (iv) of Lemma 3.2, we see that

$$\langle \Delta x_{\mathrm{H}}, D\psi_{\mathrm{H}}(x_{\mathrm{H}})\Delta x_{\mathrm{H}}\rangle_{\mathrm{H}} = (\tau+\alpha\Delta\tau)^2\langle \Delta z, D\psi(z)\Delta z\rangle = (\tau+\alpha\Delta\tau)^2\langle \widetilde{\Delta z}, D\widetilde{\psi}(\tilde{z})\widetilde{\Delta z}\rangle. \quad (5.5)$$

Thus, by (5.4) and (5.5), we obtain

$$\|\tilde{x}_{\mathrm{H}}(\alpha) \circ_{\mathrm{H}} \tilde{d}_{\mathrm{H}}(\alpha)\|_{\mathrm{F}} \leq (2\sqrt{r}\,\theta+1)\alpha^2\langle \Delta x_{\mathrm{H}}, D\psi_{\mathrm{H}}(x_{\mathrm{H}})\Delta x_{\mathrm{H}}\rangle_{\mathrm{H}}$$

which completes the proof of the theorem. $\square$

6 Step Sizes

In this section, we determine a range of step sizes for which the next iterate stays in the neighborhoods (1.11). Similarly as in Sections 3 and 4, we use the symbols $x, y, \psi, \langle\cdot,\cdot\rangle$ to denote $x_{\mathrm{H}}, y_{\mathrm{H}}, \psi_{\mathrm{H}}, \langle\cdot,\cdot\rangle_{\mathrm{H}}$.

Again, we introduce some useful results in [22].

Proposition 6.1 (Lemma 14, Proposition 29 and Lemma 30 (with its proof) of [22]).

(i) *Let* $x, y \in V$. $\lambda_{\min}(x+y) \geq \lambda_{\min}(x) - \|y\|_{\mathrm{F}}$, $\lambda_{\max}(x+y) \leq \lambda_{\max}(x) + \|y\|_{\mathrm{F}}$.

(ii) $\mathcal{N}_{\mathrm{F}}(\beta)$, $\mathcal{N}_2(\beta)$ *and* $\mathcal{N}_{-\infty}(\beta)$ *are scaling invariant, i.e.,* (x, y) *in the neighborhood if and only if* $(\tilde{x}, \tilde{y})$ *in the neighborhood.*

(iii) *If* $x, y \in \mathrm{int}K$ *operator commute then* $w := Q_{x^{1/2}} y = x \circ y$.

(iv) *If* $x, y \in \mathrm{int}K$ *then* $\|x \circ y - \mu e\|_{\mathrm{F}} \geq \|w - \mu e\|_{\mathrm{F}}$.

(v) *If* $x, y \in \mathrm{int}K$ *then* $\lambda_{\min}(x \circ y) \leq \lambda_{\min}(w)$.

Using the above proposition, we show the theorem below:

Theorem 6.2. *Let* $\beta \in (0,1)$ *and* $1 - \eta - \gamma = 0$. *Suppose that* $\psi : K \to V$ *satisfies Assumption 1.2. Define*

$$\bar{\alpha} := \frac{\beta\gamma\mu}{(2 + 2\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}}.$$

Then

(i) *If* $(x, y) \in \mathcal{N}_{\mathrm{F}}(\beta)$ *then* $(x(\alpha), y(\alpha)) \in \mathcal{N}_{\mathrm{F}}(\beta)$ *for any* $0 \leq \alpha \leq \bar{\alpha}$.

(ii) *If* $(x, y) \in \mathcal{N}_2(\beta)$ *then* $(x(\alpha), y(\alpha)) \in \mathcal{N}_2(\beta)$ *for any* $0 \leq \alpha \leq \bar{\alpha}$.

(iii) *If* $(x, y) \in \mathcal{N}_{-\infty}(\beta)$ *then* $(x(\alpha), y(\alpha)) \in \mathcal{N}_{-\infty}(\beta)$ *for any* $0 \leq \alpha \leq \bar{\alpha}$.

