REPORTS
IN
INFORMATICS

ISSN 0333-3590

Primal-Dual IPMS for Semidefinite Optimization Based on Finite Barrier Functions

M. El Ghami T. Steihaug C. Roos

REPORT NO 341 **December 2006**

Department of Informatics

UNIVERSITY OF BERGEN

Bergen, Norway

This report has URL http://www.ii.uib.no/publikasjoner/texrap/pdf/2006-341.pdf

Reports in Informatics from Department of Informatics, University of Bergen, Norway, is available at http://www.ii.uib.no/publikasjoner/texrap/.

Requests for paper copies of this report can be sent to:

Department of Informatics, University of Bergen, Høyteknologisenteret,
P.O. Box 7803, N-5020 Bergen, Norway

Primal-Dual IPMS for Semidefinite Optimization Based on Finite Barrier Functions *

M. El Ghami[1] T. Steihaug[1] C. Roos[2]

December 21, 2006

[1] *Department of Informatics, University of Bergen Høyteknologisenteret, N-5020, Bergen, Norway*

e-mail: melghami, Trond.Steihaug@ii.uib.no

[2] *Faculty of Electrical Engineering, Mathematics, and Computer Science Delft University of Technology*

P.O. Box 5031, 2600 GA Delft, The Netherlands

e-mail: C.Roos@ewi.tudelft.nl

Abstract In this paper we extend the results obtained for a class of finite kernel functions by Y.Q. Bai M. El Ghami and C.Roos published in SIAM Journal of Optimization, 13(3):766–782, 2003 [3] for linear optimization to semidefinite optimization. We show that the iteration bound for primal dual methods is $O(\sqrt{n}\log n \log \frac{n}{\epsilon})$, for large-update methods and $O(\sqrt{n}\log \frac{n}{\epsilon})$, for small-update methods. The iteration complexity obtained for semidefinite programming is the same as the best bound for primal-dual interior point methods in linear optimization.

Keywords. Interior-point; semidefinite optimization; primal-dual method.

AMS Subject Classification: **90C22** **90C31**

1 Introduction

A semidefinite optimization problem (SDO) is a convex optimization problem in the space of symmetric matrices. We consider the standard semidefinite programming problem

$$(SDP) \quad p^* = \inf_X \left\{ \mathrm{Tr}(CX) : \mathrm{Tr}(A_i X) = b_i (1 \le i \le m), X \succeq 0 \right\},$$

and its dual problem (SDD)

$$(SDD) \quad d^* = \sup_{y,S} \left\{ b^T y : \sum_{i=1}^{m} y_i A_i + S = C, S \succeq 0 \right\},$$

where C and A_i are symmetric $n \times n$ matrices, $b, y \in \mathbf{R}^m$, and $X \succeq 0$ means that X is symmetric positive semidefinite and $\mathrm{Tr}(A)$ denotes the trace of A (i.e., the sum of its diagonal elements). The matrices A_i are further assumed to be linearly independent (without loss of generality). Recall that for any two $n \times n$ matrices, A and B

$$\mathrm{Tr}(A^T B) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} B_{ij}.$$

*Will appear in proceedings of the conference FRANCORO V / ROADEF 2007, Grenoble, France.

Interior point methods ($IPMs$) provide a powerful approach for solving SDO problems. A comprehensive list of publications on this topic can be found in the SDO homepage maintained by Alizadeh [1]. The pioneering works in this direction are due to Alizadeh [1, 2] and Nesterov and Nemirovskii [9]. Most IPMs for SDO can be viewed as natural extensions of $IPMs$ for LO, and have similar polynomial complexity results. However, to obtain valid search directions is much more difficult than in the LO case. Below we describe how the usual search directions are obtained for primal-dual methods for solving SDO problems. Our aim is to show in this section that the kernel-function-based approach that we presented for LO in [3] can be applied also to SDO problems. For self-regular kernel functions this has been earlier in [11]. Just as in the LO case, the new methods have the same iteration complexity when small-updates are used, but the iteration complexity is better for large-updates methods.

1.1 Classical search direction

We assume that a strictly feasible pair ($X \succ 0, S \succ 0$) exists, which is the interior-point condition (IPC) for SDO. This ensures the existence of an optimal primal-dual pair (X^*, S^*) with zero duality gap. Hence one can write the optimality conditions for the primal-dual pair of problems as follows.

$$\begin{aligned} \mathrm{Tr}(A_i X) &= b_i, \quad i = 1, \dots, m \\ \sum_{i=1}^{m} y_i A_i + S &= C \\ XS &= 0 \\ X, S &\succeq 0. \end{aligned} \tag{1}$$

The basic idea of primal-dual $IPMs$ is to replace the above complementarity condition $XS = 0$ by the parameterized equation

$$XS = \mu E; \quad X, S \succ 0,$$

where E denotes the $n \times n$ identity matrix and $\mu > 0$. The resulting system has a unique solution for each $\mu > 0$. This solution is denoted by $(X(\mu), y(\mu), S(\mu))$ for each $\mu > 0$; $X(\mu)$ is called the μ-center of (SDP) and $(y(\mu), S(\mu))$ is the μ-center of (SDD). The set of μ-centers (with $\mu > 0$) defines a homotopy path, which is called *the central path* of (SDP) and (SDD) [5, 11]. The principal idea of $IPMs$ is to follow this central path and approach the optimal set of SDP as μ goes to zero. Newton's method amounts to linearizing the system (1), thus yielding the following system of equations.

$$\begin{aligned} \mathrm{Tr}(A_i \Delta X) &= b_i, \quad i = 1, \dots, m. \\ \sum_{i=1}^{m} \Delta y_i A_i + \Delta S &= 0 \\ X\Delta S + \Delta X S &= \mu E - XS. \end{aligned} \tag{2}$$

This so-called Newton system has a unique solution $(\Delta X, \Delta y, \Delta S)$. Note that ΔS is symmetric, due to the second equation in (2). However, a crucial point is that ΔX may be not symmetric. Many researchers have proposed various ways of 'symmetrizing' the third equation in the Newton system so that the new system has a unique symmetric solution. All these proposals can be described by using a symmetric nonsingular scaling matrix P and by replacing (2) by the system

$$\begin{aligned} \mathrm{Tr}(A_i \Delta X) &= b_i, \quad i = 1, \dots, m \\ \sum_{i=1}^{m} \Delta y_i A_i + \Delta S &= 0 \\ \Delta X + P\Delta S P^T &= \mu S^{-1} - X \end{aligned} \tag{3}$$

Now ΔX is automatically a symmetric matrix.

