

A filter algorithm for nonlinear semidefinite programming

WALTER GÓMEZ^{1*} and HÉCTOR RAMÍREZ^{2†}

¹Departamento de Ingeniería Matemática, Universidad de La Frontera
Av. Francisco Salazar 01145, Casilla 54-D, Temuco, Chile

²Departamento de Ingeniería Matemática, Centro de Modelamiento Matemático
(CNRS UMI 2807), FCFM, Universidad de Chile, Avda. Blanco Escalada 2120, Santiago, Chile

E-mails: wgozmez@ufro.cl / hramirez@dim.uchile.cl

Abstract. This paper proposes a filter method for solving nonlinear semidefinite programming problems. Our method extends to this setting the filter SQP (sequential quadratic programming) algorithm, recently introduced for solving nonlinear programming problems, obtaining the respective global convergence results.

Mathematical subject classification: 90C30, 90C55.

Key words: nonlinear semidefinite programming, filter methods.

1 Introduction

We consider the following nonlinear semidefinite programming problem (NLSDP)

$$\min_{x \in \mathbb{R}^n} \{f(x); h(x) = 0, G(x) \preceq 0\}, \quad (\text{P})$$

where $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $h: \mathbb{R}^n \rightarrow \mathbb{R}^p$ and $G: \mathbb{R}^n \rightarrow S^m$ are C^2 -functions, S^m denotes the linear space of $m \times m$ real symmetric matrices equipped with the inner product $A \cdot B := \text{trace}(AB) = \sum_{i,j=1}^m A_{ij}B_{ij}$ for all matrices $A = (A_{ij})$, $B = (B_{ij}) \in S^m$, and \preceq denotes the negative semidefinite order, that is, $A \preceq B$ iff

#CAM-186/10. Received: 08/II/10. Accepted: 10/III/10.

*Supported by CONICYT (Chile) under FONDECYT grant 1050696.

†Supported by CONICYT (Chile) under FONDECYT grant 1070297 and BASAL project.

$A - B$ is a negative semidefinite matrix. The order relations $<$, \succeq and $>$ are defined similarly.

A particular important case of the above model is the *Linear SemiDefinite Programming* (LSDP) problem, which appears when constraint $h(x) = 0$ is omitted, and $f(x)$ and $G(x)$ are linear mappings (e.g. [47, 50, 53]). The LSDP model is one of the most active research field in the last 15 years with a huge amount of applications in control theory, optimal design, among others, and links to relaxation methods for combinatorial problems. See for instance [5, 8, 43].

The mentioned research effort has produced a lot of different approaches for solving the LSDP problem. Let us just mention the fundamental approach of interior point polynomial methods [1, 5, 39], penalty/barrier and augmented Lagrangian methods [26, 37], spectral bundle methods [21], cutting plane methods [2, 20, 28], among others.

The LSDP model is however insufficient for solving important application problems, see e.g. [38, 42]. The more general nonlinear problem NLSDP has been then also used for modelling in, for instance, feedback control, structural optimization and truss design problems. See [4, 22], and more recently [3, 24, 27] and references therein.

The first- and second-order optimality conditions for NLSDP have been widely characterized, see for instance [7] and the references therein. Concerning numerically efficient solution algorithms the situation for NLSDP is completely different to the one of the LSDP case. In fact, few existing approaches have been recently proposed and are mainly still in a development stage.

Leibfritz and Mostafa developed in [34, 35] an interior point constrained trust region algorithm for a special class of NLSDP problems.

Another approach was developed by Jarre [23]. It generalizes the predictor-corrector interior-point method to nonconvex semidefinite programming.

An augmented Lagrangian algorithm and a code (PENNON) supporting NLSDP problems without equality constraints were developed by Kočvara and Stingl [25, 26]. The code was successfully tested for LSDP but recent experiments with nonconvex NLSDP [27] reported still unsatisfactory performance.

In a serie of papers, Apkarian, Noll and others have suggested different approaches for solving NLSDP problems. For instance, the partially augmented

Lagrangian approach [3, 11, 42], the spectral bundle methods [40, 41], and the sequential semidefinite programming algorithm [12]. In this paper we are mainly interested in the last idea.