Proof. Since $(\tilde{x}(0), \tilde{y}(0)) = (\tilde{x}, \tilde{y}) \in \mathrm{int}K \times \mathrm{int}K$ and the function $(\tilde{x}(\alpha), \tilde{y}(\alpha))$ is continuous with respect to α, there exists

$$\alpha^* := \sup\{\hat{\alpha} \in (0,1] \mid (\tilde{x}(\alpha), \tilde{y}(\alpha)) \in \mathrm{int}K \times \mathrm{int}K,\ \forall \alpha \in [0, \hat{\alpha})\} > 0.$$

Define

$$\tilde{w}(\alpha) := Q_{\tilde{x}(\alpha)^{1/2}} \tilde{y}(\alpha), \quad \mu(\alpha) := \langle \tilde{x}(\alpha), \tilde{y}(\alpha)\rangle/(r+1)$$

for $\alpha \in [0, \alpha^*)$.

By (ii) of Proposition 6.1, it is enough to show that for each $\mathcal{N}(\beta) \in \{\mathcal{N}_{\mathrm{F}}(\beta), \mathcal{N}_2(\beta), \mathcal{N}_{-\infty}(\beta)\}$, if $(\tilde{x}, \tilde{y}) \in \mathcal{N}(\beta)$ then $(\tilde{x}(\alpha), \tilde{y}(\alpha)) \in \mathcal{N}(\beta)$ for any $\alpha \in [0, \bar{\alpha}]$.

Since $\tilde{x}$ and $\tilde{y}$ operator commute (cf. (1.6)) and we set $1 - \eta - \gamma = 0$, by Lemma 3.6, (ii) of Lemma 3.5 and (iii) of Proposition 6.1, we have

$$\begin{aligned}
&\tilde{x}(\alpha) \circ \tilde{y}(\alpha) - \mu(\alpha)e \\
&= (1-\alpha)\tilde{x} \circ \tilde{y} + \alpha\gamma\mu e + \alpha^2 \widetilde{\Delta x} \circ \widetilde{\Delta y} + \tilde{x}(\alpha) \circ \tilde{d}(\alpha) - \{1 - \alpha(1-\gamma)\}\mu e \\
&= (1-\alpha)(\tilde{x} \circ \tilde{y} - \mu e) + \alpha^2 \widetilde{\Delta x} \circ \widetilde{\Delta y} + \tilde{x}(\alpha) \circ \tilde{d}(\alpha) \\
&= (1-\alpha)(\tilde{w} - \mu e) + \alpha^2 \widetilde{\Delta x} \circ \widetilde{\Delta y} + \tilde{x}(\alpha) \circ \tilde{d}(\alpha)
\end{aligned}$$

and

$$\begin{aligned}
\|\tilde{x}(\alpha)\circ\tilde{y}(\alpha)-\mu(\alpha)e\|_{\mathrm{F}} &\le (1-\alpha)\|\tilde{w}-\mu e\|_{\mathrm{F}}+\alpha^2\|\widetilde{\Delta x}\circ\widetilde{\Delta y}\|_{\mathrm{F}}+\|\tilde{x}(\alpha)\circ\tilde{d}(\alpha)\|_{\mathrm{F}}\\
&\le (1-\alpha)\|\tilde{w}-\mu e\|_{\mathrm{F}}+\alpha^2\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}+\|\tilde{x}(\alpha)\circ\tilde{d}(\alpha)\|_{\mathrm{F}}\\
&\le (1-\alpha)\|\tilde{w}-\mu e\|_{\mathrm{F}}+\alpha^2\left(1+(2\sqrt{r}\,\theta+1)\right)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}\\
&= (1-\alpha)\|\tilde{w}-\mu e\|_{\mathrm{F}}+\alpha^2(2+2\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}
\end{aligned}$$