1.2 Nesterov-Todd direction

In this paper we consider the symmetrization schema of Nesterov-Todd [10]. So we use

$$P = X^{\frac{1}{2}} \left(X^{\frac{1}{2}} S X^{\frac{1}{2}} \right)^{-\frac{1}{2}} X^{\frac{1}{2}} = S^{-\frac{1}{2}} \left(S^{\frac{1}{2}} X S^{\frac{1}{2}} \right)^{\frac{1}{2}} S^{-\frac{1}{2}}, \tag{4}$$

where the last equality can be easily verified. Let $D = P^{\frac{1}{2}}$, where $P^{\frac{1}{2}}$ denotes the symmetric square root of P. Now, the matrix D can be used to scale X and S to the same matrix V, defined by [6, 14]:

$$V := \frac{1}{\sqrt{\mu}} D^{-1} X D^{-1} = \frac{1}{\sqrt{\mu}} DSD. \tag{5}$$

Obviously the matrices D and V are symmetric, and positive definite. Let us further define

$$\bar{A}_i := D A_i D, \quad i = 1, 2, \ldots, m;$$

and

$$D_X := \frac{1}{\mu} D^{-1} \Delta X D^{-1}; \quad D_S := \frac{1}{\mu} D \Delta S D \tag{6}$$

Then it follows from (3)

$$\begin{aligned} \operatorname{Tr}(\bar{A}_i D_X) &= 0, \quad i = 1, \ldots, m. \\ \sum_{i=1}^{m} \Delta y_i \bar{A}_i + D_S &= 0 \\ D_X + D_S &= V^{-1} - V. \end{aligned} \tag{7}$$

In the sequel, we use the following notational conventions. Throughout this paper, $\|\cdot\|$ denotes the 2-norm of a vector. The nonnegative and the positive orthants are denoted as $\mathbf{R}_+^n$ and $\mathbf{R}_{++}^n$, respectively, and $\mathbf{S^n}$, $\mathbf{S_+^n}$, and $\mathbf{S_{++}^n}$ denote the cone of symmetric, symmetric positive semidefinite and symmetric positive definite $n \times n$ matrices, respectively. For any $V \in \mathbf{S^n}$, we denote by $\lambda(V)$ the vector of eigenvalues of V arranged in increasing order, that is, $\lambda_1(V) \le \lambda_2(V) \le, \ldots, \lambda_n(V)$. For any matrix A, we denote by $\eta_1(A) \le \eta_2(A) \le, \ldots, \le \eta_n(A)$ the singular values of A; if A is symmetric, then one has $\eta_i(A) = |\lambda_i(A)|$, $i = 1, 2, \ldots, n$, if $z \in \mathbf{R}^n$ and $f : \mathbf{R} \to \mathbf{R}$, then $f(z)$ denotes the vector in $\mathbf{R}^n$ whose i-th component is $f(z_i)$, with $1 \le i \le n$, if D is a diagonal matrix, then $f(D)$ denotes a diagonal matrix with $f(D_{ii})$ as i diagonal component. For $X \in S^n$, $X = Q^{-1} D Q$, where Q is orthogonal, and D diagonal matrices, $f(X) = Q^{-1} f(D) Q$. Finally if v is a vector, $diag(v)$ denotes the diagonal matrix with the diagonal elements v_i.

2 New search direction

In this section we introduce the definition of a matrix function [8, 13].

Definition 2.1 *Let X be a symmetric matrix, and let*

$$X = Q_X^{-1} diag(\lambda_1(X), \lambda_2(X), \ldots, \lambda_n(X)) Q_X, \tag{8}$$

be an eigenvalue decomposition of X, where $\lambda_i(X)$, $1 \le i \le n$ denote the eigenvalues of X, and Q_X is orthogonal. If $\psi(t)$ is any univariante function whose domain contains $\{\lambda_i(X); 1 \le i \le n\}$ then the matrix function $\psi(X)$ is defined by

$$\psi(X) = Q_X^{-1} diag(\psi(\lambda_1(X)), \psi(\lambda_2(X)), \ldots, \psi(\lambda_n(X))) Q_X. \tag{9}$$

Define the barrier function $\Psi(X)$ as follows [11].

$$\Psi(X) := \sum_{i=1}^{n} \psi(\lambda_i(X)) = Tr(\psi(X)). \tag{10}$$

In this paper, when we use the function $\psi(.)$ and its first three derivatives $\psi'(.)$, $\psi''(.)$, and $\psi'''(.)$ without any specification, it denotes a matrix function if the argument is a matrix and a univariate function (from $\mathbf{R}$ to $\mathbf{R}$) if the argument is in $\mathbf{R}$.

Following [11] we describe the kernel-function-based approach to SDO. Given the kernel function $\psi(t)$ and the associated $\psi(V)$ and $\psi'(V)$ as defined in Definition 2.1, we replace the right-hand side $V - V^{-1}$ in the third equation in (7) by $-\psi'(V)$. Thus we consider the following system.

$$\begin{aligned} \text{Tr}(\bar{A}_i D_X) &= 0, \quad i = 1, \ \ldots, m. \\ \sum_{i=1}^{m} \Delta y_i \bar{A}_i + D_S &= 0 \\ D_X + D_S &= -\psi'(V). \end{aligned} \tag{11}$$

Having D_X and D_S, $\triangle X$ and $\triangle S$ can be calculated from (6). Due to the orthogonality of $\triangle X$ and $\triangle S$, it is trivial to see that $D_X \perp D_S$, and so

$$\text{Tr}(D_X D_S) = \text{Tr}(D_S D_X) = 0. \tag{12}$$

The algorithm considered in this paper is described in Figure 1.

Generic Primal-Dual Algorithm for SDO

Input:
  a kernel function $\psi(t)$;
  a threshold parameter $\tau > 0$;
  an accuracy parameter $\epsilon > 0$;
  a barrier update parameter θ, $0 < \theta < 1$;
begin
  $X := X_0$; $S := S_0$; $\mu := \mu_0$;
  while $n\mu \geq \epsilon$ **do**
  begin
    $\mu := (1 - \theta)\mu$;
    while $\Phi\,(X, S, \mu) \geq \tau$ **do**
    begin
      Solve system (3) for $\triangle X$, $\triangle y$, $\triangle S$;
      Determine a step size α;
      $X := X + \alpha\triangle X$;
      $S := S + \alpha\triangle S$;
      $y := y + \alpha\Delta y$;
    end
  end
end

Figure 1: Generic primal-dual interior-point algorithm for SDO.