The sequential semidefinite programming (SSDP) method [12] is an extension of the classical *Sequential Quadratic Programming* (SQP) algorithm to the cone S_+^p of symmetric positive semidefinite matrices. Here a sequence of Karush-Kuhn-Tucker points (see the definition in Section 2) is generated by solving at each step a tangent problem to NLSDP. The subproblems can be stated as LSDP and efficiently solved. In [12] it was shown the local fast convergence of the SSDP method.

The SSDP method was later deduced from a different viewpoint by Freund and Jarre in [16]. They also provided in [17] another proof of local quadratic convergence. Another recent paper dealing with local convergence properties of the SSDP algorithm is [10], where the difficulties associated with carrying the quadratic local convergence from the SQP approach are discussed.

Correa and Ramírez proved in [9] the global convergence of SSDP using a merit function (called Han penalty function) and a line search strategy.

Motivated by the local convergence properties obtained for the SSDP algorithm and the global ones obtained for the filter SQP, the aim of this paper is to exploit this recent filter methodology in order to propose a globally convergent algorithm for solving NLSDP.

In [17] the authors report a good experience with a filter strategy for SSDP, but a proof of the global convergence is not provided.

The filter methods were first introduced by Fletcher and Leyffer in [14]. In this technique the trial points are accepted when they improve either the objective function or a constraint violation measure. These criteria are compared to previous iterates collected in a filter. This idea offers an efficient alternative to traditional merit functions with penalty terms which adjustment can be problematic.

The filter approach was first used as a tool for proving global convergence of algorithms for NLP, but soon it was successfully exploited in many ways for quite different algorithmic ideas in NLP. For instance, global convergence results of trust region filter SQP methods were established in [13, 15]. Local conver-

gence of a particular filter trust region SQP, avoiding the Maratos effect, was given in [49]. Global and local convergence results for filter SQP algorithms using line search were also presented in [51, 52]. Furthermore, filter methods have been successfully combined with many approaches like, for instance, interior point [6, 48], merit functions [19], etc.

The paper is organized as follows. In the next section some notations and preliminary results are introduced. In the third section the filter SDP algorithm is given in details and some first auxiliary results are stated. The fourth section is devoted to the proof of the main global convergence results. Finally, the fifth section contains numerical tests on selected problems from the COMPElib library and the sixth section establishes some short concluding remarks.

2 Preliminaries

We say that $(\bar{\lambda}, \bar{Y})$ is a Lagrange multiplier associated with \bar{x} , if $(\bar{x}, \bar{\lambda}, \bar{Y}) \in \mathbb{R}^n \times \mathbb{R}^p \times S^m$ satisfies the following Karush-Kuhn-Tucker (KKT) conditions:

$$\nabla_x L(\bar{x}, \bar{\lambda}, \bar{Y}) = \nabla f(\bar{x}) + Dh(\bar{x})^\top \bar{\lambda} + DG(\bar{x})^\top \bar{Y} = 0, \quad (1a)$$

$$G(\bar{x})\bar{Y} = 0, \quad (1b)$$

$$G(\bar{x}) \preceq 0, h(\bar{x}) = 0, \bar{Y} \succeq 0, \quad (1c)$$

where $L : \mathbb{R}^n \times \mathbb{R}^p \times S^m \rightarrow \mathbb{R}$ is the Lagrangian function of problem (P)

$$L(x, \lambda, Y) := f(x) + h(x)^\top \lambda + Y \cdot G(x). \quad (2)$$

Here and throughout the paper $Dh(x) := \left[\frac{\partial h_i(x)}{\partial x_j} \right]_{ij}$ denotes the $p \times n$ Jacobian matrix of h and $\nabla f(x) \in \mathbb{R}^n$ the column gradient vector of f at x . Note that, for a linear operator $Ay := \sum_{i=1}^m y_i A_i$ with $A_i \in S^m$, as $DG(x)$, we have for its adjoint operator A^\top the formula:

$$A^\top Z = (A_1 \cdot Z, \dots, A_n \cdot Z)^\top, \quad \forall Z \in S^m. \quad (3)$$

A pair $(\bar{x}, \bar{\lambda}, \bar{Y})$ satisfying (1) will be also called critical point or KKT-point of problem (P), and the set of Lagrange multipliers associated with \bar{x} will be denoted by $\Lambda(\bar{x})$. Finally, \bar{x} is called a critical point or critical solution of (P) if $\Lambda(\bar{x}) \neq \emptyset$.