where the second inequality follows from the fact that $\|u\circ v\|_{\mathrm{F}}\le\|u\|_{\mathrm{F}}\|v\|_{\mathrm{F}}$ holds for any $u,v\in V$ (cf. Lemma 2.9 of [20]) and the last inequality follows from Theorem 5.1, Lemma 3.4 and $1-\eta-\gamma=0$. Since if $(x,y)\in\mathcal{N}_{\mathrm{F}}(\beta)$ then $(\tilde{x},\tilde{y})\in\mathcal{N}_{\mathrm{F}}(\beta)$ by (ii) of Proposition 6.1, this yields that

$$\|\tilde{x}(\alpha)\circ\tilde{y}(\alpha)-\mu(\alpha)e\|_{\mathrm{F}}\le(1-\alpha)\beta\mu+\alpha^2(2+2\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}.$$

So, by (ii) of Lemma 3.5 and by $1-\eta-\gamma=0$,

$$\|\tilde{x}(\alpha)\circ\tilde{y}(\alpha)-\mu(\alpha)e\|_{\mathrm{F}}\le\beta\mu(\alpha)$$

holds if

$$(1-\alpha)\beta\mu+\alpha^2(2+2\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}\le\beta\{1-\alpha(1-\gamma)\}\mu$$

or equivalently,

$$0\le\alpha\le\frac{\beta\gamma\mu}{(2+2\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}}=:\bar{\alpha}.$$

By a similar discussion to the proof of Lemma 32 of [22], we can see that for any $\alpha\in[0,\bar{\alpha}]$, $\tilde{x}(\alpha)$ and $\tilde{y}(\alpha)$ are positive definite and $\bar{\alpha}\in(0,\alpha^*)$. Thus, by (iv) and (ii) of Proposition 6.1, we have $(x(\alpha),y(\alpha))\in\mathcal{N}_{\mathrm{F}}(\beta)$ for $\alpha\in[0,\bar{\alpha}]$.

Similarly,

$$\begin{aligned}
&\lambda_{\min}\left(\tilde{x}(\alpha)\circ\tilde{y}(\alpha)-\mu(\alpha)e\right)\\
&\quad\ge \lambda_{\min}\left((1-\alpha)\left(\tilde{x}\circ\tilde{y}-\mu e\right)\right)-\alpha^2\|\widetilde{\Delta x}\circ\widetilde{\Delta y}\|_{\mathrm{F}}-\|\tilde{x}(\alpha)\circ\tilde{d}(\alpha)\|_{\mathrm{F}}\\
&\qquad\qquad\text{(by Lemma 3.6 and (i) of Proposition 6.1)}\\
&\quad\ge (1-\alpha)\lambda_{\min}\left(\tilde{x}\circ\tilde{y}-\mu e\right)-\alpha^2\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}-\|\tilde{x}(\alpha)\circ\tilde{d}(\alpha)\|_{\mathrm{F}}\\
&\quad\ge (1-\alpha)\lambda_{\min}\left(\tilde{x}\circ\tilde{y}-\mu e\right)-\alpha^2(2+2\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}\\
&\qquad\qquad\text{(by Theorem 5.1, Lemma 3.4 and } 1-\eta-\gamma=0)\\
&\quad= (1-\alpha)\lambda_{\min}\left(\tilde{w}-\mu e\right)-\alpha^2(2+2\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}\\
&\qquad\qquad\text{(by (iii) of Proposition 6.1)}
\end{aligned}$$

and by the same discussion we have

$$\lambda_{\max}\left(\tilde{x}(\alpha)\circ\tilde{y}(\alpha)-\mu(\alpha)e\right)\le(1-\alpha)\lambda_{\max}\left(\tilde{w}-\mu e\right)+\alpha^2(2+2\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}.$$