Just as in the LO case, the parameters τ, θ, and the step size α should be chosen in such a way that the algorithm is 'optimized' in the sense that the number of iterations required by the algorithm is as small as possible. Obviously, the resulting iteration bound will depend on the kernel function underlying the algorithm, and our main task becomes to find a kernel function that minimizes the iteration bound.

The paper is organized as follows. In Section 3 we start by deriving some properties of the kernel function $\Psi(t)$. In Section 4 we derive the properties of the barrier function $\Psi(V)$. The estimate of the step size and the decrease behavior of the barrier function are discussed in Section 5. The total iteration bound of the algorithm and the complexity results are derived in Section 6. Finally, some concluding remarks follow in Section 7.

3 Properties of kernel functions

In this paper we consider a kernel functions analyzed in [3] namely,

$$\psi(t) = \frac{t^2 - 1}{2} + \frac{1}{\sigma}\left(e^{\sigma(1-t)} - 1\right) \quad \text{for some} \quad \sigma \geq 1. \tag{13}$$

Recall that all known kernel functions have the property that $\lim_{t\downarrow 0}\psi(t) = \infty$ and $\lim_{t\to\infty}\psi(t) = \infty$. Our new function has the second property, but it fails to have the first property, because

$$\lim_{t\downarrow 0}\psi(t) = \psi(0) = \frac{e^{\sigma} - 1}{\sigma} - \frac{1}{2} < \infty. \tag{14}$$

This means that if either X or S approaches the boundary of the feasible region then

$$\Phi(X, S; \mu) := 2\Psi(V),$$

converges to a finite value, depending on the value of σ. In the analysis of the algorithm based on the present kernel functions $\psi(t)$ we need its three derivatives. For ease of reference we give them here. One has

$$\psi'(t) = t - e^{\sigma(1-t)}, \quad \psi''(t) = 1 + \sigma e^{\sigma(1-t)}, \quad \psi'''(t) = -\sigma^2 e^{\sigma(1-t)} \tag{15}$$

It follows that $\psi(1) = \psi'(1) = 0$ and $\psi''(t) \geq 0$, proving that $\psi(t)$ is defined by $\psi''(t)$

$$\psi(t) = \int_1^t \int_1^{\xi} \psi''(\zeta)\, d\zeta d\xi. \tag{16}$$

Lemma 3.1 *Let ψ be as defined in (13). Then,*

$$t\psi''(t) + \psi'(t) > 0, \qquad \text{if} \quad t \geq \frac{1}{\sigma}, \tag{17-a}$$

$$\psi'''(t) < 0, \qquad \text{if} \quad t > 0, \tag{17-b}$$

$$\psi''(t)\psi'(\beta t) - \beta\psi'(t)\psi''(\beta t) > 0, \qquad \text{if} \quad t > 1, \quad \beta > 1. \tag{17-c}$$

Proof. Using (15) we write, also using $t \geq \frac{1}{\sigma}$,

$$\psi'(t) + t\psi''(t) = 2t + (t\sigma - 1)\, e^{-\sigma(t-1)} \geq \frac{2}{\sigma} > 0.$$

Thus (17-a) follows. Inequality (17-b) immediately follows from (15). By (15),

$$\psi''(t)\psi'(\beta t) - \beta\psi'(t)\psi''(\beta t) = \sigma(\beta - 1)\, e^{-\sigma(t-2+\beta t)} + g(\beta) \geq g(\beta),$$

where

$$g(\beta) = \beta(1 + \sigma t)\, e^{\sigma(1-t)} - (\beta\sigma t + 1)\, e^{-\sigma(\beta t - 1)}.$$

One has $g(1) = 0$ and

$$g'(\beta) = (1 + \sigma t)\, e^{\sigma(1-t)} + \beta(\sigma t)^2\, e^{\sigma(\beta t - 1)} \geq 0.$$

From this (17-c) follows. □ The first property (17-a) in Lemma 3.1 is related to Definition 2.1.1 and Lemma 2.1.2 in [11]. This property is equivalent to convexity of the composed function $z \mapsto \psi(e^z)$ and this holds if and only if $\psi(\sqrt{t_1 t_2}) \leq \frac{1}{2}(\psi(t_1) + \psi(t_2))$ for any $t_1, t_2 \geq \frac{1}{\sigma}$. Following [3], we therefore say that ψ is exponentially convex, or shortly, e-convex, whenever $t \geq \frac{1}{\sigma}$.

Lemma 3.2 *One has*

$$\psi(t) < \frac{1}{2}\psi''(1)\,(t-1)^2\,, \qquad \textit{if} \quad t > 1.$$

Proof. By using Taylor's theorem and $\psi(1) = \psi'(1) = 0$, we obtain

$$\psi(t) = \frac{1}{2}\psi''(1)\,(t-1)^2 + \frac{1}{6}\psi'''(\xi)\,(\xi-1)^3\,,$$

where $1 < \xi < t$ if $t > 1$. Since $\psi'''(\xi) < 0$. Thus the lemma follows. □

Lemma 3.3 *One has*

$$t\psi'(t) \geq \psi(t), \;\; \textit{if} \;\; t \geq 1.$$

Proof. Defining $g(t) := t\psi'(t) - \psi(t)$ one has $g(1) = 0$ and $g'(t) = t\psi''(t) \geq 0$. Hence $g(t) \geq 0$ and the lemma follows. □

Not that the second inequality in (17-a) ($t \geq \frac{1}{\sigma}$) is more restrictive than in [4], where we only assumed t nonnegative. This means that we must ensure that t is large enough, before using inequality (17-a). The same problem was encountered in [3], and we will see that we can deal with it in the same way as we did for linear case in [3].

We denote by $\varrho : [0, \infty) \rightarrow [1, \infty)$ and $\rho : [0, -\frac{1}{2}\psi'(0)) \rightarrow (0, 1]$ the inverse functions of $\psi(t)$ for $t \geq 1$, and $-\frac{1}{2}\psi'(t)$ for $t \leq 1$, respectively. In other words

$$s = \psi(t) \quad \Leftrightarrow \quad t = \varrho(s), \quad t \geq 1, \tag{18}$$

$$s = -\tfrac{1}{2}\psi'(t) \quad \Leftrightarrow \quad t = \rho(s), \quad t \leq 1. \tag{19}$$

We recall from [4, 7] two other theorems that we need in the next section. In the first theorem the function Ψ is applied to a positive vector v. The definition is compatible with Definition 2.1 when identifying the vector v with its diagonal matrix $\operatorname{diag}(v)$ and applying Ψ to this matrix to obtain

$$\Psi(v) = \sum_{i=1}^{n} \psi(v_i), \quad v \in \mathbf{R}^n_{++}.$$

The next theorem is due to the fact that $\psi(t)$ satisfies (17-c).

