Department of Social Systems and Management

Discussion Paper Series

A Primal Barrier Function Phase I Algorithm for Nonsymmetric Conic Optimization Problems

by

Yasuaki MATSUKAWA and Akiko YOSHISE

November 2011

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

A Primal Barrier Function Phase I Algorithm for Nonsymmetric Conic Optimization Problems

Yasuaki Matsukawa* and Akiko Yoshise†

Abstract

We call a positive semidefinite matrix whose elements are nonnegative a *doubly nonnegative matrix*, and the set of those matrices the *doubly nonnegative cone* (DNN cone). The DNN cone is not symmetric but can be represented as the projection of a symmetric cone embedded in a higher dimension. In [16], the authors demonstrated the efficiency of the DNN relaxation using the symmetric cone representation of the DNN cone. They showed that the DNN relaxation gives significantly tight bounds for a class of quadratic assignment problems, but the computational time is not affordable as long as we employ the symmetric cone representation. They then suggested a primal barrier function approach for solving the DNN optimization problem directly, instead of using the symmetric cone representation. However, most of existing studies on the primal barrier function approach have assumed the availability of a feasible interior point. This fact means that those studies are not inextricably tied to the practical usage. Motivated by these observations, we propose a primal barrier function Phase I algorithm for solving conic optimization problems over the closed convex cone K having the following properties: (a) its interior $\operatorname{int} K$ is not necessarily symmetric, (b) a self-concordant function f is defined over $\operatorname{int} K$, and (c) its dual cone K^* is not explicit or is intractable, all of which are observed when K is the DNN cone. We analyze the algorithm and provide a sufficient condition for finite termination.

1 Introduction

This paper deals with the conic optimization problem given by

$$
\begin{array}{ll}
\text{Minimize} & \langle c, x \rangle \\
\text{subject to} & Ax = b, \\
& x \in K
\end{array}
\tag{1}
$$

where $K \subseteq \mathbb{R}^p$ is a closed convex cone, $A : \mathbb{R}^p \to \mathbb{R}^q$ is a linear operator, $b \in \mathbb{R}^q$, $c \in \mathbb{R}^p$ and $\langle \cdot, \cdot \rangle$ is an inner product on $\mathbb{R}^p$. In the last two decades, many studies have been done on the symmetric conic

optimization where the closed convex cone K is self-dual, i.e., K satisfies

$$K = K^* := \{y \mid \langle x, y \rangle \geq 0 \text{ for all } x \in K\}.$$

A typical example of such problems is the semidefinite optimization problem with the semidefinite cone (SDP cone) given by

$$K = \mathcal{S}_+^n := \left\{X \in \mathbb{R}^{n \times n} \mid X = X^T, \ \forall d \in \mathbb{R}^n, \ d^T X d \geq 0\right\}.$$

The effectiveness of semidefinite relaxation (SDP relaxation) approach for combinatorial optimization problems has been supported in many papers (see, e.g., [2, 4, 15, 5, 3, 10, 7]).

In the SDP relaxation, the solution matrix X is restricted to $X \in \mathcal{S}_+^n$ and often meant to be nonnegative, i.e.,

$$X \in \mathbb{R}_+^{n \times n} := \left\{X \in \mathbb{R}^{n \times n} \mid x_{ij} \geq 0 \ (i, j = 1, 2, \dots, n)\right\}.$$

We call a positive semidefinite matrix whose elements are nonnegative a *doubly nonnegative matrix*, and the set

$$\mathcal{D}^n := \mathcal{S}_+^n \cap \mathbb{R}_+^{n \times n}$$

of those matrices the *doubly nonnegative cone* (DNN cone). The DNN cone can be represented as the projection of the direct sum of the SDP cone and the nonnegative orthant as follows:

$$\mathcal{D}^n = \{X \mid (X, Y) \in \mathcal{S}_+^n \times \mathbb{R}_+^{n \times n}, \ X = Y\}. \tag{2}$$

Note that the set $\mathcal{S}_+^n \times \mathbb{R}_+^{n \times n}$ is a symmetric cone, and we can adopt existing symmetric conic optimization solvers to solve the DNN relaxation. Matsukawa and Yoshise [16] demonstrated the efficiency of this approach and reported that

- the DNN relaxation gives significantly tighter bounds than the SDP relaxation for a class of quadratic assignment problems: more than 95% accuracy can be achieved for many instances from the QAPLIB library [1], but

- the size of the problem grows too much large due to the symmetric cone representation (2): even a mid-size problem with $n = 15$ requires more than 7 hours to solve the relaxation problem with SDPA Online Solver [13] in 2010.

Motivated by the above observations, in [16], the authors provided basic properties of the DNN cone aiming to develop another approach, and showed that

(i) the dual cone $(\mathcal{D}^n)^*$ of the DNN cone is

$$(\mathcal{D}^n)^* = \mathcal{S}_+^n + \mathbb{R}_+^{n \times n} := \{x + y \mid x \in \mathcal{S}_+^n, \ y \in \mathbb{R}_+^{n \times n}\} \tag{3}$$

(see Proposition 4.1 of [16]) and

(ii) the interior of the DNN cone $\operatorname{int} \mathcal{D}^n$ is a hyperbolic cone characterized by the self-concordant function $f : \operatorname{int} \mathcal{D}^n \to \mathbb{R}$

$$f(X) := -\log \det (X) - \sum_{i=1}^{n} \sum_{i=1}^{n} \log x_{ij} \tag{4}$$

(see Proposition 4.2 of [16] and cf. also [6, 12]).

The property (ii) above ensures that we can apply the primal barrier function approach provided in, e.g., [8, 11, 9] to solve the DNN optimization problem. In particular, the paper [9] dealt with the nonsymmetric conic optimization where only the primal cone K has a self-concordant function on its interior int K. However, all of these studies assume that a feasible interior point x^0 satisfying

$$Ax^0 = b, \quad x^0 \in \operatorname{int} K \tag{5}$$

is obtained a priori, which means that those studies are not inextricably tied to the practical usage. In fact, with the appearance of the primal-dual interior point algorithm in 1989 for linear optimization, the number of studies on the primal barrier function algorithm diminished and there are few papers on the primal infeasible interior point algorithm which may start from an infeasible interior point.

