RAIRO Operations Research
RAIRO Oper. Res. **40** (2006) 303–323
DOI: 10.1051/ro:2006022

A NEW BARRIER FOR A CLASS OF SEMIDEFINITE PROBLEMS

Erik A. Papa Quiroz[1,*] and Paolo Roberto Oliveira[1, 2,**]

Abstract. We introduce a new barrier function to solve a class of Semidefinite Optimization Problems (SOP) with bounded variables. That class is motivated by some (SOP) as the minimization of the sum of the first few eigenvalues of symmetric matrices and graph partitioning problems. We study the primal-dual central path defined by the new barrier and we show that this path is analytic, bounded and that all cluster points are optimal solutions of the primal-dual pair of problems. Then, using some ideas from semi-analytic geometry we prove its full convergence. Finally, we introduce a new proximal point algorithm for that class of problems and prove its convergence.

Keywords. Interior point methods, barrier function, central path, semidefinite optimization.

1. Introduction

Various properties have been obtained through standard assumptions applied to the analysis of the primal-dual central path in Semidefinite Optimization (SO), using the classical logarithmic barrier. For example, it has been shown that the path is bounded and all cluster points are optimal solutions of the primal-dual pair of problems, see de Klerk *et al.* [8] and Luo *et al.* [10]. Convergence results and its characterization with respect to the analytic center of the optimal set are obtained when the strict complementarity condition is satisfied. Halická *et al.* [4]

Received April 4, 2005. Accepted March 28, 2006.

[1] Programa de Engenharia de Sistemas e Computacão, COPPE, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil; erik@cos.ufrj.br

[2] Rua Honorio, 1144 c 5, 20771-421, Rio de Janeiro, Brazil; poliveir@cos.ufrj.br

* *Ph.D. Student.*

** *Partially supported by CNPq/Brazil.*

show that the path does not always converge to the analytic center, and they gave a convergence proof, using some ideas from the algebraic sets.

In this paper we are interested in solving the special class of problems

$$\min_{X\in\mathcal{S}^n}\{C \bullet X : \mathcal{A}X = b, 0 \preceq X \preceq I\}$$

where $\mathcal{S}^n$ is the set of all real symmetric $n \times n$ matrices, $C \in \mathcal{S}^n$, $0 \preceq X \preceq I$ means that X and $I - X$ are positive semidefinite and the operator $\mathcal{A} : \mathcal{S}^n \to I\!R^m$ is defined as $\mathcal{A}X := (A_i \bullet X)_{i=1,\ldots,m}$ with $A_i \in \mathcal{S}^n$. The classical barrier solves this problem, by interior point methods, through $-\ln\det X - \ln\det(I - X)$. Convergence results of the central path, obtained by the KKT optimal conditions of the penalized problem with respect to that barrier, can be easily adapted from the general case when the problem has only the constraint $X \succeq 0$. In a different way, in this paper, we introduce the new barrier function

$$B(X) = \mathrm{Tr}[(2X - I)(\ln X - \ln(I - X))]$$

and study, essentially, convergence properties of the central path induced by B.

This work is a natural extension to (SO) of our recent paper Papa Quiroz and Oliveira [12], where we proposed the new self-concordant barrier applied to the hypercube $(0, 1)^n \subset I\!R^n$, given by $\sum_{i=1}^n(2x_i - 1)(\ln x_i - \ln(1 - x_i))$. Although it does not ensure a theoretical gain, in terms of the self-concordance property, we expect that the representation of the intrinsic structure of real problems that are naturally formulated in some hypercube, could lead to good computational performance. The same could be considered in the case of (SO), even if in the latter case, we are not able to ensure the self-concordance property. We think that the classes of problems presented in Section 3, which are naturally embedded in the set $0 \preceq X \preceq I$, also justify the proposed barrier.

The paper is divided as follows. In Section 1.1 we give preliminaries and notations that we will use along the paper and in Section 1.2 we present basic aspects of semi-analytic theory that will be used in the proof of the convergence of the path. In Section 2 we define the problem and we give the main assumptions. In Section 3 we present some examples of the class of problems studied in this paper. In Section 4 we present the new barrier and study the primal-dual central path obtaining its full convergence. Finally in Section 5 we introduce a new proximal point Bregman type algorithm with convergence result.

1.1. Preliminaries and notation

We denote by $I\!R^{n\times n}$ the vector space of all $n\times n$ real matrices and $\mathcal{S}^n$ the set of all real symmetric $n\times n$ matrices. For any $X, Y \in \mathcal{S}^n$, the inner product of X and Y is defined by $X \bullet Y := \mathrm{Tr}(XY)$ (the trace operator) and $||X||_F := (X \bullet X)^{1/2}$ is the Frobenius norm. We also denote by $\mathcal{S}^n_+$ the convex cone of symmetric positive semidefinite matrices and by $\mathcal{S}^n_{++}$ the cone of symmetric positive definite matrices.

Every matrix $X \in \mathcal{S}^n$ can be written in the form

$$X = Q^T D Q$$

where $D = \mathtt{diag}(\lambda_1(X), ..., \lambda_n(X))$, $\{\lambda_i(X)\}_{i=1,...,n}$ being the eigenvalues of X, and Q is the corresponding orthonormal matrix of the eigenvectors of X. Hence for any analytic scalar valued function g we can define a function of a matrix $X \in \mathcal{S}^n$ as

$$g(X) = Q^T g(D) Q$$

whenever the scalar functions $g(\lambda_i(X))$ are well defined, see for example Horn and Johnson, [5], Section 6.2. In particular we obtain for any $X \in \mathcal{S}^n_{++}$:

$$\ln X = Q^T D_{ln} Q$$

where $D_{ln} = \mathtt{diag}(\ln \lambda_1(X), ..., \ln \lambda_n(X))$. This matrix is the inverse of the exponential matrix exp, see [5]. Thus for any $B \in \mathcal{S}^n_{++}$, $\log B = A$ if and only if $B = \exp(A)$.

Finally, we introduce the operator $P \odot Q : I\!R^{n\times n} \to I\!R^{n\times n}$, where $P, Q \in I\!R^{n\times n}$, defined by

$$(P \odot Q)U := (1/2)(PUQ^T + QUP^T).$$

1.2. Basic aspects of semi-analytic sets

In this subsection we introduce the definition of (s)-analytic curves and the curve selection lemma that we are going to use in the proof of the convergence of the primal-dual central path. We refer the reader to Lojasiewicz, [9], for more details.

Definition 1.1. We say that a subset M of $I\!R^n$ is a semi-analytic set if it is determined by a finite number of equalities and inequalities of analytic functions; *i.e.* M can be written as

$$\bigcup_{j=1}^{k} \bigcap_{i=1}^{l} \{x : f_{ij}(x) \sigma_{ij} 0\},$$

where $k, l \in I\!N$, f_{ij} are analytic functions and σ_{ij} corresponds to one of the signs $\{<\}$, $\{>\}$ or $\{=\}$. If M has only equalities, ($\sigma_{ij} = 0$, for all i, j) then M is called an analytic set.

Definition 1.2. Let M be an analytic set of $I\!R^n$. A curve $\gamma \subset M$ is an (s)-analytic curve when it is the image of an analytic embedding of $(0, 1]$ into relatively compact and semi-analytic of M.