Therefore, by (ii) of Proposition 6.1, if $(x,y)\in\mathcal{N}_{-\infty}(\beta)$ then

$$\lambda_{\min}\left(\tilde{x}(\alpha)\circ\tilde{y}(\alpha)-\mu(\alpha)e\right)\ \ge\ -(1-\alpha)\beta\mu-\alpha^2(2+2\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}},$$

holds, and if $(x, y) \in \mathcal{N}_2(\beta)$ then

$$\begin{aligned} \lambda_{\min}\left(\tilde{x}(\alpha) \circ \tilde{y}(\alpha) - \mu(\alpha)e\right) &\geq -(1-\alpha)\beta\mu - \alpha^2(2 + 2\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}} \quad \text{and} \\ \lambda_{\max}\left(\tilde{x}(\alpha) \circ \tilde{y}(\alpha) - \mu(\alpha)e\right) &\leq (1-\alpha)\beta\mu + \alpha^2(2 + 2\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}} \end{aligned}$$

hold. Since $\alpha \in [0, \bar{\alpha}]$ implies

$$(1-\alpha)\beta\mu + \alpha^2(2 + 2\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}} \leq \beta\{1 - \alpha(1-\gamma)\}\mu,$$

it follows from (ii) of Lemma 3.5 and (iv) and (ii) of Proposition 6.1 that $(x(\alpha), y(\alpha)) \in \mathcal{N}_2(\beta)$ if $(x, y) \in \mathcal{N}_2(\beta)$, and $(x(\alpha), y(\alpha)) \in \mathcal{N}_{-\infty}(\beta)$ if $(x, y) \in \mathcal{N}_{-\infty}(\beta)$ for $\alpha \in [0, \bar{\alpha}]$. □

7 Homogeneous Algorithms and Their Complexity Bounds

Here, we give a detailed description of the homogeneous algorithms:

Input **1.** Choose $\epsilon > 0$ and $\beta \in (0, 1)$.

2. Choose the neighborhood $\mathcal{N}(\beta) \in \{\mathcal{N}_{\mathrm{F}}(\beta),\ \mathcal{N}_2(\beta),\ \mathcal{N}_{-\infty}(\beta)\}$.

3. Set $\gamma := 1 - 1/\sqrt{r+1}$ if $\mathcal{N}(\beta) = \mathcal{N}_{\mathrm{F}}(\beta)$, and set $\gamma := 1/2$ if $\mathcal{N}(\beta) = \mathcal{N}_2(\beta)$ or $\mathcal{N}(\beta) = \mathcal{N}_{-\infty}(\beta)$.

4. Set $\eta = 1 - \gamma$.

5. Let $k := 0$. Let $(x_{\mathrm{H}}^{(0)}, y_{\mathrm{H}}^{(0)}) := (e_{\mathrm{H}}, e_{\mathrm{H}}) \in \mathcal{N}(\beta)$ be the initial point. Let $\mu^{(0)} := \langle x^{(0)}, y^{(0)}\rangle_{\mathrm{H}}/(r+1) = 1$.

begin

while $\mu_{\mathrm{H}}^{(k)} > \epsilon$ **do**

1. Set $(x_{\mathrm{H}}, y_{\mathrm{H}}) := (x_{\mathrm{H}}^{(k)}, y_{\mathrm{H}}^{(k)})$ and $\mu_{\mathrm{H}} := \mu_{\mathrm{H}}^{(k)}$.

2. Choose a scaling element $p \in \mathcal{P}(x_{\mathrm{H}}, y_{\mathrm{H}})$ and compute $(\tilde{x}, \tilde{y})$ by (1.6).

3. Compute the Newton direction $(\Delta x_{\mathrm{H}}, \Delta y_{\mathrm{H}})$ by solving the scaled Newton system (1.8) and applying the inverse scaling (1.7).

4. Choose the largest step-size $\tilde{\alpha} \in (0, 1]$ such that $(x_{\mathrm{H}}(\alpha), y_{\mathrm{H}}(\alpha)) \in \mathcal{N}(\beta)$ where $x_{\mathrm{H}}(\alpha)$ and $y_{\mathrm{H}}(\alpha)$ are defined by (1.10).