Theorem 3.4 (Theorem 3.2 in [4]) *For any positive vector v and any $\beta > 1$, we have*

$$\Psi(\beta v) \leq n\psi\left(\beta\varrho\left(\frac{\Psi(v)}{n}\right)\right).$$

The following theorem gives a lower bound of $\delta(v)$ in terms of $\Psi(v)$. This is due to the fact that $\psi(t)$ satisfies (17-b).

Theorem 3.5 (Theorem 4.9 in [4]) *One has*

$$\delta(v) \geq \tfrac{1}{2}\psi'\left(\varrho\left(\Psi(v)\right)\right).$$

4 Properties of $\Psi(V)$ and $\delta(V)$

In this section we extend two Theorem 3.4 and Theorem 3.5 to positive definite matrices. Recall that these theorems follows from (17-c) and (17-b) respectively. [4].

Theorem 4.1 *Let ϱ be as defined in (18). Then for any positive vector v and any $\beta > 1$ we have:*

$$\Psi(\beta V) \leq n\psi\left(\beta\varrho\left(\frac{\Psi(V)}{n}\right)\right).$$

Proof. Let $v_i := \lambda_i(V)$, $1 \leq i \leq n$. Then $v > 0$ and

$$\Psi(\beta V) = \sum_{i=1}^{n}\psi(\beta\lambda_i(V)) = \sum_{i=1}^{n}\psi(\beta v_i) = \Psi(\beta v).$$

Using Theorem 3.4 we get

$$\begin{aligned}\Psi(\beta v) \leq n\psi\left(\beta\varrho\left(\frac{\Psi(v)}{n}\right)\right) &= n\psi\left(\beta\varrho\left(\frac{\sum_{i=1}^{n}\psi(v_i)}{n}\right)\right)\\ &= n\psi\left(\beta\varrho\left(\frac{\sum_{i=1}^{n}\psi(\lambda_i(V))}{n}\right)\right)\\ &= n\psi\left(\beta\varrho\left(\frac{\Psi(V)}{n}\right)\right).\end{aligned}$$

This proves the theorem. □

The next theorem gives a lower bound on the norm-based proximity measure $\delta(V)$, defined by

$$\delta(V) = \tfrac{1}{2}\|\psi'(V)\| = \frac{1}{2}\sqrt{\sum_{i=1}^{n}\psi'(\lambda_i(V))^2} = \frac{1}{2}\|D_X + D_S\|, \tag{20}$$

in terms of $\Psi(V)$. Since $\Psi(V)$ is strictly convex and attains its minimal value zero at $V = E$, we have

$$\Psi(V) = 0 \quad\Leftrightarrow\quad \delta(V) = 0 \quad\Leftrightarrow\quad V = E.$$

Theorem 4.2 *Let ϱ be as defined in (18). Then*

$$\delta(V) \geq \tfrac{1}{2}\psi'\left(\varrho\left(\Psi(V)\right)\right).$$

Proof. The statement in the lemma is obvious if $V = E$ since then $\delta(V) = \Psi(V) = 0$. Otherwise we have $\delta(V) > 0$ and $\Psi(V) > 0$. To deal with the nontrivial case. Again, let $v_i := \lambda_i(V)$, $1 \leq i \leq n$. Then $v > 0$ and

$$\delta(V)^2 = \tfrac{1}{4}\sum_{i=1}^{n}\psi'(\lambda_i(V))^2 = \tfrac{1}{4}\sum_{i=1}^{n}\psi'(v_i)^2 = \delta(v)^2.$$

Using Theorem 3.5 we get

$$\begin{aligned}\delta(v) \geq \tfrac{1}{2}\psi'\left(\varrho\left(\Psi(v)\right)\right) &= \tfrac{1}{2}\psi'\left(\varrho\left(\sum_{i=1}^{n}\psi(v_i)\right)\right)\\ &= \tfrac{1}{2}\psi'\left(\varrho\left(\sum_{i=1}^{n}\psi(\lambda_i(V))\right)\right)\\ &= \tfrac{1}{2}\psi'\left(\varrho\left(\Psi(V)\right)\right).\end{aligned}$$

This completes the proof of the theorem. □

Corollary 4.3 *Let ϱ be as defined in (18). Thus we have*

$$\delta(V) \geq \frac{\Psi(V)}{2\varrho\left(\Psi(V)\right)}.$$

Proof. Using Theorem 4.2, i.e., $\delta(V) \geq \frac{1}{2}\psi'(\varrho(\Psi(V)))$, we obtain from Lemma 3.3

$$\delta(V) \geq \frac{1}{2}\psi'(\varrho(\Psi(V))) \geq \frac{\psi\left(\varrho(\Psi(V))\right)}{2\varrho\left(\Psi(V)\right)} = \frac{\Psi(V)}{2\varrho\left(\Psi(V)\right)}.$$

This proves the corollary. □

Lemma 4.4 *If $\Psi(V) \geq 1$, then*

$$\delta(V) \geq \frac{1}{6}\Psi(V)^{\frac{1}{2}}. \tag{21}$$

Proof. The inverse function of $\psi(t)$ for $t \in [1, \infty)$ is obtained by solving t from the equation

$$\psi(t) = \frac{t^2 - 1}{2} + \frac{e^{\sigma(1-t)} - 1}{2} = s, \qquad t \geq 1.$$

Since it is hard to solve this equation explicitly, we derive an upper bound for t, as this suffices for our goal. One has from (16) and $\psi''(t) \geq 1$, we have

$$s = \psi(t) = \int_1^t \int_1^\xi \psi''(\zeta)\, d\zeta d\xi \geq \int_1^t \int_1^\xi d\zeta d\xi = \frac{1}{2}(t-1)^2,$$

which implies

$$t = \varrho\left(s\right) \leq 1 + \sqrt{2s}. \tag{22}$$

Assuming $s \geq 1$, we get

$$t = \varrho\left(s\right) \leq \sqrt{s} + \sqrt{2s} \leq 3s^{\frac{1}{2}}.$$

Omitting the argument V, and assuming $\Psi(V) \geq 1$, we thus have

$$\varrho(\Psi(V)) \leq 3\Psi(V)^{\frac{1}{2}}.$$

Now, using Corollary 4.3, we have

$$\delta(V) \geq \frac{\Psi(V)}{2\varrho\left(\Psi(V)\right)} \geq \frac{1}{6}\Psi(V)^{\frac{1}{2}}.$$