Recently, Skajaa, Jørgensen and Hansen [14] proposed a homogeneous interior-point algorithm for non-symmetric conic optimization which does not need to assume the existence of a feasible interior point. Note that, in either [14] or [9], it has been assumed that we can easily check whether an element s lies in the dual cone K^* of K. However, as described in (ii) above, the dual cone $(\mathcal{D}^n)^*$ of the DNN cone $\mathcal{D}^n$ is given by (3) and it is not necessarily easy to find whether $s \in (\mathcal{D}^n)^*$ or not. A new and purely primal algorithm would be expected to find a feasible interior point satisfying (5) for the DNN optimization problem.

In this paper, we propose an algorithm to find a feasible interior point of the conic optimization (1) where the closed convex cone has the following properties:

(a) int K is not necessarily symmetric, i.e., it can be not self-dual or not homogeneous,

(b) a self-concordant function f is defined over int K, and

(c) its dual cone K^* is not explicit or is intractable.

Our algorithm is a purely primal barrier function algorithm based on the self-concordant function f, and does not require any knowledge of the dual cone K^*.

The paper is organized as follows. In Section 2, we summarize basic properties of the self-concordant function according to [11]. The primal barrier function Phase I algorithm is provided in Section 3. The algorithm consists of the *feasibility step* and the well-known *centering steps*. After analyzing the *feasibility step* in Section 4, we provide a sufficient condition for finite termination of the algorithm and estimate the required number of iterations in Section 5. Concluding remarks are given in Section 6.

2 Basic properties of the self-concordant function

Our analysis crucially relies on Section 2 of [11] where many properties of the self-concordant function have been investigated. In this section, we provide some basic properties of the self-concordant function used in the paper.

Let D_f be an open convex set and $f : D_f \to \mathbb{R}$ be a functional $f \in \mathcal{C}^2$. We denote by $g(x)$ and $H(x)$ the gradient and the Hessian of $f(x)$ at x, respectively. We obtain the following proposition.

Proposition 2.1 (Proposition 1.5.7 of [11]). *If* $x, y \in D_f$ *then*

$$g(y) = g(x) + H(x)(y - x) + \int_0^1 \big(H(x + t(y - x)) - H(x) \big)(y - x) dt.$$

In what follows, we suppose that $H(x)$ is positive definite for all $x \in D_f$ which implies that f is strictly convex. The local inner (or the intrinsic) product and the induced norm are defined as follows:

$$\langle u, v \rangle_x := \langle u, H(x)v \rangle, \quad \|u\|_x := \sqrt{\langle u, u \rangle_x}. \tag{6}$$

Let $B_x(x, \lambda)$ be the open ball given by

$$B_x(x, \lambda) := \{y \mid \|y - x\|_x < \lambda\}.$$

In [11], the self-concordant function is defined as follows.

Definition 2.2 (Section 2.2.1 of [11]). *A functional $f : D_f \to \mathbb{R}$ is said to be (strongly nondegenerate) self-concordant if for all $x \in D_f$ we have*

$$B_x(x, 1) \subseteq D_f \tag{7}$$

and if whenever $y \in B_x(x, 1)$ we have

$$1 - \|y - x\|_x \leq \frac{\|v\|_y}{\|v\|_x} \leq \frac{1}{1 - \|y - x\|_x} \quad \text{for all } v \neq 0. \tag{8}$$

Since $H(x)$ is positive definite, $H(x)^{-1}$ exists for any $x \in D_f$. By introducing the notation

$$H_x(y) := H(x)^{-1}H(y), \tag{9}$$

we see the following properties.

Theorem 2.3 (Theorem 2.2.1 of [11]). *Assume that the functional f has the property that $B_x(x, 1) \subseteq D_f$ for all $x \in D_f$.*

(i) *f is self-concordant iff for all $x \in D_f$ and $y \in B_x(x, 1)$*

$$\|H_x(y)\|_x, \|H_x(y)^{-1}\|_x \leq \frac{1}{(1 - \|y - x\|_x)^2}.$$

(i) *f is self-concordant iff for all $x \in D_f$ and $y \in B_x(x, 1)$*

$$\|I - H_x(y)\|_x, \|I - H_x(y)^{-1}\|_x \leq \frac{1}{(1 - \|y - x\|_x)^2} - 1.$$

We denote by $n(x)$ the Newton direction for f at x which is given by

$$n(x) = -H(x)^{-1}g(x) \tag{10}$$

and by $\bar{n}(x)$ the projected Newton direction which is the unique solution of the system

$$H(x)\bar{n}(x) + g(x) = A^*y, \tag{11}$$
$$A\bar{n}(x) = 0, \tag{12}$$

where A^* denotes the adjoint of A satisfying $\langle x, A^*y \rangle = \langle Ax, y \rangle$ for any $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$. The projected Newton direction $\bar{n}(x)$ and the Newton direction $n(x)$ have the following relation:

$$\begin{aligned}
\langle \bar{n}(x), n(x) \rangle_x &= \langle \bar{n}(x), \bar{n}(x) - H(x)^{-1}A^*y \rangle_x & \text{(by (10) and (11))} \\
&= \langle \bar{n}(x), \bar{n}(x) \rangle_x - \langle \bar{n}(x), A^*y \rangle & \text{(by (6))} \\
&= \|\bar{n}(x)\|_x^2 - \langle A\bar{n}(x), y \rangle \\
&= \|\bar{n}(x)\|_x^2. \quad \text{(by (12))} \tag{13}
\end{aligned}$$

The theorem below gives an upper bound of the difference between the function value $f(y)$ and its second-order approximation at x of the self-concordant function f.