Lemma 1.1 (Curve selection lemma)**.** *If M is a semi-analytic set of $I\!R^n$ and if $a \in \overline{M}$ (the closure of M) is not an isolated point on M, then M contains an (s)-analytic curve that converges to the point a. Equivalently, there exists $\epsilon > 0$ and a real analytic curve $\gamma : [0, \epsilon) \to I\!R^n$ with $\gamma(0) = a$ and $\gamma(t) \in M$ for $t > 0$.*

Proof. See [9], Proposition 2, page 103. ∎

2. Definition of the problem and assumptions

We are interested in solving

$$\min_{X\in\mathcal{S}^n}\{C\bullet X : \mathcal{A}X = b, 0\preceq X\preceq I\} \tag{2.1}$$

where $X,\ C\in\mathcal{S}^n$, $\mathcal{A}:\mathcal{S}^n\to I\!R^m$ is a linear operator defined by $\mathcal{A}X := (A_i\bullet X)_{i=1}^m\in I\!R^m$, with $A_i\in\mathcal{S}^n$, $0\preceq X$ (respectively $X\preceq I$) meaning that X (respectively $I-X$) are positive semidefinite matrices.

We denote by

$$\mathcal{P}=\{X\in\mathcal{S}^n : \mathcal{A}X=b,\ 0\preceq X\preceq I\}$$

the feasible set of the problem (2.1) and its relative interior by

$$\mathcal{P}^0=\{X\in\mathcal{S}^n : \mathcal{A}X=b,\ 0\prec X\prec I\},$$

where $0\prec X$ (respectively $X\prec I$) means that X (respectively $I-X$) are positive definite matrices. The optimal solution set of the problem (2.1) is denoted by $\mathcal{P}^*$.

Due to the fact that the primal objective function is continuous on the compact set $\mathcal{P}$, it achieves a global minimum point on $\mathcal{P}$. Besides, the linearity (and therefore convexity) of the primal objective function implies that any local minimum is a global minimum. Thus, the set of optimal solutions of the problem (2.1), is a non-empty and bounded convex set.

The dual problem of (2.1) is:

$$\max_{(y,S,W)\in I\!R^m\times\mathcal{S}^n\times\mathcal{S}^n}\{b^Ty - W\bullet I : \mathcal{A}^*y+S-W=C, W\succeq 0, S\succeq 0\} \tag{2.2}$$

where $\mathcal{A}^*: I\!R^m\to\mathcal{S}^n$ is a linear operator defined by $\mathcal{A}^*y=\sum\limits_{i=1}^m y_iA_i$.

We denote by

$$\mathcal{D}=\{(y,S,W)\in I\!R^m\times\mathcal{S}^n\times\mathcal{S}^n :\ \ \mathcal{A}^*y+S-W=C,\ \ W\succeq 0, S\succeq\ 0\}$$

the feasible set of the problem (2.2) and

$$\mathcal{D}^0=\{(y,S,W)\in I\!R^m\times\mathcal{S}^n\times\mathcal{S}^n :\ \ \mathcal{A}^*y+S-W=C,\ \ W\succ 0, S\succ\ 0\}$$

its relative interior and by $\mathcal{D}^*$ the optimal solutions of the problem (2.2). We impose the following assumptions on the problem (2.1).

Assumptions:

(1) The set $\mathcal{P}^0$ is non-empty.
(2) The matrices A_i are linearly independent.

We point out that (1) is a standard assumption in interior point methods for Semidefinite Optimization and under the assumption (2) y is uniquely determined for (S,W).

Given X and (y, S, W) two feasible points of (2.1) and (2.2) respectively, the duality gap in objective values between X and (y, S, W) is:

$$\mathtt{gap} := C \bullet X - (b^T y - W \bullet I) = X \bullet S + (I - X) \bullet W \geq 0.$$

Also, due to assumption (1) and the fact that (2.1) has a solution, the dual problem (2.2) has also a solution and the duality gap is zero, see de Klerk, [7], Theorem 2.2, that is, if X^* is optimal in (2.1), then there exists (y^*, S^*, W^*) that is optimal for (2.2), with $C \bullet X^* = b^T y^* - W^* \bullet I$. Those solutions are characterized by the Karush-Kuhn-Tucker (KKT) conditions:

$$\begin{array}{rcl} \mathcal{A}^* y + S - W & = & C \\ \mathcal{A}X & = & b \\ XS & = & 0 \\ (I - X)W & = & 0 \\ S, W & \succeq & 0 \\ 0 \preceq & X & \preceq I. \end{array}$$

3. Examples

In this section we list some examples of problems found in the literature that can be written in the form (2.1).

Example 1. Finding the sum of the first few eigenvalues.

Overton and Womersley (Th. 3.4, p. 329, [11]) give the following characterization for the sum of the first k eigenvalues of a symmetric matrix:

$$\lambda_1(A) + ... + \lambda_k(A) \ = \ \max\{A \bullet X\colon \ \mathrm{Tr} X = k, \ 0 \preceq X \preceq I\}, \tag{3.3}$$

where $\lambda_j(A), j = 1, ..., k$, denote the jth largest eigenvalue of A. We may write (3.3) as $\max\{A \bullet X\colon \ \mathcal{A}X = b \ \ 0 \preceq X \preceq I\}$, where $\mathcal{A}X = I \bullet X$ and $b = k$. Therefore, (3.3) is a particular case of (2.1).

Example 2. Minimizing the sum of the first few eigenvalues.

We consider the minimization of the sum of the first k eigenvalues of a symmetric matrix:

$$\min \lambda_1(A(y)) + ... + \lambda_k(A(y)) \quad \text{where} \quad A(y) = A_0 + \sum_{i=1}^{m} y_i A_i, \ y \in I\!\!R^m. \tag{3.4}$$

It can be proved, see Alizadeh, [1] Theorem 4.3, that the dual version of this problem is

$$\max\{A_0 \bullet X\colon \ \mathrm{Tr} X = k, \ \ A_j \bullet X = 0 \ \text{ for } \ j = 1, ..., m, \ \ 0 \preceq X \preceq I\}. \tag{3.5}$$

Note that we may write (3.5) as $\max\{A_0 \bullet X : \ \mathcal{A}X = b, \ \ 0 \preceq X \preceq I\}$, where $\mathcal{A}X = (I \bullet X, A_1 \bullet X, ..., A_m \bullet X)^T$ and $b = (k, 0, 0, ...0) \in I\!R^{m+1}$. Therefore, that is an example of (2.1).

Example 3. Minimization of the weighted sum of eigenvalues.

Consider the following problem:

$$\min \ m_1\lambda_1(A) + ... + m_k\lambda_k(A) \quad \text{where} \quad m_1 \geq ... \geq m_k > 0. \tag{3.6}$$

Observe that due to condition $m_1 \geq ... \geq m_k > 0$, that problem is convex. Donath and Hoffman in [3] rewrote that sum as follows:

$$\begin{aligned} m_1\lambda_1 + ... + m_k\lambda_k = (m_1 - m_2)\lambda_1 + (m_2 - m_3)(\lambda_1 + \lambda_2) + ... \\ + (m_{k-1} - m_k)(\lambda_1 + ... + \lambda_{k-1}) + m_k(\lambda_1 + ... + \lambda_k). \end{aligned}$$

For each of the partial sums of eigenvalues in the right side of the equality we may use the SDP formulation from the example 2, obtaining:

$$\begin{array}{ll} \min & (m_1 - m_2)X_1 \bullet A + (m_2 - m_3)X_2 \bullet A + ... + m_k X_k \bullet A \\ \text{s.t} & \mathrm{Tr}(X_i) = i \quad \text{for} \ \ i = 1, ..., k \\ & 0 \preceq X_i \preceq I. \end{array}$$

Now observe that this problem can be written under form (2.1), where $C = \mathtt{diag}((m_1 - m_2)A_1, (m_2 - m_3)A_2, ..., m_k A_k)$, $X = \mathtt{diag}(X_1, X_2, ..., X_n)$, $\mathcal{A}(X) = I \bullet X$ and $b = (1, 2, ..., k)^T$. Thus (3.6) is a particular case of (2.1)

Example 4. The graph partitioning problem.