This proves the lemma. □

Note that if $\Psi(V) \geq 1$, substitution in (21) gives

$$\delta(V) \geq \frac{1}{6}. \tag{23}$$

4.1 Fixing the value of σ

After the update of μ to $(1-\theta)\mu$ we have $V_+ = \frac{V}{\sqrt{1-\theta}}$. Application of Theorem 4.1 yields that

$$\Psi(V_+) \leq L = L(n, \theta, \tau) = n\psi\left(\frac{\varrho\left(\frac{\tau}{n}\right)}{\sqrt{1-\theta}}\right). \tag{24}$$

Lemma 4.5 *Suppose that $L \geq 8$ and $\Psi(V) \leq L$. If $\sigma \geq 1 + 2\log(L+1)$, then $\lambda_i(V) > \frac{3}{2\sigma}$, for all $i = 1, ..., n$.*

Proof. First note that $\Psi(V) \leq L$ implies $\psi(\lambda_i(V)) \leq L$ for each $i = 1, ..., n$. Hence, putting $t = \lambda_i(V)$, we have

$$\frac{t^2-1}{2} + \frac{1}{\sigma}\left(e^{\sigma(1-t)} - 1\right) \leq L.$$

It follows that

$$\frac{1}{\sigma}\left(e^{\sigma(1-t)} - 1\right) \leq L + \frac{1-t^2}{2} \leq L + \frac{1}{2}, \tag{25}$$

This implies

$$e^{1-\sigma t} \leq \frac{1 + \sigma\left(L + \frac{1}{2}\right)}{e^{\sigma-1}}.$$

Since the expression at the right-hand side is monotonically decreasing in σ, it follows that

$$e^{1-\sigma t} \leq \frac{1 + (1 + 2\log(L+1))\left(L + \frac{1}{2}\right)}{(L+1)^2}.$$

The expression at the right-hand side is monotonically decreasing in L. The value at $L = 8$ is $0.57842... < e^{-\frac{1}{2}}$. Thus we obtain that $e^{1-\sigma t} < e^{-\frac{1}{2}}$, which implies $1 - \sigma t < -\frac{1}{2}$, or $t > \frac{3}{2\sigma}$, proving the lemma. □

Note that at the start of each inner iteration $\tau < \Psi(V) \leq L$. To ensure that L satisfies the conditions of Lemma 4.5, we assume from now that $L \geq 8$. and we choose

$$\sigma = 1 + 2\log(L+1). \tag{26}$$

5 Analysis of the algorithm

In the analysis of the algorithm the concepts of exponential convexity is crucial ingredient. Before dealing with these concepts, we start with two technical lemmas.

Lemma 5.1 *[Lemma 3.3.14 (c) in [8]] Let $A, B \in \mathbf{S}^n$ be two nonsingular matrices and $f(t)$ be given real-valued function such that $f(e^t)$ is a convex function. One has*

$$\sum_{i=1}^{n} f(\eta_i(AB)) \leq \sum_{i=1}^{n} f(\eta_i(A)\eta_i(B)), \tag{27}$$

where $\eta_i(A)$, and $\eta_i(B)$ $i = 1, 2, ..., n$ denote the singular values of A, and B, respectively

Lemma 5.2 *[Lemma 5.1 in [15]] Let A, $A + B \in \mathbf{S}^{\mathbf{n}}_{+}$, one has*

$$\lambda_i(A+B) \geq \lambda_1 - |\lambda_n(B)|, \quad i = 1, 2, ..., n. \tag{28}$$

Lemma 5.3 *Let V_1 and V_2 are two symmetric positive definite, and $\lambda_1(V_1), \lambda_1(V_2) \geq \frac{1}{\sigma}$, then*

$$\Psi\left((V_1^{\frac{1}{2}} V_2 V_1^{\frac{1}{2}})^{\frac{1}{2}}\right) \leq \frac{1}{2}\left(\Psi(V_1) + \Psi(V_2)\right). \tag{29}$$

Proof. By the definition of the singular values of a matrix we have any nonsingular matrix $U \in \mathbf{S}^n$,

$$\eta_i(U) = \left(\lambda_i(U^T U)\right)^{\frac{1}{2}} = \left(\lambda_i(UU^T)\right)^{\frac{1}{2}}, \quad i = 1, 2, ..., n.$$

From this, we can write

$$\eta_i(V_1^{\frac{1}{2}}V_2^{\frac{1}{2}}) = \left(\lambda_i(V_1^{\frac{1}{2}}V_2V_1^{\frac{1}{2}})\right)^{\frac{1}{2}} = \lambda_i\left((V_1^{\frac{1}{2}}V_2V_1^{\frac{1}{2}})^{\frac{1}{2}}\right), \quad i = 1, 2, ..., n.$$

Since V_1 and V_2 are symmetric positive definite, using Lemma 5.1 one has

$$\begin{aligned}\Psi\left((V_1^{\frac{1}{2}}V_2V_1^{\frac{1}{2}})^{\frac{1}{2}}\right) &= \sum_{i=1}^{n}\psi\left(\eta_i(V_1^{\frac{1}{2}}V_2^{\frac{1}{2}})\right) \leq \sum_{i=1}^{n}\psi\left(\eta_i(V_1^{\frac{1}{2}})\eta_i(V_2^{\frac{1}{2}})\right)\\ &\leq \frac{1}{2}\sum_{i=1}^{n}\left(\psi\left(\eta_i^2(V_1)^{\frac{1}{2}}\right)\psi\left(\eta_i^2(V_2)^{\frac{1}{2}}\right)\right)\\ &= \frac{1}{2}\sum_{i=1}^{n}\left(\psi\left(\eta_i(V_1)\right)\psi\left(\eta_i(V_2)\right)\right) = \frac{1}{2}\left(\Psi(V_1) + \Psi(V_2)\right).\end{aligned}$$

The second inequality follows from the exponential convexity of $\psi(t)$, $t \geq \frac{1}{\sigma}$. This completes the lemma.