4

Theorem 2.4 (Theorem 2.2.2 of [11]). *If f is self-concordant, $x \in D_f$ and $y \in B_x(x,1)$, then*

$$|f(y) - q_x(y)| \leq \frac{\|y - x\|_x^3}{3(1 - \|y - x\|_x)}$$

where $q_x(y)$ is the second-order approximation of $f(y)$ at x given by

$$
\begin{aligned}
q_x(y) \quad &:= \quad f(x) + \langle g(x), (y - x) \rangle + \frac{1}{2} \langle (y - x), H(x)(y - x) \rangle \\
&= \quad f(x) - \langle n(x), (y - x) \rangle_x + \frac{1}{2} \|y - x\|_x^2
\end{aligned}
$$

Here the second equality above follows from the definitions of (10) and (6). At the last of this section, we give two propositions on the complexity value of f.

Proposition 2.5 (Section 2.3.1 of [11]). *Let $f : D_f \to \mathbb{R}$ be a strongly nondegenerate self-concordant barrier. Let ϑ_f be the complexity value of f given by*

$$\vartheta_f := \sup_{x \in D_f} \|n(x)\|_x^2.$$

Then the following inequalities hold.

$$\|\bar{n}(x)\|_x \leq \|n(x)\|_x \leq \sqrt{\vartheta_f}.$$

For the self-concordant function f over the interior $\operatorname{int} \mathcal{D}^n$ of the DNN cone $\mathcal{D}^n$ in (2), we obtain the following lemma by simple calculations.

Proposition 2.6. *Let f be a self-concordant barrier on $\operatorname{int} \mathcal{D}^n$ given by (4). The complexity value*

$$\vartheta_f := \sup_{x \in \operatorname{int} \mathcal{D}^n} \|n(x)\|_x^2$$

of the function f is $n^2 + n$.

3   A primal barrier function Phase I algorithm for the nonsymmetric conic optimization problem

Let K be a closed convex cone which is not necessarily symmetric, and f be a self-concordant functional defined over $D_f := \operatorname{int} K$. In this section, we propose an algorithm for finding a feasible interior point of the primal (nonsymmetric) optimization problem of (1).

Recall that the Newton direction $n(x)$ and the projected Newton direction $\bar{n}(x)$ for f at x are given by (10) and (11) – (12), respectively. Here we outline the algorithm.

Let $\tau \in (0,1)$ and choose an initial point $x^0 \in \operatorname{int} K$. Define

$$r^0 := Ax^0 - b, \quad S_\gamma := \{x \in \operatorname{int} K \mid Ax - b = \gamma r^0\}. \tag{14}$$

Note that r^0 is not necessarily zero and x^0 may be an infeasible interior point. We impose the following assumption which is often satisfied by the DNN relaxation problems of combinatorial optimization problems (cf., e.g., [2, 4, 3, 10, 7]).

Assumption 3.1. *The set*
$$S_1 = \{x \in \text{int } K \mid Ax - b = r^0\}$$

is bounded.

If the original problem has a feasible interior point then $S_0 \neq \emptyset$ and otherwise $S_0 = \emptyset$. In any cases, we have the following lemma.

Lemma 3.2. *Suppose that Assumption 3.1 holds.*

(i) *The set*
$$\cup_{\gamma \in [0,1]} S_\gamma$$

  is a bounded convex set.

(ii) *If $S_\gamma \neq \emptyset$ then the analytical center $z(\gamma)$ for f in S_γ, which is the minimizer of f over S_γ, uniquely exists and satisfies*
$$\bar{n}(z(\gamma)) = 0$$

  for any $\gamma \in [0, 1]$.

(iii)
$$\inf\{f(x) \mid x \in \cup_{\gamma \in [0,1]} S_\gamma\} > -\infty$$

Proof. (i): The convexity of the set $\cup_{\gamma \in [0,1]} S_\gamma$ directly follows from the definition and the convexity of K. Suppose that $\cup_{\gamma \in [0,1]} S_\gamma$ is unbounded. Then there exists a sequence $\{(\gamma^k, x^k)\}$ satisfying

$$\gamma^k \in [0, 1], \quad Ax^k - b = \gamma^k r^0, x^k \in \text{int } K, \quad \|x^k\| \to +\infty.$$

Let us consider the sequence
$$y^k := \frac{x^k}{\|x^k\|} \in \text{int } K.$$

Since $\{y^k\}$ and $\{\gamma^k\}$ are bounded and $\|y^k\| = 1$ for any k, by taking a subsequence if necessary, we find $\bar{y} \neq 0$ satisfying
$$A\bar{y} = 0, \quad \bar{y} \in K.$$

Thus, for any $\alpha > 0$, $x^0 + \alpha\bar{y}$ is in the set S_1 which contradicts the assumption that S_1 is bounded.

(ii): Since S_γ is bounded for any $\gamma \in [0, 1]$, by Theorem 2.2.8 of [11], the function $f(x)$ has a minimizer $z(\gamma) \in S_\gamma$. Since Karush-Kuhn-Tucker optimality conditions imply that $z(\gamma)$ satisfies

$$g(z(\gamma)) = A^* y,$$

it follows from the system of (11) and (12) that $\bar{n}(z(\gamma)) = 0$.

(iii): Since the function f is convex, it holds that

$$\begin{aligned} f(x) &\geq f(x^0) + \langle g(x^0), x - x^0 \rangle \\ &\geq f(x^0) - \|g(x^0)\| \|x - x^0\| \end{aligned}$$

for any x in the bounded convex set $\cup_{\gamma \in [0,1]} S_\gamma$. Thus the function $f(x)$ is bounded from below on $\cup_{\gamma \in [0,1]} S_\gamma$. $\qquad \square$

The assertion (ii) of Lemma 3.2 ensures that the value $\|\bar{n}(x)\|_x$ can be used to measure proximity to the analytical center. Moreover, it follows from (iii) of Lemma 3.2 that the value

$$\Delta := f(x^0) - \inf\{f(x) \mid x \in \cup_{\gamma \in [0,1]} S_\gamma\} \geq 0 \tag{15}$$

is well defined. It is not required to compute the value Δ in our algorithm, but plays an important role to estimate the iteration number of *centering steps* described below.