5. Set $(x_{\mathrm{H}}^{(k+1)}, y_{\mathrm{H}}^{(k+1)}) := (x_{\mathrm{H}}(\tilde{\alpha}), y_{\mathrm{H}}(\tilde{\alpha}))$ and $\mu_{\mathrm{H}}^{(k+1)} := \langle x_{\mathrm{H}}^{(k+1)}, y_{\mathrm{H}}^{(k+1)}\rangle_{\mathrm{H}}/(r+1)$. $k := k + 1$.

end

end.

Theorem 7.1. *Suppose that $\psi : K \to V$ satisfies Assumption 1.2, and that the condition number $\sqrt{K_G}$ (see (4.1) for the definition of K_G) can be bounded from above by $\kappa < \infty$ for all iterations of the algorithm.*

(i) *The short-step algorithm terminates in $\mathcal{O}\left(\kappa\sqrt{r}(1 + \sqrt{r}\,\theta)\log\epsilon^{-1}\right)$ iterations.*

(ii) *The semi-long-, and long-step-algorithms terminate in $\mathcal{O}\left(\kappa r(1 + \sqrt{r}\,\theta)\log\epsilon^{-1}\right)$ iterations.*

Proof. Since $1-\eta-\gamma=0$ holds in the algorithm, by Lemma 3.5 and the definition of μ, we have

$$\mu^{(k+1)} \le \{1-\tilde{\alpha}(1-\gamma)\}\mu^{(k)}. \tag{7.1}$$

First we analyze the algorithm using the neighborhood $\mathcal{N}(\beta)=\mathcal{N}_{\mathrm{F}}(\beta)$. Theorems 6.2 and 4.7 and the assumption $\sqrt{K_G}\le\kappa$ ensure that

$$\begin{aligned}\tilde{\alpha}\ge\bar{\alpha} \quad &:= \quad \frac{\beta\gamma\mu}{(2+2\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}} \\ &\ge \quad \frac{\beta\gamma}{(2+2\sqrt{r}\,\theta)}\,\frac{2(1-\beta)}{\kappa\{\beta^2+(1-\gamma)^2(r+1)\}}.\end{aligned}$$

Since we set $1-\gamma=1/\sqrt{r+1}$, $\gamma\ge 1/4$ holds, it follows from (7.1) that

$$\mu^{(k+1)} \le \left\{1-\frac{\beta(1-\beta)}{\sqrt{r+1}(4+4\sqrt{r}\,\theta)\kappa(\beta^2+1)}\right\}\mu^{(k)}.$$

Since $\mu^{(0)}=1$, the algorithm terminates after $\mathcal{O}\left(\kappa\sqrt{r}(1+\sqrt{r}\,\theta)\log\epsilon^{-1}\right)$ number of iterations.

Similarly, if we use $\mathcal{N}(\beta)=\mathcal{N}_2(\beta)$ or $\mathcal{N}(\beta)=\mathcal{N}_{-\infty}(\beta)$ then

$$\begin{aligned}\tilde{\alpha}\ge\bar{\alpha} \quad &:= \quad \frac{\beta\gamma\mu}{(4+4\sqrt{r}\,\theta)\|\widetilde{\Delta x}\|_{\mathrm{F}}\|\widetilde{\Delta y}\|_{\mathrm{F}}} \\ &\ge \quad \frac{\beta\gamma}{(4+4\sqrt{r}\,\theta)}\,\frac{2(1-\beta)}{\kappa(r+1)\left\{(1-2\gamma)(1-\beta)+\gamma^2\right\}}.\end{aligned}$$

Since $\gamma=1/2$ in these cases,

$$\mu^{(k+1)} \le \left\{1-\frac{\beta(1-\beta)}{(1+\sqrt{r}\,\theta)\kappa(r+1)}\right\}\mu^{(k)}$$

holds and the algorithm terminates after $\mathcal{O}\left(\kappa r(1+\sqrt{r}\,\theta)\log\epsilon^{-1}\right)$ number of iterations. $\square$