5.1 Decrease of the proximity during a (damped) Newton step

In this section we start to compute the step size. After a damped step, with step size α. Following [15, 4], using (6) we have

$$\begin{aligned}&X_+ = X + \alpha\triangle X = X + \alpha\sqrt{\mu}DD_XD = \sqrt{\mu}D\left(V + \alpha D_X\right)D,\\ &y_+ = y + \alpha\Delta y,\\ &S_+ = S + \alpha\triangle S = X + \alpha\sqrt{\mu}D^{-1}D_SD^{1} = \sqrt{\mu}D^{-1}\left(V + \alpha D_S\right)D^{-1}.\end{aligned}$$

One has [11]

$$V_+ = \frac{1}{\sqrt{\mu}}\left(D^{-1}X_+S_+D\right)^{\frac{1}{2}}. \tag{30}$$

Note that V_+^2 is unitarily similar to the matrix $X_+^{\frac{1}{2}}S_+X_+^{\frac{1}{2}}$ and thus to

$$\left(V + \alpha D_X\right)^{\frac{1}{2}}\left(V + \alpha D_S\right)\left(V + \alpha D_X\right)^{\frac{1}{2}}.$$

This implies that the eigenvalues of V_+ are the same as those of the matrix

$$\tilde{V}_+ := \left(\left(V + \alpha D_X\right)^{\frac{1}{2}}\left(V + \alpha D_S\right)\left(V + \alpha D_X\right)^{\frac{1}{2}}\right)^{\frac{1}{2}}. \tag{31}$$

By the definition of $\Psi(V)$, we have $\Psi\left(\tilde{V}_+\right) = \Psi\left(V\right)$.

Our aim is to find an upper bound for $f(\alpha) := \Psi\left(V_+\right) - \Psi\left(V\right) = \Psi\left(\tilde{V}_+\right) - \Psi\left(V\right)$. To do this we assume for the moment that the step size α is such that:

$$\lambda_i(V + \alpha D_X) \geq \frac{1}{\sigma}, \quad \lambda_i(V + \alpha D_S) \geq \frac{1}{\sigma}, \quad i = 1, 2..., n. \tag{32}$$

Now ψ is e-convex, and using (29), so

$$\Psi\left(\tilde{V}_+\right) = \Psi\left(\left(\left(V + \alpha D_X\right)^{\frac{1}{2}}\left(V + \alpha D_S\right)\left(V + \alpha D_X\right)^{\frac{1}{2}}\right)^{\frac{1}{2}}\right) \leq \tfrac{1}{2}\left[\Psi\left(V + \alpha D_X\right) + \Psi\left(V + \alpha D_S\right)\right].$$

Thus we have $f(\alpha) \leq f_1(\alpha)$, where

$$f_1(\alpha) := \tfrac{1}{2}\left[\Psi\left(V + \alpha D_X\right) + \Psi\left(V + \alpha D_S\right)\right] - \Psi\left(V\right)$$

is convex in α, since Ψ is convex. Obviously, $f(0) = f_1(0) = 0$. Taking the derivative to α, we get

$$f_1'(\alpha) = \tfrac{1}{2} Tr \left(\psi' \left(V + \alpha D_X \right) D_X + \psi' \left(V + \alpha D_S \right) D_S \right).$$

This gives, using the last equality in (11) and (20),

$$f_1'(0) = \tfrac{1}{2} Tr \psi'(V) \left(D_X + D_S \right) = -\tfrac{1}{2} Tr \left(\psi'(V)^2 \right) = -2\delta(V)^2. \tag{33}$$

Differentiating once more, we obtain

$$f_1''(\alpha) = \tfrac{1}{2} Tr \left(\psi'' \left(V + \alpha D_X \right) D_X^2 + \psi'' \left(V + \alpha D_S \right) D_S^2 \right). \tag{34}$$

Below we use the following notation:

$$\lambda_1(V) := \min(\lambda_i(V)), \quad \delta := \delta(V).$$

Lemma 5.4 *One has*

$$f_1''(\alpha) \le 2\delta^2 \, \psi'' \left(\lambda_1(V) - 2\alpha\delta \right).$$

Proof. The last equality in (11) and (20) imply that $\|D_X + D_S\|^2 = \|D_X\|^2 + \|D_S\|^2 = 4\delta^2$. Thus we have $|\lambda_n(D_X)| \le 2\delta$ and $|\lambda_n(D_S)| \le 2\delta$. Using Lemma 5.2 and $V + \alpha D_X \succeq 0$, Therefore,

$$\lambda_i(V + \alpha D_X) \ge \lambda_1(V) - \alpha \left| \lambda_n(D_X) \right| \ge \lambda_1(V) - 2\alpha\delta, \quad 1 \le i \le n, \tag{35}$$

$$\lambda_i(V + \alpha D_S) \ge \lambda_1(V) - \alpha \left| \lambda_n(D_S) \right| \ge \lambda_1(V) - 2\alpha\delta, \quad 1 \le i \le n. \tag{36}$$

Due to (15), ψ'' is monotonically decreasing, so using the above inequalities, we have

$$\psi''(\lambda_i(V + \alpha D_X)) \le \psi''(\lambda_1(V) - 2\alpha\delta), \quad \psi''(\lambda_i(V + \alpha D_S)) \le \psi''(\lambda_1(V) - 2\alpha\delta), \tag{37}$$

this implies that

$$\psi''(V + \alpha D_X) \le \psi''(\lambda_1(V) - 2\alpha\delta)E, \quad \psi''(V + \alpha D_S) \le \psi''(\lambda_1(V) - 2\alpha\delta)E, \tag{38}$$

Now, using (12), and $\|D_X\|^2 + \|D_S\|^2 = 4\delta^2$, by (34) we obtain

$$f_1''(\alpha) \le \tfrac{1}{2} \, \psi'' \left(\lambda_1(V) - 2\alpha\delta \right) \sum_{i=1}^{n} \left(\lambda_i(D_X^2) + \lambda_i(D_X^2) \right) = 2\delta^2 \, \psi'' \left(\lambda_1(V) - 2\alpha\delta \right).$$

This proves the lemma. □

Putting $v_i = \lambda_i(X)$, $1 \le i \le n$, we have

$$f_1''(\alpha) \le 2\delta^2 \, \psi'' \left(v_1 - 2\alpha\delta \right),$$

which is the same inequality as inequality (41) in [3]. From this stage on we can apply exactly the same argument as in the LO case [3] to obtain the following results which require no further proof.