Let x be the current iterate of our algorithm and define $r := Ax - b$. In our algorithm, the residual vector r is decreasing on a line segment $(0, r^0]$ whenever $r^0 = Ax^0 - b \neq 0$ by using the shifted projected Newton direction given below (see (20)). So, we can take a $\gamma \in (0, 1]$ so that $r = \gamma r^0$.

If x is close to the analytical center $z(\gamma)$, e.g., $\|\bar{n}(x)\|_x \leq 1$ and γ is sufficiently small, i.e., $\gamma \leq \epsilon$, our algorithm will stop with a sufficiently approximated feasible interior point x. We may start several primal interior point algorithm with x as an initial point.

If $\|\bar{n}(x)\|_x \leq 1$ but $\gamma > \epsilon$, our algorithm will proceed to the *feasibility step* where we compute the shifted projected Newton direction $\bar{n}^r(x)$ which is the solution of

$$\begin{cases} H(x)\bar{n}^r(x) + g(x) = A^* y, \\ A\bar{n}^r(x) = -r. \end{cases} \tag{16}$$

Define

$$d^r(x) := -H(x)^{-1} A^* (AH(x)^{-1} A^*)^{-1} r. \tag{17}$$

Then it holds that

$$\bar{n}^r(x) = \bar{n}(x) + d^r(x). \tag{18}$$

We calculate the new iteration x^+ as

$$x^+ := x + \alpha \bar{n}^r(x), \quad \alpha := \frac{1 - \tau}{1 + \|d^r(x)\|_x} \in (0, 1 - \tau]. \tag{19}$$

Note that by (16), x^+ satisfies

$$Ax^+ - b = A(x + \alpha \bar{n}^r(x)) - b = (Ax - b) + \alpha A\bar{n}^r(x) = (1 - \alpha)r. \tag{20}$$

The above relation and the fact $\alpha \in (0, 1)$ ensure that the residual vector r is decreasing on the line segment $(0, r^0]$.

If $\|\bar{n}(x)\|_x > 1$ our algorithm will proceed to the *centering steps* described as in Theorem 2.2.3 of [8] until we find a point y such that $\|\bar{n}(y)\|_x \leq 1$.

We summarize our algorithm below.

A primal barrier function Phase I algorithm

Step 0: Choose an accuracy $\epsilon > 0$, a step size parameter $\tau \in (0, 1)$, and an initial point $x^0 \in \text{int}\, K$. Set $x := x^0$.

Step 1: Define $r := Ax - b$ and let $\gamma \in (0, 1]$ be the value satisfying $r = \gamma r^0$. If

$$\|\bar{n}(x)\|_x \leq 1, \quad \gamma \leq \epsilon$$

then stop. We have obtained a sufficiently approximated feasible interior point x. Here $\|\bar{n}(x)\|_x$ can be calculated by solving the system (11) – (12) and by (6).

Step 2: If
$$\|\bar{n}(x)\|_x \le 1, \quad \gamma > \epsilon$$
then proceed to the *feasibility step*. Compute the shifted projected Newton direction $\bar{n}^r(x)$ by solving (16). Set the new iterate x^+ as in (19). Set $x := x^+$. Go to **Step 1**.

Step 3: If
$$\|\bar{n}(x)\|_x > 1$$
then proceed to the *centering steps* described as in Theorem 2.2.3 of [8] until we find a point $y \in S_\gamma$ such that $\|\bar{n}(y)\|_x \le 1$. Set $x := y$. Go to **Step 1**.

Note that at the beginning of the centering steps, we have an interior initial point $x \in S_\gamma$ to find a point $y \in S_\gamma$ which is sufficiently close to the analytical center $z(\gamma)$ of S_γ. This fact implies that the *centering steps* in **Step 3** exactly corresponds to the ones in existing primal feasible algorithms as in [11] or in [8] and has been well analyzed. Thus we only have to know the behavior of the *feasibility step*. In the succeeding sections, we will analyze the *feasibility step* and give a sufficient condition for finding an approximate feasible interior point in a finite number of iterations.

4   Analysis of the feasibility step

In this section, we analyze the *feasibility step* in Section 3. Recall that K is a closed convex cone and f is a self-concordant function over $D_f := \operatorname{int} K$ with the complexity value ϑ_f. At the feasibility step, we choose a point x^+ as in (19). As we will see in Theorem 4.6, the value of the barrier function f will be increased at most $\mathcal{O}(\vartheta_f)$ at x^+.

Lemma 4.1. *At the feasibility step, we have*

(i)
$$\alpha\|\bar{n}^r(x)\|_x \le \alpha\left(\|\bar{n}(x)\|_x + \|d^r(x)\|_x\right) \le 1 - \tau.$$

(ii)
$$\begin{aligned} x^+ &:= x + \alpha\left(\bar{n}(x) + d^r(x)\right) \in B_x(x,1) \subseteq \operatorname{int} K, \\ &\quad x + \alpha\bar{n}(x) \in B_x(x,1) \subseteq \operatorname{int} K. \end{aligned}$$

(iii)
$$0 < \alpha\|d^r(x)\|_{x+\alpha\bar{n}(x)} \le \frac{\alpha\|d^r(x)\|_x}{1 - \alpha\|\bar{n}(x)\|_x} \le 1 - \tau.$$

(iv)
$$x^+ \in B_{x+\alpha\bar{n}(x)}(x + \alpha\bar{n}(x), 1) \subseteq \operatorname{int} K.$$