Remark 7.2. *Note that our homogeneous algorithm does not take the feasibility of* $(x_{\mathrm{H}}^{(k)},y_{\mathrm{H}}^{(k)})$ *into account, but the obtained* $(x_{\mathrm{H}}^{(k)},y_{\mathrm{H}}^{(k)})$ *will be sufficiently feasible. In fact, by (i) of Lemma 3.5 with* $\eta=1-\gamma$ *and the same discussion as in the proof of Theorem 7.1, we can see that*

$$\|s_{\mathrm{H}}^{(k)}\|_{\mathrm{F}}=\|y_{\mathrm{H}}^{(k)}-\psi_{\mathrm{H}}(x_{\mathrm{H}}^{(k)})\|_{\mathrm{F}}\le\epsilon$$

holds after $\mathcal{O}\left(\kappa r(1+\sqrt{r}\,\theta)\log\epsilon^{-1}\right)$ *number of iterations.*

By Lemma 4.5 and the above theorem, we obtain the following corollary.

Corollary 7.3. *Suppose that* $\psi: K\to V$ *satisfies Assumption 1.2. Suppose that we use the NT, xy or yx method for determining the search direction. Then the number of iterations of each homogeneous algorithm is bounded as follows:*

	NT method	*xy or yx method*
Short-step using $\mathcal{N}_{\mathrm{F}}(\beta)$	$\mathcal{O}\left(\sqrt{r}(1+\sqrt{r}\,\theta)\log\epsilon^{-1}\right)$	$\mathcal{O}\left(\sqrt{r}(1+\sqrt{r}\,\theta)\log\epsilon^{-1}\right)$
Semi-long-step using $\mathcal{N}_2(\beta)$	$\mathcal{O}\left(r(1+\sqrt{r}\,\theta)\log\epsilon^{-1}\right)$	$\mathcal{O}\left(r(1+\sqrt{r}\,\theta)\log\epsilon^{-1}\right)$
Long-step using $\mathcal{N}_{-\infty}(\beta)$	$\mathcal{O}\left(r(1+\sqrt{r}\,\theta)\log\epsilon^{-1}\right)$	$\mathcal{O}\left(r^{1.5}(1+\sqrt{r}\,\theta)\log\epsilon^{-1}\right)$

8 Concluding Remarks

In this paper, we provided a class of homogeneous algorithms for monotone complementarity problems (CPs) based on the homogeneous model in [24]. The algorithms (a) start from an infeasible interior point, (b) use the commutative class of search directions including the xy, the yx and the Nesterov and Todd (NT) directions, and (c) use the $\mathcal{N}_{\mathrm{F}}$, $\mathcal{N}_2$ and $\mathcal{N}_{-\infty}$ neighborhoods. To analyze their iteration complexity, a scaled Lipschitz property of the function ψ over int$K$ and an associated parameter $\theta \geq 0$ have been introduced. We showed that the scaled Lipschitz property of the function ψ is inherited by the homogeneous function ψ_{H} with the same order of θ. We also showed that the curve search technique for various problems over nonnegative orthants (cf. [15, 19, 25]) can be extended to CPs over symmetric cones. Consequently, we derived polynomial iteration-complexity bounds of the algorithms which are the best obtained so far when the function ψ is affine.

Acknowledgment

The author would like to thank two anonymous referees for their valuable comments. In particular, the author is indebted to one of the referees for his/her careful reading of the manuscript and for many suggestions which helped to improve the paper.

References

[1] F. Alizadeh, J.P. Haebely, M.L. Overton, Primal-dual interior-point methods foe semidefinite programming: Convergence rates, stability and numerical results, *SIAM Journal on Optimization* 8 (1998) 746–768.

[2] F. Alizadeh and S.H. Schmieta, Symmetric cones, potential reduction methods and word-by-word extensions, in *Handbook of Semidefinite Programming. Theory, Algorithms, and Applications*, H. Wolkowicz, R. Saigal and L. Vandenberghe (eds.), Kluwer Academic Publishers, Boston, MA, 2000.