Lemma 5.5 *One has $f_1'(\alpha) \le 0$ if α satisfies the inequality*

$$-\psi' \left(\lambda_1(V) - 2\alpha\delta \right) + \psi' \left(\lambda_1(V) \right) \le 2\delta. \tag{39}$$

Lemma 5.6 *With ρ as defined in (19), as the inverse function of $-\frac{1}{2}\psi'(t)$ for $t \in (0, 1]$, the step size*

$$\bar{\alpha} := \frac{1}{2\delta} \left[\rho\left(\delta\right) - \rho\left(2\delta\right) \right] \tag{40}$$

is the largest possible solution of inequality (39).

Lemma 5.7 *Let ρ and $\bar{\alpha}$ be as defined in Lemma 5.6. Then*

$$\bar{\alpha} \geq \frac{1}{\psi''\left(\rho\left(2\delta\right)\right)}. \tag{41}$$

As in the LO case, we use

$$\widetilde{\alpha} = \frac{1}{\psi''\left(\rho(2\delta)\right)}, \tag{42}$$

as the default step size. By Lemma 5.7 we have $\widetilde{\alpha} \leq \bar{\alpha}$.

Lemma 5.8 *If the step size α is such that $\alpha \leq \bar{\alpha}$ then*

$$f(\alpha) \leq -\alpha\, \delta^2. \tag{43}$$

Theorem 5.9 *Let ρ be as defined in Lemma 5.6 and $\widetilde{\alpha}$ as in (42). Then*

$$f(\widetilde{\alpha}) \leq -\frac{\delta^2}{\psi''\left(\rho(2\delta)\right)} \leq -\frac{\delta}{16\sigma}. \tag{44}$$

Proof. Since $\widetilde{\alpha} \leq \bar{\alpha}$, Lemma 5.8 gives $f(\bar{\alpha}) \leq -\widetilde{\alpha}\, \delta^2$, where $\widetilde{\alpha} = \frac{1}{\psi''(\rho(2\delta))}$. Thus the first inequality follows. To obtain the inverse function $t = \rho(s)$ of $-\frac{1}{2}\psi'(t)$ for $t \in (\frac{1}{\sigma}, 1]$ we need to solve t from the equation

$$-\left(t - e^{\sigma(t-1)}\right) = 2s.$$

This implies, using $t \leq 1$,

$$e^{\sigma(1-t)} = 2s + t \leq 2s + 1.$$

Hence, putting $t = \rho(2\delta)$, which is equivalent to $4\delta = -\psi'(t)$, we get

$$e^{\sigma(1-t)} \leq 4\delta + 1. \tag{45}$$

Using (45), and $\sigma \geq 1$, thus we have

$$\widetilde{\alpha} = \frac{1}{\psi''(t)} = \frac{1}{1 + \sigma e^{\sigma(1-t)}} \geq \frac{1}{1 + \sigma\left(4\delta + 1\right)} \geq \frac{1}{\sigma\left(4\delta + 2\right)}.$$

Also using (23) (i.e., $6\delta \geq 1$) we get,

$$\widetilde{\alpha} \geq \frac{1}{\sigma\left(2 + 4\delta\right)} = \frac{1}{2\sigma\left(1 + 2\delta\right)} \geq \frac{1}{2\sigma\left(6\delta + 2\delta\right)} = \frac{1}{16\sigma\delta}.$$

Denote

$$\bar{\alpha} := \frac{1}{16\sigma\delta}; \tag{46}$$

this will be our default step size. Hence

$$f(\widetilde{\alpha}) \leq -\frac{\delta^2}{16\sigma\delta} = -\frac{\delta}{16\sigma}.$$

Thus the theorem follows. □

Substitution in (21) gives

$$f(\tilde{\alpha}) \leq -\frac{\delta}{16\sigma} \leq -\frac{\Psi^{\frac{1}{2}}}{96\sigma}.$$

Finally, to validate the above analysis we need to show that $\widetilde{\alpha} = \overline{\alpha}$ satisfies (32). Using (46) and Lemma 4.5, we may write

$$\lambda_i(V + \alpha D_X) \geq \lambda_1(V) - 2\overline{\alpha}\delta \geq \frac{3}{2\sigma} - \frac{2\delta}{16\delta\sigma} \geq \frac{3}{2\sigma} - \frac{1}{8\sigma} = \frac{11}{8\sigma} \geq \frac{1}{\sigma}, \quad i = 1, 2, ..., n, \tag{47}$$

and

$$\lambda_i(V + \alpha D_S) \geq \lambda_1(V) - 2\overline{\alpha}\delta \geq \frac{3}{2\sigma} - \frac{2\delta}{16\delta\sigma} \geq \frac{3}{2\sigma} - \frac{1}{8\sigma} = \frac{11}{8\sigma} \geq \frac{1}{\sigma}, \quad i = 1, 2, ..., n. \tag{48}$$

6 Complexity

In this section we drive the complexity bounds for large-update methods and small-update methods. An upper bound for the total number of iterations is obtained by multiplying (the upper bound for) the number K by the number of barrier parameter updates, which is bounded above by (cf. [12] Lemma II.17, page 116)

$$\frac{1}{\theta}\log\frac{n}{\epsilon}.$$

Lemma 6.1 *(Proposition 1.3.2 in [11]) Let $t_0, t_1, \cdots, t_K$ be a sequence of positive numbers such that*

$$t_{k+1} \le t_k - \kappa t_k^{1-\gamma}, \qquad k = 0, 1, \cdots, K-1, \tag{49}$$

where $\kappa > 0$ and $0 < \gamma \le 1$. Then $K \le \left\lfloor \frac{t_0^\gamma}{\kappa\gamma} \right\rfloor$.