An important special case of the graph partitioning problem, see [1], can be written as

$$\min\{C \bullet X : \ X_{ii} = k/n, \ \ 0 \preceq X \preceq I\}. \tag{3.7}$$

We can write (3.7) as (2.1) with $A_1 = \mathtt{diag}(1, 0, ..., 0)$, $A_2 = \mathtt{diag}(0, 1, 0, ..., 0)$,..., $A_n = (0, 0, ..., 0, 1)$ and $b = (k/n, k/n, ..., k/n) \in I\!R^n$.

4. A new barrier and the central path

In this section we introduce a new barrier and propose a new interior penalized problem to solve (2.1) and study the behavior of the central path, obtained by its optimality conditions. We will observe that a property of this path is its primal feasibility, with respect to (2.1), and dual infeasibility, with respect to the dual problem (2.2), by a term $\mu[\ln X(\mu) - \ln(I - X(\mu))]$, that converges to zero when μ converges to zero, see Corollary 4.1. We prove that this path converges by using some ideas from semi-analytic geometry.

4.1. The new barrier

We start denoting the *matrix cube* as $\mathcal{S}^n_{[0,I]} = \{X \in \mathcal{S}^n : 0 \preceq X \preceq I\}$ and its interior as $\mathcal{S}^n_{(0,I)}$. Now, we define on $\mathcal{S}^n_{(0,I)}$ the following function:

$$B(X) = \mathrm{Tr}[(2X - I)(\ln X - \ln(I - X))]. \tag{4.1}$$

Such function generalizes for $\mathcal{S}^n_{(0,I)}$ the self-concordant barrier function for the hypercube introduced in [12]. We observe that $B(X)$ is invariant under the decomposition of the matrix X. Moreover, we can write $B(X)$ as:

$$B(X) = \sum_{i=1}^{n}(2\lambda_i(X) - 1)(\ln \lambda_i(X) - \ln(1 - \lambda_i(X))), \tag{4.2}$$

where $0 < \lambda_i(X) < 1$, $i = 1, ..., n$, are the eigenvalues of X. We have immediately the following properties:

- If $X \to \mathtt{bound}\mathcal{S}^n_{(0,I)}$ (X approaches the boundary, that is, $\lambda_i(X) \to 0$ or $\lambda_i(X) \to 1$) then $B(X) \to \infty$. Thus B is a barrier function on $\mathcal{S}^n_{(0,I)}$.
- $B(X) \geq 0$, for all $X \in \mathcal{S}^n_{(0,I)}$.
- B is a strictly convex function continuously differentiable on $\mathcal{S}^n_{(0,I)}$.

4.2. Definition of the central path

To solve the problem (2.1) we propose a new penalized problem:

$$\begin{array}{ll} \min & \phi_B(X,\mu) = C \bullet X + \mu \mathrm{Tr}[(2X - I)(\ln X - \ln(I - X))] \\ & \mathcal{A}X = b \\ & (0 \prec X \prec I) \end{array} \tag{4.3}$$

where $\mu > 0$ is a positive parameter. The first and second derivatives of $\phi_B(.,\mu)$ are:

$$\begin{aligned} \nabla\phi_B(X,\mu) &:= C + \mu[2(\ln X - \ln(I - X)) - X^{-1} + (I - X)^{-1}] \\ \nabla^2\phi_B(X,\mu) &:= 2\mu[X^{-1} + (I - X)^{-1}] + \mu(X^{-1} \odot X^{-1}) + \mu((I - X)^{-1} \odot (I - X)^{-1}), \end{aligned} \tag{4.4}$$

where

$$\begin{aligned} [X^{-1} + (I - X)^{-1}]Y &= X^{-1}Y + (I - X)^{-1}Y \\ (X^{-1} \odot X^{-1})Y &= X^{-1}YX^{-1} \\ ((I - X)^{-1} \odot (I - X)^{-1})Y &= (I - X)^{-1}Y(I - X)^{-1}. \end{aligned}$$

Since the objective function is linear and $B(X) = \mathrm{Tr}[(2X - I)(\ln X - \ln(I - X))]$ is strictly convex we have that $\phi_B(.,\mu)$ is strictly convex on the relative interior of the feasible set. In addition, that function takes infinite values on the boundary of $\mathcal{P}$. We conclude that $\phi_B(.,\mu)$ achieves the minimal value in its domain (for fixed

μ) at a unique point. The KKT (first order) optimality conditions for this problem are therefore necessary and sufficient, and are given by:

$$\begin{array}{rcl} \mathcal{A}^*y + S - W & = & C + 2\mu[\ln X - \ln(I - X)] \\ \mathcal{A}X & = & b \\ S - \mu X^{-1} & = & 0 \\ W - \mu(I - X)^{-1} & = & 0 \\ S, W & \succeq & 0 \\ 0 \prec & X & \prec I. \end{array} \tag{4.5}$$

The unique solution of this system, denoted by $(X(\mu), y(\mu), S(\mu), W(\mu))$, $\mu > 0$, we call the primal-dual central path. Clearly, this path is primal feasible and dual infeasible with respect to the problems (2.1) and (2.2) respectively. Furthermore, we call $(X(\mu))$ and $(y(\mu), S(\mu), W(\mu))$ the primal and dual central path respectively. The duality gap in that solution satisfies

$$\mathtt{gap}(\mu) := C \bullet X(\mu) - (b^T y(\mu) - W(\mu) \bullet I) = 2n\mu - 2\mu(\ln X(\mu) - \ln(I - X(\mu))) \bullet X(\mu).$$

4.3. Analyticity of the central path

Here we show that the analyticity of the central path follows from a straightforward application of the implicit function theorem.

Theorem 4.1. *The function $f_{cp} : \mu \to (X(\mu), y(\mu), S(\mu), W(\mu))$ is an analytic function for $\mu > 0$.*

Proof. The proof follows from the implicit function theorem if we show that the equations defining the central path are analytic, with a derivative (with respect to (X, y, S, W)) that is square and nonsingular at points on the path. Obviously the equations of the central path are analytic, so we should prove the second part.

Let $\Phi : \mathcal{S}^n \times I\!R^m \times \mathcal{S}^n \times \mathcal{S}^n \times I\!R \longrightarrow \mathcal{S}^n \times I\!R^m \times \mathcal{S}^n \times \mathcal{S}^n$ such that

$$\Phi(X, y, S, W, \mu) = \begin{pmatrix} C + 2\mu[\ln X - \ln(I - X)] - S + W - \mathcal{A}^*y \\ \mathcal{A}X - b \\ S - \mu X^{-1} \\ W - \mu(I - X)^{-1} \end{pmatrix}.$$

The derivative of Φ (with respect to (X, y, S, W)) is:

$$\Phi'(X, y, S, W) = \begin{pmatrix} 2\mu[X^{-1} + (I - X)^{-1}] & -\mathcal{A}^* & -I & I \\ \mathcal{A} & 0 & 0 & 0 \\ \mu(X^{-1} \odot X^{-1}) & 0 & I & 0 \\ -\mu((I - X)^{-1} \odot (I - X)^{-1}) & 0 & 0 & I \end{pmatrix},$$

where I denotes the identity operator. We have been rather loose in writing this in matrix form, since the blocks are operators rather than matrices, but the meaning is clear. We want to show that this derivative is nonsingular, and for this it suffices to prove that its null-space is the trivial one.