Proof. (i): The first inequality follows from (18). Since $\|\bar{n}(x)\|_x \le 1$ holds at **Step 2**, by the definition (19) of α, we have
$$\alpha\left(\|\bar{n}(x)\|_x + \|d^r(x)\|_x\right) \le \alpha\left(1 + \|d^r(x)\|_x\right) \le 1 - \tau.$$
(ii): By the definitions (19) and (18), we see that
$$\|x^+ - x\|_x = \alpha\|\bar{n}^r(x)\|_x = \alpha\|\bar{n}(x) + d^r(x)\| \le \alpha\left(\|\bar{n}(x)\|_x + \|d^r(x)\|_x\right).$$

8

Thus, the assertion (i) and the property (7) imply that $x^+ \in B_x(x, 1) \subseteq D_f = \operatorname{int} K$. By a similar discussion, we also see that

$$\alpha \|\bar{n}(x)\|_x \leq 1 - \tau \text{ and } x + \alpha \bar{n}(x) \in B_x(x, 1) \subseteq D_f = \operatorname{int} K.$$

(iii): Note that $\|d^r(x)\| > 0$ whenever the algorithm enters the feasibility step. Since $x + \alpha \bar{n}(x) \in D_f = \operatorname{int} K$ by (ii) above, by substituting $x + \alpha \bar{n}(x)$ and $d^r(x) \neq 0$ into y and v in (8), respectively, we see that

$$\frac{\|d^r(x)\|_{x+\alpha\bar{n}(x)}}{\|d^r(x)\|_x} \leq \frac{1}{1 - \alpha\|\bar{n}(x)\|_x}$$

or equivalently

$$0 < \|d^r(x)\|_{x+\alpha\bar{n}(x)} \leq \frac{\|d^r(x)\|_x}{1 - \alpha\|\bar{n}(x)\|_x}$$

holds. Since the assertion (i) ensures that

$$\frac{\alpha\|d^r(x)\|_x}{1 - \|\alpha\bar{n}(x)\|_x} \leq 1 - \tau,$$

we obtain the assertion (iii).

(iv): We see that

$$\|x^+ - (x + \alpha\bar{n}(x))\|_{x+\alpha\bar{n}(x)} = \|\alpha d^r(x)\|_{x+\alpha\bar{n}(x)}.$$

Thus the assertion follows from (iii) and the property (7). □

We will derive an upper bound of $f(x^+)$ using the above and the results in Section 2.

First we observe the relation between $f(x^+)$ and $f(x + \alpha\bar{n}(x))$. Since $x^+ \in B_{x+\alpha\bar{n}(x)}(x + \alpha\bar{n}(x), 1)$ by (iv) of Lemma 4.1, by substituting $x^+ := x + \alpha\bar{n}(x) + \alpha d^r(x)$ and $x + \alpha\bar{n}(x)$ into y and x in Theorem 2.4, we obtain the following inequality:

$$\begin{align}
f(x_+) \quad &\leq \quad f(x + \alpha\bar{n}(x)) \tag{21}\\
&\quad + \langle g(x + \alpha\bar{n}(x)), \alpha d^r(x)\rangle \tag{22}\\
&\quad + \frac{1}{2}\langle \alpha d^r(x), H(x + \alpha\bar{n}(x))\alpha d^r(x)\rangle \tag{23}\\
&\quad + \frac{\alpha^3\|d^r(x)\|_{x+\alpha\bar{n}(x)}^3}{3(1 - \alpha\|d^r(x)\|_{x+\alpha\bar{n}(x)})}. \tag{24}
\end{align}$$

We derive the following bounds for each term in (21) - (24).

Lemma 4.2 (An upperbound of the term in (21)).

$$f(x + \alpha\bar{n}(x)) \leq f(x) + \frac{(1-\tau)^2}{\tau}$$

Proof. By (ii) of Lemma 4.1, we know that $x + \alpha\bar{n}(x) \in B_x(x, 1)$. Thus by Theorem 2.4, it holds that

$$
\begin{aligned}
f(x + \alpha\bar{n}(x)) \quad &\leq \quad f(x) + \langle g(x), \alpha\bar{n}(x)\rangle + \frac{1}{2}\langle \alpha\bar{n}(x), H(x)\alpha\bar{n}(x)\rangle + \frac{\alpha^3\|\bar{n}(x)\|_x^3}{3(1 - \alpha\|\bar{n}(x)\|_x)} \\
&= \quad f(x) - \langle n(x), \alpha\bar{n}(x)\rangle_x + \frac{1}{2}\|\alpha\bar{n}(x)\|_x^2 + \frac{\alpha^3\|\bar{n}(x)\|_x^3}{3(1 - \alpha\|\bar{n}(x)\|_x)} \\
&\qquad \text{(by (6) and (10))} \\
&\leq \quad f(x) - \alpha\|\bar{n}(x)\|_x^2 + \frac{1}{2}\alpha^2\|\bar{n}(x)\|_x^2 + \frac{\alpha^3\|\bar{n}(x)\|_x^3}{3(1 - \alpha\|\bar{n}(x)\|_x)} \\
&\qquad \text{(by (13))} \\
&\leq \quad f(x) + \frac{1}{2}\alpha^2\|\bar{n}(x)\|_x^2 + \frac{\alpha^3\|\bar{n}(x)\|_x^3}{3(1 - \alpha\|\bar{n}(x)\|_x)}.
\end{aligned}
$$

Note that at the feasibility step, we have $\|\bar{n}(x)\|_x \leq 1$, $\alpha \leq 1 - \tau$ by (19), and

$$
\frac{1}{1 - \alpha\|\bar{n}(x)\|_x} \leq \frac{1}{1 - \alpha\left(\|\bar{n}(x)\|_x + \|d^r(x)\|_x\right)} \leq \frac{1}{\tau}
$$

by (i) of Lemma 4.1. These facts and $\tau \in (0, 1)$ imply that

$$
\begin{aligned}
\frac{1}{2}\alpha^2\|\bar{n}(x)\|_x^2 + \frac{\alpha^3\|\bar{n}(x)\|_x^3}{3(1 - \alpha\|\bar{n}(x)\|_x)} \quad &\leq \quad \frac{(1 - \tau)^2}{2} + \frac{(1 - \tau)^3}{3\tau} \\
&< \quad \frac{(1 - \tau)^2}{2\tau} + \frac{(1 - \tau)^2}{3\tau} \\
&< \quad \frac{(1 - \tau)^2}{\tau}.
\end{aligned}
$$