Lemma 6.2 *if K denotes the number of inner iterations, we have*

$$K \le 192\Psi_0^{\frac{1}{2}}.$$

Proof. The definition of K implies $\Psi_{K-1} > \tau$ and $\Psi_K \le \tau$ and

$$\Psi_{k+1} \le \Psi_k - \kappa\left(\Psi_k\right)^{1-\gamma}, \qquad k = 0, 1, \cdots, K-1,$$

with $\kappa = \frac{1}{96}$ and $\gamma = \frac{1}{2}$. Application of Lemma 6.1, with $t_k = \Psi_k$ yields the desired inequality. □

Using $\psi_0 \le L$, and Lemma 6.2 we obtain the following upper bound on the total number of iterations:

$$\frac{192\sigma L^{\frac{1}{2}}}{\theta}\log\frac{n}{\epsilon}. \tag{50}$$

6.1 Large-update

We just established that (50) is an upper bound for the total number of iterations, where the number L is as given in (24):

$$L = n\psi\left(\frac{\rho\left(\frac{\tau}{n}\right)}{\sqrt{1-\theta}}\right). \tag{51}$$

Using

$$\psi(t) = \frac{t^2-1}{2} + \frac{e^{\sigma(1-t)}-1}{\sigma} \le \frac{t^2-1}{2}, \quad \text{for} \quad t \ge 1,$$

and (22), by substitution in (51) we obtain

$$L \le n\frac{\left(\frac{\varrho\left(\frac{\tau}{n}\right)}{\sqrt{1-\theta}}\right)^2 - 1}{2} \le \frac{n}{2(1-\theta)}\left(\theta + 2\sqrt{2\frac{\tau}{n}} + \frac{2\tau}{n}\right) = \frac{\left(\theta n + 2\sqrt{2\tau n} + 2\tau\right)}{2(1-\theta)}.$$

Using (50), thus the total number of iterations is bounded above by

$$\frac{K}{\theta}\log\frac{n}{\epsilon} \le \frac{192\sigma}{\theta\sqrt{2(1-\theta)}}\left(\theta n + 2\sqrt{2\tau n} + 2\tau\right)^{\frac{1}{2}}\log\frac{n}{\epsilon}.$$

A large-update methods uses $\tau = O(n)$ and $\theta = \Theta(1)$ and $\sigma = O(\log n)$. The right-hand side expression is then $O\left(\sqrt{n}\log n \log\frac{n}{\epsilon}\right)$, as easily may be verified.

6.2 Small-update methods

For small-update methods one has $\tau = O(1)$ and $\theta = \Theta\left(\frac{1}{\sqrt{n}}\right)$. Using Lemma 3.2, with $\psi''(1) = 1 + \sigma$, we then obtain

$$L = n\psi\left(\frac{\varrho\left(\frac{\tau}{n}\right)}{\sqrt{1-\theta}}\right) \leq \frac{n(1+\sigma)}{2}\left(\frac{\varrho\left(\frac{\tau}{n}\right)}{\sqrt{1-\theta}} - 1\right)^2.$$

Using (22), then

$$L \leq \frac{n(1+\sigma)}{2}\left(\frac{1+\sqrt{\frac{2\tau}{n}}}{\sqrt{1-\theta}} - 1\right)^2.$$

Using $1 - \sqrt{1-\theta} = \frac{\theta}{1+\sqrt{1-\theta}} \leq \theta$, this leads to

$$L \leq \frac{(1+\sigma)}{2(1-\theta)}\left(\theta\sqrt{n} + \sqrt{2\tau}\right)^2 = \sigma O(1) = O(\sigma). \tag{52}$$

Using (26) (i.e., $\sigma = 1 + 2\log(1+L)$), by (52), we have

$$\sigma \leq 1 + 2\log(1 + O(\sigma)). \tag{53}$$

This implies that $\sigma = O(1)$. Then $L = O(1)$. Using (50), thus the total number of iterations is bounded above by

$$\frac{K}{\theta}\log\frac{n}{\epsilon} \leq \frac{192}{\theta}O(1)\log\frac{n}{\epsilon} = O\left(\sqrt{n}\log\frac{n}{\epsilon}\right).$$

7 Concluding Remarks

If a better iteration complexity can be proved for linear optimization, the tools developed in this paper will make it possible to extend these results to somidefinite optimization problem. We showed that the iteration bound of a large-update interior-point method for SDO based on the finite kernel functions considered in this paper is $O\left(\sqrt{n}\log n \log\frac{n}{\epsilon}\right)$, which improves the classical iteration complexity with a factor almost $\sqrt{n}$. For small-update method we obtain the iteration bound, $O\left(\sqrt{n}\log\frac{n}{\epsilon}\right)$. In both cases these are the best known iteration bounds.

References

[1] F. Alizadeh. *Combinatorial Optimization with Interior Point Methods and Semi-Definite Matrices*. PhD thesis, University of Minnesota, Minneapolis, Minnesota, USA, 1991. 1

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

[3] Y.Q. Bai, M. El Ghami, and C. Roos. A new efficient large-update primal-dual interior-point method based on a finite barrier. *SIAM Journal on Optimization*, 13(3):766–782, 2003. (document), 1, 3, 3, 3, 5.1

[4] Y.Q. Bai, M. El Ghami, and C. Roos. A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization. *SIAM Journal on Optimization*, 15(1):101–128, 2004. 3, 3, 3.4, 3.5, 4, 5.1

[5] E. de Klerk. *Interior Point Methods for Semidefinite Programming*. PhD thesis, TU Delft, The Netherlands, 1997. 1.1

[6] E. de Klerk. *Aspects of Semidefintie Programming*, volume 65 of *Applied Optimization*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002. 1.2

[7] M. El Ghami. *New Primal-dual Interior-point Methods Based on Kernel Functions.* PhD thesis, TU Delft, The Netherlands, 2005. 3

[8] R.A. Horn, and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, UK, 1985. 2, 5.1

[9] Y.E. Nesterov, and A.S. Nemirovskii. *Interior Point Polynomial Methods in Convex Programming : Theory and Algorithms*. SIAM, Philadelphia, USA, 1993. 1

[10] Y.E. Nesterov, and M.J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*, 22(1):1–42, 1997. 1.2

[11] J. Peng, C. Roos, and T. Terlaky. *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms*. Princeton University Press, 2002. 1, 1.1, 2.1, 2, 3, 5.1, 6.1

[12] C. Roos, T. Terlaky, and J.-ph. Vial. *Theory and Algorithms for Linear Optimization. An Interior-Point Approach*. Springer Science, 2005. 6

[13] W. Rudin. *Principles of Mathematical Analysis*. Mac-Graw Hill Book Company, New York, 1978. 2

[14] J. F. Sturm, and S. Zhang. Symmetric primal-dual path following algorithms for semidefinite programming. *Applied Numirical Mathematics*, 29:301–315, 1999. 1.2

[15] G. Q. Wang, Y. Q. Bai, and C. Roos. Primal-dual interior-point algorithms for semidefinite optimization based on a simple kernel function. *Journal of Mathematical Modelling and Algorithms, 4:409–433, 2005.* 5.2, 5.1