Let U, V, P any matrices in $I\!R^{n\times n}$ and $v \in I\!R^m$, consider the following system of equations:

$$
\begin{aligned}
2\mu[X^{-1} + (I-X)^{-1}]U - \mathcal{A}^*v - V + P &= 0 & (4.6)\\
\mathcal{A}U &= 0 & (4.7)\\
\mu(X^{-1} \odot X^{-1})U + V &= 0 & (4.8)\\
-\mu((I-X)^{-1} \odot (I-X)^{-1})U + P &= 0. & (4.9)
\end{aligned}
$$

From (4.8) and (4.9) we have

$$
\begin{aligned}
V &= -\mu[X^{-1} \odot X^{-1}]U & (4.10)\\
P &= \mu((I-X)^{-1} \odot (I-X)^{-1})U. & (4.11)
\end{aligned}
$$

Substituting (4.10) and (4.11) in (4.6), we obtain:

$$[2\mu[X^{-1} + (I-X)^{-1}] + \mu(X^{-1} \odot X^{-1}) + \mu((I-X)^{-1} \odot (I-X)^{-1})]U - \mathcal{A}^*v = 0.$$

Observe that the coefficient of U is the Hessian, see (4.4), then,

$$\nabla^2\phi_B(X,\mu)U - \mathcal{A}^*v = 0. \tag{4.12}$$

Applying $(\nabla^2\phi_B(X,\mu))^{-1}$ in (4.12)

$$U - (\nabla^2\phi_B(X,\mu))^{-1}\mathcal{A}^*v = 0. \tag{4.13}$$

Now, multiplying (4.13) by $\mathcal{A}$, and taking into consideration (4.7) gives

$$\mathcal{A}(\nabla^2\phi_B(X,\mu))^{-1}\mathcal{A}^*v = 0.$$

Since that A_i are linearly independent matrices we have

$$v = 0. \tag{4.14}$$

Substituting (4.14) in (4.12)

$$U = 0. \tag{4.15}$$

Finally, substituting (4.15) in (4.10) and (4.11), we obtain

$$V = 0 \quad \text{and} \quad P = 0. \tag{4.16}$$

From (4.14), (4.15) and (4.16) we conclude that $\Phi'(X, y, S, W, \mu)$ is nonsingular on the central path (and throughout $\mathcal{S}^n_{(0,I)} \times I\!R^m \times \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$). Thus the central path is indeed an analytical path. $\square$

4.4. The primal central path

In this subsection we present some properties of the primal central path. Those results are easy extensions from general nonnegative barriers in $\mathbb{R}^n$ to $\mathcal{S}^n$, but for the sake of completeness we give the proofs.

Lemma 4.1. *The function* $0 < \mu \to B(X(\mu))$*, where* B *is defined as in (4.1), is non increasing.*

Proof. Suppose $\mu_2 < \mu_1$. We will prove that $B(X(\mu_1)) \leq B(X(\mu_2))$. Since $X(\mu_1)$ minimizes $\phi_B(X, \mu_1)$ and $X(\mu_2)$ minimizes $\phi_B(X, \mu_2)$, see (4.3), we have:

$$C \bullet X(\mu_1) + \mu_1 B(X(\mu_1)) \leq C \bullet X(\mu_2) + \mu_1 B(X(\mu_2))$$

and

$$C \bullet X(\mu_2) + \mu_2 B(X(\mu_2)) \leq C \bullet X(\mu_1) + \mu_2 B(X(\mu_1)).$$

Adding those inequalities gives

$$\mu_2 B(X(\mu_2)) + \mu_1 B(X(\mu_1)) \leq \mu_2 B(X(\mu_1)) + \mu_1 B(X(\mu_2)),$$

that implies

$$(\mu_1 - \mu_2) B(X(\mu_1)) \leq (\mu_1 - \mu_2) B(X(\mu_2))$$

as $\mu_2 < \mu_1$ we have

$$B(X(\mu_1)) \leq B(X(\mu_2)).$$

□

Lemma 4.2. *If* $\mu_2 < \mu_1$ *then*

i. $\phi_B(X(\mu_2), \mu_2) \leq \phi_B(X(\mu_1), \mu_1)$;
ii. $C \bullet X(\mu_2) \leq C \bullet X(\mu_1)$.

Proof. Since $\mu_2 < \mu_1$, $X(\mu_2)$ minimizes $\phi_B(X, \mu_2)$ and $B \geq 0$:

$$\begin{aligned}
\phi_B(X(\mu_2), \mu_2) &= C \bullet X(\mu_2) + \mu_2 B(X(\mu_2)) \\
&\leq C \bullet X(\mu_1) + \mu_2 B(X(\mu_1)) \\
&\leq C \bullet X(\mu_1) + \mu_1 B(X(\mu_1)) \\
&= \phi_B(X(\mu_1), \mu_1).
\end{aligned}$$

This shows i. Let us consider the following inequality

$$C \bullet X(\mu_2) + \mu_2 B(X(\mu_2)) \leq C \bullet X(\mu_1) + \mu_2 B(X(\mu_1)).$$

Since $B(X(\mu_2)) \leq B(X(\mu_1))$, the inequality above gives

$$C \bullet X(\mu_2) \leq C \bullet X(\mu_1),$$

concluding the proof. □

Lemma 4.3. *If X^* is an optimal solution to the problem (2.1), and $\mu > 0$, then*

$$C \bullet X^* \leq C \bullet X(\mu) \leq \phi_B(X(\mu), \mu).$$

Proof. Since X^* is a optimal solution of the primal problem, we have that $C \bullet X^* \leq C \bullet X(\mu)$. As $B \geq 0$,

$$C \bullet X(\mu) \leq C \bullet X(\mu) + \mu B(X(\mu)),$$

that implies

$$C \bullet X(\mu) \leq \phi_B(X(\mu), \mu).$$

Then

$$C \bullet X^* \leq C \bullet X(\mu) \leq \phi_B(X(\mu), \mu). \qquad \square$$

Proposition 4.1. *All cluster points of the primal path $\{X(\mu)\}$ are optimal solutions of the problem (2.1).*

Proof. Let $\bar{X}$ be a cluster point of $\{X(\mu)\}$. Note that $\mathcal{A}\bar{X} = b$ and $0 \preceq \bar{X} \preceq I$. Let $\{\mu_k\}$ be a sequence of positive numbers such that

$$\lim_{k \to \infty} \mu_k = 0$$

and

$$\lim_{k \to \infty} X(\mu_k) = \bar{X}.$$

Fix $X \in \mathcal{P}$, an arbitrary feasible solution of the problem (2.1). Due to assumption 1, we can take X^0 primal feasible. Note that, for all $\epsilon \in (0, 1)$,

$$X(\epsilon) = (1 - \epsilon)X + \epsilon X^0 \in \mathcal{P}^0.$$

The optimality conditions for $\{X(\mu_k)\}$ give

$$C \bullet X(\mu_k) + \mu_k B(X(\mu_k)) \leq C \bullet X(\epsilon) + \mu_k B(X(\epsilon)).$$

That implies

$$\mu_k [B(X(\mu_k)) - B(X(\epsilon))] \leq C \bullet X(\epsilon) - C \bullet X(\mu_k).$$

As B is convex, we get

$$\mu_k [\nabla B(X(\epsilon))(X(\mu_k) - X(\epsilon))] \leq C \bullet X(\epsilon) - C \bullet X(\mu_k).$$