Thus we obtain the lemma. $\qquad\square$

Lemma 4.3 (An upperbound of the term in (22))**.**

$$
\langle g(x + \alpha\bar{n}(x)), \alpha d^r(x)\rangle \leq \vartheta_f + \frac{(1 - \tau)^2}{\tau}.
$$

Proof. In what follows, we denote $H(x)^{-1}g(x)$ by $g_x(x)$ for ease of notation. It follows from (10) that

$g_x(x) = -n(x)$ and $\|g_x(x)\|_x = \|n(x)\|_x$. Then, by the definition (6), we have

$$
\begin{aligned}
&\langle g(x + \alpha\bar{n}(x)), \alpha d^r(x)\rangle \\
&= \langle g_x(x + \alpha\bar{n}(x)), \alpha d^r(x)\rangle_x \\
&\leq \|\alpha d^r(x)\|_x \|g_x(x + \alpha\bar{n}(x))\|_x \\
&= \|\alpha d^r(x)\|_x \left(\left\| g_x(x) + \alpha\bar{n}(x) + \alpha\bar{n}(x) \int_0^1 (H_x(x + t\alpha\bar{n}(x)) - I)dt \right\|_x \right) \\
&\qquad \text{(by Proposition 2.1)} \\
&\leq \|\alpha d^r(x)\|_x \left(\|g_x(x)\|_x + \|\alpha\bar{n}(x)\|_x + \|\alpha\bar{n}(x)\|_x \int_0^1 \|H_x(x + t\alpha\bar{n}(x)) - I\|_x dt \right) \\
&\leq \|\alpha d^r(x)\|_x \left(\|g_x(x)\|_x + \|\alpha\bar{n}(x)\|_x + \|\alpha\bar{n}(x)\|_x \int_0^1 \left(\frac{1}{(1 - t\|\alpha\bar{n}(x)\|_x)^2} - 1 \right) dt \right) \\
&\qquad \text{(by Theorem 2.3)} \\
&= \|\alpha d^r(x)\|_x \left(\|n(x)\|_x + \|\alpha\bar{n}(x)\|_x + \|\alpha\bar{n}(x)\|_x \left[\frac{1}{\|\alpha\bar{n}(x)\|_x - \|\alpha\bar{n}(x)\|_x^2 t} - t \right]_0^1 \right) \\
&= \|\alpha d^r(x)\|_x \left(\|n(x)\|_x + \|\alpha\bar{n}(x)\|_x + \frac{\|\alpha\bar{n}(x)\|_x^2}{1 - \|\alpha\bar{n}(x)\|_x} \right) \\
&= \|\alpha d^r(x)\|_x \left(\|n(x)\|_x + \frac{\|\alpha\bar{n}(x)\|_x}{1 - \|\alpha\bar{n}(x)\|_x} \right) \\
&= \alpha\|d^r(x)\|_x \left(\|n(x)\|_x + \frac{\alpha\|\bar{n}(x)\|_x}{1 - \alpha\|\bar{n}(x)\|_x} \right).
\end{aligned}
$$

Note that at the feasibility step, we have $\alpha\|\bar{n}(x)\|_x \leq 1 - \tau$ by (19), and

$$
\alpha\|d^r(x)\|_x \leq 1 - \tau,
$$
$$
\frac{1}{1 - \alpha\|\bar{n}(x)\|_x} \leq \frac{1}{1 - \alpha(\|\bar{n}(x)\|_x + \|d^r(x)\|_x)} \leq \frac{1}{\tau}
$$

by (i) of Lemma 4.1. The above facts ensure that

$$
\begin{aligned}
\alpha\|d^r(x)\|_x \left(\|n(x)\|_x + \frac{\alpha\|\bar{n}(x)\|_x}{1 - \alpha\|\bar{n}(x)\|_x} \right) &\leq (1 - \tau)\left(\vartheta_f + \frac{1 - \tau}{\tau} \right) \\
&\leq \vartheta_f + \frac{(1 - \tau)^2}{\tau}.
\end{aligned}
$$

and we obtain the lemma. $\qquad\square$

Lemma 4.4 (An upperbound of the term in (23))**.**

$$
\frac{1}{2}\langle \alpha d^r(x), H(x + \alpha\bar{n}(x))\alpha d^r(x)\rangle \leq \frac{(1 - \tau)^2}{2\tau^2}.
$$

Proof. Recall that $H_x(y)$ is defined by (9). By the definition (6), we see that

$$\frac{1}{2}\langle \alpha d^r(x), H(x+\alpha\bar{n}(x))\alpha d^r(x)\rangle$$

$$= \frac{1}{2}\langle \alpha d^r(x), H_x(x+\alpha\bar{n}(x))\alpha d^r(x)\rangle_x$$

$$\leq \frac{1}{2}\|\alpha d^r(x)\|_x^2\|H_x(x+\alpha\bar{n}(x))\|_x$$

$$\leq \frac{1}{2}\|\alpha d^r(x)\|_x^2\frac{1}{(1-\|\alpha\bar{n}(x)\|_x)^2}$$

$$= \frac{1}{2}\left(\frac{\alpha\|d^r(x)\|_x}{1-\alpha\|\bar{n}(x)\|_x}\right)^2$$

where the last inequality follows from $x+\alpha\bar{n}(x) \in B_x(x,1)$ by (ii) of Lemma 4.1 and from Theorem 2.3. As we have seen in the proof of Lemma 4.3, it holds that

$$\alpha\|d^r(x)\|_x \leq 1-\tau \text{ and } \frac{1}{1-\alpha\|\bar{n}(x)\|_x} \leq \frac{1}{\tau}.$$