[3] E. Andersen and Y. Ye, On a homogeneous algorithm for the monotone complementarity problems, *Mathematical Programming* 84 (1999) 375–400.

[4] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems: Volume I*, Springer, New York, 2003.

[5] J. Faraut and A. Korányi, *Analysis on Symmetric Cones* Oxford Science Publishers, 1994.

[6] L. Faybusovich, Euclidean Jordan algebras and interior-point algorithms, *Positivity* 1 (1997)331–357.

[7] L. Faybusovich, Linear systems in Jordan algebras and primal-dual interior point algorithms, *Journal of Computational and Applied Mathematics* 86 (1997) 149–175.

[8] L. Faybusovich and T. Tsuchiya, Primal-dual algorithms and infinite-dimensional Jordan algebras of finite rank, *Mathematical Programming* 97 (2003) 471–493.

[9] L. Faybusovich, Several Jordan-algebraic aspects of optimization, *Optimization* 57 (2008) 379–393.

[10] M.S. Gowda, R. Sznajder and J. Tao, Some P-properties for linear transformations on Euclidean Jordan algebras, *Linear Algebra and its Applications* 393 (2004) 203–232.

[11] M.S. Gowda and R. Sznajder, Some global uniqueness and solvability results for linear complementarity problems over symmetric cones, *SIAM Journal on Optimization* 18 (2007) 461–481.

[12] O. Güler, Barrier functions in interior-point methods, *Mathematics of Operations Research* 21 (1996) 860–885.

[13] R.A. Hauser and O. Güler, Self-scaled barrier functions on symmetric cones and their classification, *Foundations of Computational Mathematics* 2 (2002) 121–143.

[14] B. Jansen, K. Roos, T, Terlaky and A. Yoshise, Polynomiality of primal-dual affine scaling algorithms for nonlinear complementarity problems, *Mathematical Programming* 77 (1997) 315–345.

[15] R.D.C. Monteiro and I. Adler, An extension of Karmarkar type algorithm to a class of convex separable programming problems with global linear rate of convergence, *Mathematics of Operations Research* 15 (1990) 408–422.

[16] R.D.C. Monteiro and Y. Zhang, A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming, *Mathematical Programming* 81 (1998) 281–299.

[17] M. Muramatsu, On commutative class of search directions for linear programming over symmetric cones, *Journal of Optimization Theory and Applications* 112 (2002) 595–625.

[18] Yu.E. Nesterov and M.J. Todd, Self-scaled barriers and interior-point for convex programming, *Mathematics of Operations Research* 22 (1997) 1–42.

[19] F. Potra and Y. Ye, Interior-point methods for nonlinear complementarity problem, *Journal of Optimization Theory and Application*, 88 (1996) 617–642.

[20] B.K. Rangarajan, Polynomial convergence of infeasible-interior-point methods over symmetric cones, *SIAM Journal of Optimization* 16 (2006) 1211–1229.

[21] S.H. Schmieta, Complete classification of self-scaled barrier functions, Technical report, Department of IEOR, Columbia University, New York, 2000.

[22] S.H. Schmieta and F. Alizadeh, Extension of primal-dual interior-point algorithm to symmetric cones, *Mathematical Programming* 96 (2003) 409–438.

[23] J. Tao and M.S. Gowda, Some P-properties for nonlinear transformations on Euclidean Jordan algebras, *Mathematics of Operations Research* 30 (2005) 985–1004.

[24] A. Yoshise, Interior point trajectories and a homogeneous model for nonlinear complementarity problems over symmetric cones, *SIAM Journal on Optimization* 17 (2006) 1129–1153.

[25] J. Zhu, A path following algorithm for a class of convex programming problems, *Zeitschrift fur Operations Research* 36 (1992) 359–377.

Manuscript received 29 February 2008
revised 10 January 2009
accepted for publication 6 March 2009

Akiko Yoshise
Graduate School of Systems and Information Engineering, University of Tsukuba
Tsukuba, Ibaraki 305-8573, Japan
E-mail address: yoshise@sk.tsukuba.ac.jp