Taking limit when $k \to \infty$, leads to

$$0 \cdot [\nabla B(X(\epsilon)) \bullet (\bar{X} - X(\epsilon))] \leq C \bullet X(\epsilon) - C \bullet \bar{X},$$

that is

$$0 \leq C \bullet X(\epsilon) - C \bullet \bar{X}.$$

Now, taking $\epsilon \to 0$, we obtain

$$C \bullet \bar{X} \leq C \bullet X.$$

Since X is an arbitrary feasible solution, the last inequality implies that $\bar{X}$ is an optimal point of (2.1). □

Proposition 4.2. *Let μ_k be a sequence of positive real numbers such that $\mu_k \to 0$ as $k \to 0$. Then:*

$$\mu_k B(X(\mu_k)) \to 0, \quad \textit{as} \ \ k \to \infty.$$

Proof. We simplify the notation, letting $X(\mu_k)$ for X^k. Now, suppose that $\bar{X}$ is an optimal point of the primal problem (2.1). Let X^k a sequence such that

$$X^k \to \bar{X}, \ \text{ as } \ k \to \infty.$$

By continuity we have

$$C \bullet X^k \to C \bullet \bar{X}, \ \text{ as } \ k \to \infty. \tag{4.17}$$

From Lemma 4.3, we have

$$C \bullet \bar{X} \leq C \bullet X^k \leq \phi_B(X^k, \mu_k)$$

and therefore the sequence $\phi_B(X^k, \mu_k)$ is lower bounded. Besides, by Lemma 4.2, part ii, we know that ϕ_B is non increasing. So, there exists $\zeta^* \in I\!R^n$ such that

$$\phi_B(X^k, \mu_k) \to \zeta^* \ \text{ as } \ k \to \infty, \tag{4.18}$$

where $C \bullet \bar{X} \leq \zeta^*$.

Now from (4.17) and (4.18) we have

$$C \bullet X^k - \phi_B(X^k, \mu_k) \to C \bullet \bar{X} - \zeta^*, \ \text{ as } \ k \to \infty.$$

Since $\phi_B(X^k, \mu_k) = C \bullet X^k + \mu_k B(X^k)$, we have

$$\mu_k B(X^k) \to \zeta^* - C \bullet \bar{X}, \ \text{ as } \ k \to \infty.$$

As $B \geq 0$ and non increasing we should have

$$\mu_k B(X^k) \to 0, \ \text{ as } \ k \to \infty$$

since $\mu_k \to 0, \ \text{ as } \ k \to \infty.$ □

Corollary 4.1. *Let (X^k) be a sequence of the primal central path. Then:*

$$\mu_k \| \ln X^k - \ln(I - X^k) \|_F \ \to \ 0, \ \ \textit{as} \ \ k \ \to \ \infty.$$

Proof. From Proposition 4.2, we have

$$\mu_k \mathrm{Tr}[(2X^k - I)(\ln X^k - \ln(I - X^k))] \to 0, \ \ k \to \infty.$$

From (4.2), this is equivalent to

$$\mu_k \sum_{i=1}^{n} (2\lambda_i(X^k) - 1)(\ln \lambda_i(X^k) - \ln(1 - \lambda_i(X^k))) \to 0, \ \ k \to \infty,$$

where $\lambda_i(X^k)$, $i = 1, ..., n$, are the eigenvalues of X^k.

As $(2\lambda_i(X^k) - 1)(\ln \lambda_i(X^k) - \ln(1 - \lambda_i(X^k))) \geq 0$, that implies:

$$\mu_k(2\lambda_i^k - 1)(\ln \lambda_i^k - \ln(1 - \lambda_i^k)) \to 0, \ \ k \to \infty, \ \ \forall i = 1, ..., n,$$

where we use the notation λ_i to mean $\lambda_i(X^k)$. We rewrite the last limit as

$$2\mu_k(\lambda_i^k \ln \lambda_i^k + (1-\lambda_i^k)\ln(1-\lambda_i^k)) - \mu_k \ln \lambda_i^k - \mu_k \ln(1-\lambda_i^k) \to 0, \ \ k \to \infty, \ \ i = 1, ..., n. \tag{4.19}$$

The sequence $\alpha(\lambda_i^k) := \lambda_i^k \ln \lambda_i^k + (1 - \lambda_i^k)\ln(1 - \lambda_i^k)$ is bounded and since $\mu_k \to 0$ as $k \to \infty$, we obtain

$$2\mu_k \alpha(\lambda_i^k) \to 0, \ \ k \to \infty, i = 1, ..., n. \tag{4.20}$$

Subtracting the expression (4.20) from (4.19), as both are convergent sequences, it is true that

$$-\mu_k \ln \lambda_i^k - \mu_k \ln(1 - \lambda_i^k) \to 0, \ \ k \to \infty, \ \ i = 1, ..., n.$$

Besides, as $0 < \lambda_i^k < 1$ we have that $\ln \lambda_i^k < 0$ and $\ln(1 - \lambda_i^k) < 0$, $i = 1, ..., n$, so,

$$-\mu_k \ln \lambda_i^k \to 0, \ \ k \to \infty, \ \ i = 1, ...n, \tag{4.21}$$

and

$$-\mu_k \ln(1 - \lambda_i^k) \to 0, \ \ k \to \infty, \ \ i = 1, ..., n. \tag{4.22}$$

Therefore, adding (4.21) and (4.22) we conclude that

$$\mu_k(\ln \lambda_i^k - \ln(1 - \lambda_i^k)) \to 0, \ \ k \to \infty, \ \ i = 1, ..., n.$$

It follows that

$$\sum_{i=1}^{n} \mu_k^2(\ln \lambda_i(X^k) - \ln(1 - \lambda_i(X^k)))^2 \to 0, \ \ k \to \infty,$$

and therefore

$$\mu_k^2 \sum_{i=1}^{n} (\ln \lambda_i(X^k) - \ln(1 - \lambda_i(X^k)))^2 \to 0, \ \ k \to \infty,$$

which implies

$$\mu_k\|(\ln X^k - \ln(1-X^k))\|_F \to 0, \quad k \to \infty. \qquad \square$$

4.5. The primal-dual central path

Proposition 4.3. *For all $r > 0$ the set $\{(X(\mu), S(\mu), W(\mu)) : \ \mu \le r\}$ is bounded.*

Proof. As $X(\mu) \in \mathcal{S}^n_{(0,I)}$, the set $\{X(\mu)\}$ is bounded, in particular, when $\mu \le r$. Now, we will prove that $\{(S(\mu), W(\mu)) : \ \mu \le r\}$ is also bounded. We know that $(X(\mu), y(\mu), S(\mu), W(\mu))$ solves the equations:

$$\mathcal{A}^*y + S - W = C + 2\mu[\ln X - \ln(I-X)]$$

$$W = \mu(I-X)^{-1}$$

$$S = \mu X^{-1}.$$

So, defining the function on $\mathcal{S}^n_{(0,I)} \times I\!R^m \times \mathcal{S}^n \times \mathcal{S}^n$

$$L(X,y,S,W) = C \bullet X - y^T(\mathcal{A}X - b) - S \bullet X - W \bullet (I-X) + 2\mu \mathrm{Tr}[X \ln X + (I-X)\ln(I-X)]$$

we have

$$\nabla_X L(X(\mu), y(\mu), S(\mu), W(\mu)) = 0, \tag{4.23}$$

where ∇_X denotes the gradient of L with respect to X. Now, observing that $L(.\ , y(\mu), S(\mu), W(\mu))$ is a strictly convex function on $\mathcal{S}^n_{(0,I)}$, (4.23) implies that $X(\mu)$ is a unique minimum point of $L(X, y(\mu), S(\mu), W(\mu)))$ on $\mathcal{S}^n_{(0,I)}$. On the other hand, we have:

$$\begin{aligned}
&C \bullet X(\mu) - L(X(\mu), y(\mu), S(\mu), W(\mu)) = \\
&S(\mu) \bullet X(\mu) - W(\mu) \bullet (I - X(\mu)) - 2\mu[\mathrm{Tr}[X(\mu)\ln X(\mu) + (I - X(\mu))\ln(I - X(\mu))] \\
&\qquad = n\mu + n\mu - 2\mu \mathrm{Tr}[X(\mu)\ln X(\mu) + (I - X(\mu))\ln(I - X(\mu))] \le 2n\mu(1 + \ln 2),
\end{aligned}$$

where the second equality comes from (4.5).