Therefore, we have

$$\left(\frac{\alpha\|d^r(x)\|_x}{1-\alpha\|\bar{n}(x)\|_x}\right)^2 \leq \left(\frac{1-\tau}{\tau}\right)^2$$

which completes the proof. $\qquad\square$

Lemma 4.5 (An upperbound of the term in (24)).

$$\frac{\alpha^3\|d^r(x)\|_{x+\alpha\bar{n}(x)}^3}{3(1-\alpha\|d^r(x)\|_{x+\alpha\bar{n}(x)})} \leq \frac{(1-\tau)^3}{3\tau}.$$

Proof. The definition of the local norm (6) and the assertion (iii) of Lemma 4.1 ensure that

$$\alpha\|d^r(x)\|_{x+\alpha\bar{n}(x)} \leq 1-\tau.$$

Thus, it immediately follows that

$$\frac{\alpha^3\|d^r(x)\|_{x+\alpha\bar{n}(x)}^3}{3(1-\alpha\|d^r(x)\|_{x+\alpha\bar{n}(x)})} \leq \frac{(1-\tau)^3}{3\tau}$$

which completes the proof. $\qquad\square$

Combining the results in Lemmas 4.2 to 4.5, we obtain the following theorem which gives an upperbound of the value $f(x^+)$.

Theorem 4.6. *At the feasibility step, the new iterate $x^+ \in D_f = \text{int } K$ satisfies*

$$f(x^+) \leq f(x) + \vartheta_f + 3\left(\frac{1-\tau}{\tau}\right)^2.$$

Thus

$$f(x^+) - f(x) = \mathcal{O}(1+\vartheta_f).$$

Proof. By replacing the terms in (21) – (24) by the upperbounds in Lemmas 4.2 to 4.5, we have

$$
\begin{aligned}
f(x_+) \quad &\leq \quad f(x) + \frac{(1-\tau)^2}{\tau} \\
&\quad + \vartheta_f + \frac{(1-\tau)^2}{\tau} \\
&\quad + \frac{(1-\tau)^2}{2\tau^2} \\
&\quad + \frac{(1-\tau)^3}{3\tau} \\
&\leq \quad f(x) + \vartheta_f + 3\left(\frac{1-\tau}{\tau}\right)^2.
\end{aligned}
$$

Since $\tau \in (0,1)$, we obtain the theorem. $\qquad\square$

5 Finite termination of the algorithm

In this section, we analyze the total number of steps of the primal barrier function Phase I algorithm proposed in Section 3. We impose the following assumption.

Assumption 5.1. *There exists $\delta < +\infty$ such that*

$$
\sup\left\{\|d^r(x)\|_x \mid x \in \cup_{\gamma \in [0,1]} S_\gamma, \ \|\bar{n}(x)\|_x \leq 1\right\} \leq \delta.
$$

Here $\|\cdot\|_x$, $d^r(x)$, S_γ and $\bar{n}(x)$ are defined by (6), (17), (14) and (11) – (12), respectively.

Note that by the definitions $\|\cdot\|_x$ and $d^r(x)$, $\|d^r(x)\|_x$ can be represented equivalently as

$$
\|d^r(x)\|_x = \sqrt{\langle d^r(x), H(x)d^r(x)\rangle} = \sqrt{\langle r, (AH(x)^{-1}A^*)^{-1}r\rangle}.
$$

In what follows, we show that Assumption 5.1 is a sufficient condition for finite termination of the algorithm. We obtain the following lemma under Assumption 5.1.

Lemma 5.2. *Suppose that Assumptions 3.1 and 5.1 hold. Then the maximum number of executions of* **Step 2** *(the feasibility step) is bounded by*

$$
\frac{1+\delta}{1-\tau} \log \frac{1}{\epsilon} = \mathcal{O}\left((1+\delta)\log\frac{1}{\epsilon}\right).
$$

Proof. We have seen that (20) holds at the feasibility step. It follows from the definition (19) that Assumption 5.1 ensures that

$$
\alpha \geq \frac{1-\tau}{1+\delta}. \tag{25}
$$

Note that the value of γ is not changed during centering steps. Thus, at the N-th iteration x^N of the feasibility step, we have

$$
Ax^N - b = \gamma^N r^0, \quad \gamma^N \leq \left(1 - \frac{1-\tau}{1+\delta}\right)^N.
$$

13

By the construction of the algorithm, we will stop if $\gamma^N \leq \epsilon$. Thus if

$$\left(1 - \frac{1-\tau}{1+\delta} \right)^N \leq \epsilon \tag{26}$$

then the algorithm will be terminated. By taking the logarithm of both sides of (26) and using the fact

$$\log(1 - \xi) \leq -\xi \ \text{ for any } \xi \in [0, 1)$$

we obtain

$$N \left(\frac{1-\tau}{1+\delta} \right) \geq \log \frac{1}{\epsilon}.$$

Thus we complete the proof. $\qquad\qquad\square$

The above lemma ensures that if Assumptions 3.1 and 5.1 hold then the original problem has an approximated feasible interior point. Using the above lemma, we obtain the following iteration bound of the algorithm.

Theorem 5.3. *Suppose that Assumptions 3.1 and 5.1 hold. Then the total number of steps of the primal barrier function Phase I algorithm is*

$$\mathcal{O}\left(\Delta(1+\delta)\log\frac{1}{\epsilon} + (1+\vartheta_f)\left((1+\delta)\log\frac{1}{\epsilon} \right)^2 \right)$$

where Δ is defined by (15).

Proof. It is well known that the number of the centering steps bounded by $\mathcal{O}(f^0 - f^*)$ where f^0 is the value of f at the beginning of **Step 3** and f^* is the value of f at the analytical center of f (see Theorem 2.2.3 of [8] or Section 3 of [9]).