Then, due to assumption 1, we can take X^0 primal feasible, getting:

$$\begin{aligned}
C \bullet X(\mu) - 2n\mu(1+\ln 2) &\le L(X(\mu), y(\mu), S(\mu), W(\mu)) \\
&\le L(X^0, y(\mu), S(\mu), W(\mu)) \\
&\le C \bullet X^0 - S(\mu) \bullet X^0 - W(\mu) \bullet (I - X^0), \qquad (4.24)
\end{aligned}$$

where the last inequality is a consequence of the feasibility of X^0, and the fact that $\sum_{i=1}^{n}[\lambda_i(X^0)\ln\lambda_i(X^0) + (1-\lambda_i(X^0))\ln(1-\lambda_i(X^0))] \le 0$.

Now, apply $\lambda_{\min}(X)\mathrm{Tr}S \leq X \bullet S$ and $(1-\lambda_{\min}(X))\mathrm{Tr}W \leq (I-X)\bullet W$ in (4.24), to get:

$$C \bullet X(\mu) - 2n\mu(1+\ln 2) \leq C \bullet X^0 - \lambda_{\min}(X^0)\mathrm{Tr}S(\mu) - (1-\lambda_{\min}(X^0))\mathrm{Tr}W(\mu).$$

Thus

$$\lambda_{\min}(X^0)\mathrm{Tr}S(\mu) + (1-\lambda_{\min}(X^0))\mathrm{Tr}W(\mu) \leq C \bullet X^0 + 2n\mu(1+\ln 2) - C \bullet X(\mu). \tag{4.25}$$

Since $0 < \lambda_i(X^0) < 1$ for all $i = 1, ..., n$, we have, in particular $0 < \lambda_{min}(X^0) < 1$. Thus the quantity ϵ defined by:

$$\epsilon = \min\{\lambda_{min}(X^0), 1 - \lambda_{min}(X^0)\}$$

is positive. Now, let X^* be an optimal point of the problem (2.1). From (4.25), and using $\mu \leq r$, we find:

$$\begin{aligned}\mathrm{Tr}S(\mu) + \mathrm{Tr}W(\mu) &\leq \frac{1}{\epsilon}[C \bullet X^0 + 2n\mu(1+\ln 2) - C \bullet X(\mu)] \\ &\leq \frac{1}{\epsilon}(C \bullet X^0 + 2n(1+\ln 2)r - C \bullet X^*).\end{aligned}$$

As $\|S\|_F \leq \mathrm{Tr}S$ and $\|W\|_F \leq \mathrm{Tr}W$ we have:

$$\|S(\mu)\|_F + \|W(\mu)\|_F \leq \frac{1}{\epsilon}[C \bullet X^0 + 2n\mu(1+\ln 2) - C \bullet X^*].$$

Therefore, the set $\{(X(\mu), S(\mu), W(\mu)) : \ \mu \leq r\}$ is bounded. □

Although the function $\ln(X(\mu)) - \ln(I - X(\mu))$ is not bounded in $\mathcal{S}^n_{(0,I)}$, it is possible to prove the existence of cluster point of the primal-dual central path.

Lemma 4.4. *The cluster points set of the primal-dual central path* $\{(X(\mu), y(\mu), S(\mu), W(\mu))\}$ *is nonempty.*

Proof. Because $\{(X(\mu), S(\mu), W(\mu)) : \ \mu \leq r\}$ is bounded, see previous proposition, there exists a sequence $\{(X(\mu_j), S(\mu_j), W(\mu_j))\}$ and a point $(\bar{X}, \bar{S}, \bar{W})$ such that:

$$(X(\mu_j), S(\mu_j), W(\mu_j)) \to (\bar{X}, \bar{S}, \bar{W}), j \to \infty, \ \text{ and}$$
$$\mu_j \to 0, j \to \infty.$$

From the first equation of the central path (4.5) we have:

$$y(\mu_j) = (\mathcal{A}\mathcal{A}^*)^{-1}\mathcal{A}\left(C + 2\mu_j[\ln(X(\mu_j) - \ln(I - X(\mu_j))] - S(\mu_j) + W(\mu_j)\right).$$

Taking limit when j goes to ∞ and using the Corollary 4.1 we have

$$y(\mu_j) \to (\mathcal{A}\mathcal{A}^*)^{-1}\mathcal{A}\left(C - \bar{S} + \bar{W}\right), j \to \infty.$$

Defining $\bar{y} := (\mathcal{A}\mathcal{A}^*)^{-1}\mathcal{A}\left(C - \bar{S} + \bar{W}\right)$, we obtain that $(\bar{X}, \bar{y}, \bar{S}, \bar{W})$ is a cluster point of the primal-dual central path. Therefore the proof is completed. □

Proposition 4.4. *All cluster points of the primal-dual central path are optimal solutions of the primal (2.1) and dual (2.2) pair of problems respectively.*

Proof. Let $(\bar{X}, \bar{y}, \bar{S}, \bar{W})$ be a cluster point of the primal-dual central path. Then there exist sequences $\{\mu_k\}$ and $(X(\mu_k), y(\mu_k), S(\mu_k), W(\mu_k))$ such that

$$\mu_k \to 0, k \to \infty, \quad \text{and}$$

$$(X(\mu_k), y(\mu_k), S(\mu_k), W(\mu_k)) \to (\bar{X}, \bar{y}, \bar{S}, \bar{W}), k \to \infty.$$

From Proposition 4.1, $\bar{X}$ is an optimal point of the primal problem (2.1), so we should prove that $(\bar{y}, \bar{S}, \bar{W})$ is an optimal point of the dual problem (2.2). From KKT conditions

$$\mathcal{A}^* y^k + S^k - W^k = C + 2\mu[\ln X^k - \ln(I - X^k)],$$

taking limit when k goes to infinity, using continuity property and Corollary 4.1, we have:

$$\mathcal{A}^* \bar{y} + \bar{S} - \bar{W} = C. \tag{4.26}$$

Clearly, as $S^k \succ 0$ and $W^k \succ 0$ it holds:

$$\bar{S} \succeq 0 \quad \text{and} \quad \bar{W} \succeq 0. \tag{4.27}$$

From (4.26) and (4.27) we conclude that $(\bar{y}, \bar{S}, \bar{W})$ is a feasible point of the problem (2.2). To verify that $(\bar{y}, \bar{S}, \bar{W})$ is an optimal solution it is sufficient to show that the complementarity property is satisfied, that is, $\bar{W} \bullet (\bar{X} - I) = 0$ and $\bar{S} \bullet \bar{X} = 0$.

We know that $X(\mu_k) \bullet S(\mu_k) = \mu_k$, from the KKT conditions. Now, taking limit when k goes to infinity and using continuity of X^k and S^k as functions of μ_k (see Th. 4.1), we arrive at:

$$\bar{X} \bullet \bar{S} = 0.$$

In a similar way we obtain

$$\bar{W} \bullet (I - \bar{X}) = 0.$$

Therefore, the proof is complete. □

4.6. Convergence of the primal-dual central path

We show that the primal-dual central path converges. To obtain this result we will use Lemma 1.1. The arguments are similar to those used for the logarithmic barrier, in Halická *et al.* [4].

Theorem 4.2. *The primal-dual Central Path $(X(\mu), y(\mu), S(\mu), W(\mu))$ converges to a point $(X^*, y^*, S^*, W^*) \in P^* \times D^*$.*

Proof. Using the existence result of cluster points of the sequence $(X(\mu), y(\mu), S(\mu), W(\mu))$ we have that there exists a point $(X^*, y^*, S^*, W^*) \in P^* \times D^*$ and a subsequence $\{\mu_j\}$ such that

$$\lim_{\mu_j \to 0} (X(\mu_j), y(\mu_j), S(\mu_j), W(\mu_j)) = (X^*, y^*, S^*, W^*).$$

Let us define the following set:

$$M = \left\{ (\bar{X}, \bar{y}, \bar{S}, \bar{W}, \mu) \colon \begin{array}{l} \mathcal{A}\bar{X} \ = 0 \\ \mathcal{A}^*\bar{y} + \bar{S} - \bar{W} \ = 2\mu[\ln(\bar{X} + X^*) \\ \qquad - \ln(I - (\bar{X} + X^*))] \\ (\bar{X} + X^*)(\bar{S} + S^*) = \mu I \\ (I - (\bar{X} + X^*))(\bar{W} + W^*) = \mu I \\ (\bar{X} + X^*) \succ 0, \ (I - (\bar{X} + X^*)) \succ 0, \\ \qquad (\bar{S} + S^*) \succ 0, \ (\bar{W} + W^*) \succ 0 \text{ and } \mu > 0. \end{array} \right\}.$$

It is easy to show that if there exists $(\bar{X}, \bar{y}, \bar{S}, \bar{W}, \mu) \in M$ then $(\bar{X} + X^*, \bar{y} + y^*, \bar{S} + S^*, \bar{W} + W^*)$ is a point on the primal-dual central path. Now we also have that the zero element is in the closure of M. Indeed, as

$$\lim_{\mu_j \to 0} (X(\mu_j), y(\mu_j), S(\mu_j), W(\mu_j)) = (X^*, y^*, S^*, W^*), \tag{4.28}$$

we can define the sequence

$$(\bar{X}_j, \bar{y}_j, \bar{S}_j, \bar{W}_j, \bar{\mu}_j) := (X(\mu_j) - X^*, \ y(\mu_j) - y^*, \ S(\mu_j) - S^*, \ W(\mu_j) - W^*, \mu_j).$$

From (4.28) we have immediately that $(\bar{X}_j, \bar{y}_j, \bar{S}_j, \bar{W}_j, \bar{\mu}_j) \in M$ and

$$\lim_{j \to \infty} (\bar{X}_j, \bar{y}_j, \bar{S}_j, \bar{W}_j, \bar{\mu}_j) = (0_{n\times n}, 0_m, 0_{n\times n}, 0_{n\times n}, 0).$$

Therefore, $0 \in \overline{M}$.

Now, the result follows by applying the curve selection lemma. To see this, observe that Lemma 1.1 implies the existence of an $\epsilon > 0$ and an analytic function $\gamma \colon [0, \epsilon) \to \mathcal{S}^n \times I\!\!R^m \times \mathcal{S}^n \times \mathcal{S}^n \times I\!\!R$ such that

$$\gamma(t) = (\bar{X}(t), \bar{y}(t), \bar{S}(t), \bar{W}(t), \mu(t)) \to (0_{n\times n}, 0_m, 0_{n\times n}, 0_{n\times n}, 0) \ \text{ as } \ t \to 0, \tag{4.29}$$

and if $t > 0$, $(\bar{X}(t), \bar{y}(t), \bar{S}(t), \bar{W}(t), \mu(t)) \in M$, that is,

$$\begin{array}{rcl} \mathcal{A}\bar{X}(t) & = & 0, \\ \mathcal{A}^*\bar{y}(t) + \bar{S}(t) - \bar{W}(t) & = & 2\mu[\ln(\bar{X}(t) + X^*) - \ln(I - (\bar{X}(t) + X^*))], \\ (\bar{X}(t) + X^*)(\bar{S}(t) + S^*) & = & \mu I, \\ (I - (\bar{X}(t) + X^*))(\bar{W}(t) + W^*) & = & \mu I, \\ \bar{X}(t) + X^* & \succ & 0, \\ I - (\bar{X}(t) + X^*) & \succ & 0, \\ \bar{S}(t) + S^* & \succ & 0, \\ \bar{W}(t) + W^* & \succ & 0, \\ \mu(t) & > & 0. \end{array} \tag{4.30}$$

Since the system that defines the central path (4.5) has a unique solution, the system (4.30) has also a unique solution, which is given by:

$$X(\mu(t)) = \bar{X}(t) + X^*, \quad y(\mu(t)) = \bar{y}(t) + y^*, \quad S(\mu(t)) = \bar{S}(t) + S^*, \\ W(\mu(t)) = \bar{W}(t) + W^*,$$

for $t > 0$. Now, by applying limit as $t \to 0$ and using (4.29), we obtain, for $\lim_{t\to 0} \mu(t) = 0$,

$$\lim_{t\to 0} X(\mu(t)) = X^*, \quad \lim_{t\to 0} y(\mu(t)) = y^*, \quad \lim_{t\to 0} S(\mu(t)) = S^*, \quad \lim_{t\to 0} W(\mu(t)) = W^*.$$

Since $\mu(t) > 0$ on $(0, \epsilon)$, $\mu(0) = 0$, and $\mu(t)$ is analytic on $[0, \epsilon)$, there exists an interval, say $(0, \epsilon')$ where $\mu'(t) > 0$. So, the inverse function $\mu^{-1} \colon \mu(t) \to t$ exists on the interval $(0, \mu(\epsilon'))$. Besides, $\mu^{-1}(t) > 0$ for all $t \in (0, \mu(\epsilon'))$, and $\lim_{t\to 0} \mu^{-1}(t) = 0$. Therefore, it follows that

$$\lim_{t\to 0} X(t) = \lim_{t\to 0} X(\mu(\mu^{-1}(t))) = \lim_{t\to 0} \bar{X}(\mu^{-1}(t)) + X^* = X^*.$$

Similarly, we get

$$\lim_{t\to 0} y(t) = y^*, \qquad \lim_{t\to 0} S(t) = S^*, \qquad \lim_{t\to 0} W(t) = W^*.$$

Since that (X^*, y^*, S^*, W^*) was arbitrary, it must be unique, concluding the proof. □

5. A proximal point algorithm

Iusem *et al.* [6] show, for a certain class of barriers in linear optimization, that the primal central path converges to the same point, as given by the proximal point method, associated with Bregman distances. This is a motivation for this section, where we obtain a similar result in (SO), with the proposed barrier.

We notice that Cruz Neto *et al.* [2] also develops that analysis in a framework that contains a large set of barriers.