In our algorithm, the value of the function f is increased only at **Step 2** of the feasibility step. As we have shown in Theorem 4.6, the increment is $\mathcal{O}(1 + \vartheta_f)$ at each feasibility step. Moreover, Lemma 5.2 ensures that the total increment throughout the algorithm is

$$\mathcal{O}\left((1 + \vartheta_f)(1+\delta)\log\frac{1}{\epsilon} \right).$$

Thus, the function value f^0 at the beginning of **Step 3** is

$$f^0 = f(x^0) + \mathcal{O}\left((1+\vartheta_f)(1+\delta)\log\frac{1}{\epsilon} \right).$$

At each **Step 3**, the value $f(z(\gamma))$ at the analytical center $z(\gamma) \in S_\gamma$ satisfies

$$f(z(\gamma)) \geq \inf\{ f(x) \mid x \in \cup_{\gamma \in [0,1]} S_\gamma \}$$

where the set S_γ is defined by (14). Thus, by the definition (15), we have

$$f^0 - f(z(\gamma)) \leq \Delta + \mathcal{O}\left((1+\vartheta_f)(1+\delta)\log\frac{1}{\epsilon} \right)$$

14

which ensures that the number of corrector steps in each **Step 3** is

$$\mathcal{O}\left(\Delta + (1 + \vartheta_f)(1 + \delta)\log\frac{1}{\epsilon}\right).$$

Let N be the maximum number of executions of **Step 2**. Here the maximum number of executions of **Step 3** is bounded by $N + 1$ since **Step 3** is never executed consecutively by the construction of the algorithm.

Thus, by Lemma 5.2 again, the total number of centering steps throughout the algorithm is

$$\mathcal{O}\left(\Delta(1 + \delta)\log\frac{1}{\epsilon} + (1 + \vartheta_f)\left((1 + \delta)\log\frac{1}{\epsilon}\right)^2\right)$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

By Proposition 2.6, we obtain the following corollary for the special case where K is the DNN cone $\mathcal{D}^n$.

Corollary 5.4. *Let K be the DNN cone $\mathcal{D}^n$ given by (2). Suppose that Assumptions 3.1 and 5.1 hold. Then the total number of steps of the primal barrier function Phase I algorithm is*

$$\mathcal{O}\left(\Delta(1 + \delta)\log\frac{1}{\epsilon} + n^2\left((1 + \delta)\log\frac{1}{\epsilon}\right)^2\right)$$

where Δ is defined by (15).

6   Concluding remarks

In this paper, we proposed a purely primal barrier function Phase I algorithm for solving nonsymmetric conic optimization problems including the DNN optimization problem. In order to develop efficient and practical algorithms for such problems, many issues remain to be investigated.

The first one is to find much clearer justification for finite termination of the algorithm. Assumption 5.1 gives a sufficient condition, but we have to figure out which instances satisfy the assumption. The second one is to provide more sophisticated algorithms. In our algorithm, we only use the self-concordant barrier function f and do not make any use of the objective function $\langle c, x \rangle$. Using the hybrid function

$$\eta\langle c, x\rangle + f(x)$$

instead of f will be a promising research direction for developing practical primal barrier algorithms. It is of course important to conduct numerical experiments to observe the behavior of the algorithms.

References

[1] R.E. Burkard, S.E. Karisch and F. Rendl, QAPLIB: A Quadratic Assignment Problem Library, http://www.seas.upenn.edu/qaplib/

[2] I. M. Bomze, M. Dür, E. De Klerk, C. Roos, A. J. Quist and T. Terlaky, On copositive programming and standard quadratic optimization problems, *Journal of Global Optimization*, vol. 18, pp.301-320, 2000.

[3] S. Burer, On the copositive representation of binary and continuous nonconvex quadratic programs, *Mathematical Programming*, vol. 120, pp.479-495, 2009.

[4] E. De Klerk and D. V. Pasechnik, Approximation of the stability number of a graph via copositive programming, *SIAM Journal on Optimization*, vol. 12, pp.875-892, 2002.

[5] E. De Klerk, D. V. Pasechnik, R. Sotirov, On semidefinite programming relaxations of the traveling salesman problem, *SIAM Journal on Optimization* vol. 19, pp.1559-1573, 2008.

[6] O. Güler, Hyperbolic polynomials and interior point methods for convex programming, *Mathematics of Operations Research*, vol. 22, pp.350–377, 1997.

[7] G. Iyengar, D.J. Phillips and C. Stein, Feasible and accurate algorithms for covering semidefinite programs, H. Kaplan (Ed.): *SWAT 2010*, LNCS 6139, pp.150–162, 2010.

[8] Yu.E. Neasterov and A.S. Nemirovskii, *Interior-Point Polynomial Algorithm in Convex Programming*, SIAM, Philadelphia, 1994.

[9] Yu.E. Nesterov, Towards nonsymmetric conic optimization, CORE Discussion Paper No. 2006/28, 2006.

[10] J. Povh and F. Rendl, Copositive and semidefinite relaxations of the quadratic assignment problem, *Discrete Optimization* vol. 16, pp.231–241, 2009.

[11] J. Renegar, *A Mathematical View of Interior Point Methods for Convex Optimization*, MPS-SIAM Series on Optimization, 2001.

[12] J. Renegar, Hyperbolic programs, and their derivative relaxations, *Foundations of Computational Mathematics* vol. 6, pp.59-79, 2005.

[13] SDPA Online Solver. http://sdpa.indsys.chuo-u.ac.jp/portal/

[14] A. Skajaa, J. B. Jørgensen and P. C. Hansen, On implementing a homogeneous interior-point algorithm for nonsymmetric conic optimization, IMM-Technical Report 2011-02, 2011.

[15] J. F. Sturm and S. Zhang. On cones of nonnegative quadratic functions. *Mathematics of Operations Research*, vol. 28, 2003, pp.246-267.

[16] A. Yoshise and Y. Matsukawa, On optimization over the doubly nonnegative cone, *Proceedings of 2010 IEEE Multi-conference on Systems and Control*, pp.13-19, 2010.