Let introduce the proximal point algorithm to solve the problem (2.1) based on a Bregman "distance" induced by the barrier B, that is,

$$D_B(Z,Y) = B(Z) - B(Y) - \mathtt{tr}[\nabla B(Y)(Y - Z)]. \tag{5.31}$$

The algorithm generates a sequence $\{Z^k\} \subset \mathcal{S}^n_{(0,I)}$ defined as

$$Z^0 \in \mathcal{S}^n_{(0,I)} \text{ such that } \nabla B(Z^0) \in Im(A^*), \tag{5.32}$$

$$Z^{k+1} = \arg\min_{Z \in \mathcal{S}^n} \{C \bullet Z + \lambda_k D_B(Z, Z^k) : \mathcal{A}Z = b, 0 \preceq Z \preceq I\} \tag{5.33}$$

where $\{\lambda_k\} \subset I\!\!R_{++}$ satisfies

$$\sum_{k=0}^{\infty} \lambda_k^{-1} = \infty. \tag{5.34}$$

The sequence $\{Z^k\}$, $k = 1, 2, ...$ generated by the algorithm is well defined and unique (for each k) due to the strict convexity of the objective function on $\mathcal{P}^0$ and that it takes infinite values on the boundary of $\mathcal{P}$. Thus for all $k \geq 1$, $Z^k \in \mathcal{P}^0$. As Z^{k+1} is the solution of the problem (5.33) then there exists $v^k \in I\!\!R^m$ such that

$$C + \lambda_k(\nabla B(Z^{k+1}) - \nabla B(Z^k)) = \mathcal{A}^* v^k. \tag{5.35}$$

Now, it is easy to prove, using the assumption $\nabla B(Z^0) \in Im(\mathcal{A}^*)$, that the primal central path $X(\mu)$, as defined in the previous section, is the unique solution of the following problem

$$\min_{Z \in \mathcal{S}^n} \{C \bullet Z + \mu D_B(Z, Z^0) : \mathcal{A}Z = b, 0 \preceq Z \preceq I\}.$$

Thus $X(\mu)$ satisfies

$$C + \mu(\nabla B(X(\mu)) - \nabla B(Z^0)) = \mathcal{A}^* w(\mu) \tag{5.36}$$

for some $w(\mu) \in I\!\!R^m$.

Now, the relation between the sequence μ_k and λ_k is given by $\mu_k = (\sum_{j=0}^{k-1} \lambda_j^{-1})^{-1}$. As it will be clear in the proof of the next theorem, the divergence condition on λ_k implies the convergence of μ_k to 0, leading to the aimed result of an unique convergence point for both methods. The following result is obtained as a natural extension to semidefinite optimization from the results obtained by Iusem *et al.* [6].

Theorem 5.1. *The sequence $\{X^k\}$ generated by (5.32)-(5.33) and the primal central path converge to the same point.*

Proof. Let $\mu_k = (\sum_{j=0}^{k-1} \lambda_j^{-1})^{-1}$. Obviously $\{\mu_k\}$ is a decreasing sequence for each $k \geq 1$ and converges to zero when k goes to infinity. From (5.36) we have

$$C + \mu_k(\nabla B(X(\mu_k)) - \nabla B(Z^0)) = \mathcal{A}^* w(\mu_k)$$

$$C + \mu_{k+1}(\nabla B(X(\mu_{k+1})) - \nabla B(Z^0)) = \mathcal{A}^* w(\mu_{k+1}),$$

for some sequence $\{w(\mu_k)\}$.
Using the fact that $\mu_{k+1}^{-1} - \mu_k^{-1} = \lambda_k^{-1}$, the above equations imply that

$$C + \lambda_k(\nabla B(X(\mu_{k+1})) - \nabla B(X(\mu_k))) = \mathcal{A}^* v^k,$$

where

$$v^k = \lambda_k(\mu_{k+1}^{-1} w(\mu_{k+1}) - \mu_k^{-1} w(\mu_k)).$$

Now, due to the uniqueness of the minimum of (5.33), we get, from the last equation and (5.35), that $Z^k = X(\mu_k)$. Finally

$$\lim_{k\to\infty} Z^k = \lim_{k\to\infty} X(\mu_k) = X^*$$

where the last equality comes from Theorem 4.2. □

Remark 5.1. In both methods there is not a characterization of the convergence point. Particularly, we cannot affirm that it is an analytic center.

6. Concluding remarks

In this paper we proved that the main properties of the central path induced by the classical logarithmic barrier are also satisfied by the central path with respect to the new barrier $B(X) = \mathrm{Tr}[(2X - I)(\ln X - \ln(I - X))]$. Moreover, we introduced a new proximal point algorithm for solving semidefinite optimization problems with bounded variables, and we get the property that the convergence point is the same for that method and the central path. We cannot assure that property for a general proximal point algorithm. Nevertheless, there is a large class of barriers where similar results were obtained, see Cruz Neto *et al.* [2].

We observe that the self-concordance property, that we got in the $I\!R^n$ case, see [12], is an open question, consequently the polynomial complexity, too.

Furthermore, it is known that the main iteration of proximal algorithms is, in general a computationally difficult problem. The usual inexact techniques would need to be considered, in order to envisage a more tractable algorithm. We notice that, besides the using of the proposed barrier in primal-dual algorithms, it is worthwhile to verify its computational behavior, as a barrier that takes into account the structure of the "matrix hypercube" in an intrinsic way. Both are current works.

Acknowledgements. We are grateful to Professor Gregorio Malajovich, who presented us to the Lojasiewicz, [9], reference for the selection lemma for semi-analytic sets, and to the anonymous referees that suggested some important inclusions and changing.

References

[1] F. Alizadeh, Interior Point Methods in Semidefinite Programming with Applications to Combinatorial Optimization. *SIAM J. Optim.* **5** (1995) 13–51.

[2] J.X. da Cruz Neto, O.P. Ferreira, P.R. Oliveira and R.C.M. Silva, Central Paths in Semidefinite Programming, Generalized Proximal Point Method and Cauchy Trajectories in Riemannian Manifolds, submitted, `http://www.optimization-online.org/DB_FILE/2006/02/1327.pdf`

[3] W.E. Donath and A.J. Hoffman, Lower bounds for the partitioning of graphs. *IBM J. Res. Devel.* **17** (1973) 420–425.

[4] M. Halická, E. de Klerk and C. Roos, On the Convergence of the Central Path in Semidefinite Optimization. *SIAM J. Optim.* **12** (2002) 1090–1099.

[5] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis.* Cambridge University Press (1991).

[6] A.N. Iusem, B.S. Svaiter and J.X. da Cruz Neto, Central Paths, Generalized Proximal Point Methods and Cauchy Trajectories in Riemannian Manifolds. *SIAM J. Control Optim.* **37** (1999) 566–588.

[7] E. de Klerk, *Aspects of Semidefinite Programming.* Kluwer Academic Publisher (2002).

[8] E. de Klerk, C. Roos and T. Terlaky, Initialization in Semidefinite Programming Via a Self-Dual Skew-Symmetric Embedding. *Oper. Res Lett.* **20** (1997) 213–221.

[9] S. Lojasiewicz, *Ensembles semi-analytiques.* I.H.E.S., Bures-sur-Yvette (1965).

[10] Z.-Q. Luo, J.F. Sturm and S. Zhang, Superlinear Convergence of a Symmetric Primal-Dual Path Following Algorithm for Semidefinite Programming. *SIAM J. Optim.* **8** (1998) 59–81.

[11] M.L. Overton and R.S. Womersley, Optimality Conditions and Duality Theory for Minimizing Sums of the Largest Eigenvalues of Symmetric Matrices. *Math. Program.* **62** (1993) 321–357.

[12] E.A. Papa Quiroz and P.R. Oliveira, A New Self-concordant Barrier for the Hypercube. *J. Optim. Theory Appl. (JOTA)*, accepted.