Also, let $q = m - r$ with $r := \text{rank}(G(\bar{x}))$, and denote by $E = E(\bar{x}) \in \mathbb{R}^{m \times q}$ a matrix whose columns are an orthonormal basis of $\text{Ker } G(\bar{x})$. Using this notation, KKT conditions (1) can be equivalently written in terms of $\bar{\Phi} := E^\top \bar{Y} E \in S^q$ as follows

$$\nabla f(\bar{x}) + Dh(\bar{x})^\top \bar{\lambda} + DG(\bar{x})^\top (E \bar{\Phi} E^\top) = 0, \tag{4a}$$

$$G(\bar{x}) \preceq 0, h(\bar{x}) = 0, \bar{\Phi} \succeq 0. \tag{4b}$$

Hence $\bar{\Phi}$ can be seen as a *reduced* Lagrangian multiplier associated with the SDP constraint of problem (P) that obviously satisfies that $\bar{Y} = E \bar{\Phi} E^\top$.

In this paper, we will use *Robinson's constraint qualification condition* [44] defined at a feasible point \bar{x} of (P) as

$$0 \in \text{int} \left\{ \begin{pmatrix} G(\bar{x}) \\ h(\bar{x}) \end{pmatrix} + \begin{pmatrix} DG(\bar{x}) \\ Dh(\bar{x}) \end{pmatrix} \mathbb{R}^n - \begin{pmatrix} S_-^m \\ \{0\} \end{pmatrix} \right\}, \tag{5}$$

where $\text{int } C$ denotes the topological interior of the set C , and $S_-^m = \{A \in S^m \mid A \preceq 0\}$. It is easy to see that condition (5) is equivalent to *Mangasarian-Fromovitz constraint qualification condition (MFCQ)*

$$\{\nabla h_j(\bar{x})\}_{j=1}^p \text{ is linearly independent, and} \tag{6a}$$

$$\exists \bar{d} \in \mathbb{R}^n \text{ s. t. } \begin{cases} Dh(\bar{x})\bar{d} = 0 \\ \text{and } G(\bar{x}) + DG(\bar{x})\bar{d} \prec 0. \end{cases} \tag{6b}$$

Also, it can be shown that under (5) the set of Lagrange multipliers $\Lambda(\bar{x})$ is nonempty and also bounded [29]. Actually, when \bar{x} is assumed to be a local solution of (P), condition (5) is equivalent to saying that $\Lambda(\bar{x})$ is nonempty and compact.

Given a matrix $A = (A_{ij}) \in S^m$ we denote by $\|A\|_{Fr}$ its norm associated with the mentioned (trace) inner product in S^m . This is also called the Frobenius norm of A and is given by

$$\|A\|_{Fr} = \sqrt{\text{trace}(A^2)} = \sqrt{\sum_{i,j=1}^m A_{ij}^2} = \sqrt{\sum_{i=1}^m \lambda_i(A)^2},$$

where $(\lambda_1(A), \dots, \lambda_m(A))$ denotes the vector of eigenvalues of A in non-increasing order. In particular it holds that

$$|\lambda_j(A)| \leq \|A\|_{Fr}, \quad j = 1, \dots, m. \tag{7}$$

With the linear operator $DG(x)$ we can also associate a Frobenius norm defined as follows

$$\|DG(x)\|_{Fr} = \sqrt{\sum_{i=1}^n \left\| \frac{\partial G(x)}{\partial x_i} \right\|_{Fr}^2},$$

which satisfies the relation

$$\|DG(x)d\|_{Fr} \leq \|d\|_2 \|DG(x)\|_{Fr}. \tag{8}$$

Here and from now on $\|v\|_2$ denotes the Euclidian norm of a given vector v .

The operator $D^2G(x) : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow S^m$, obtained from the second derivative of $G(x)$, is defined by

$$D^2G(x)(d, \hat{d}) = d^\top D^2G(x)\hat{d} = \sum_{i,j=1}^n d_i \hat{d}_j \frac{\partial^2 G(x)}{\partial x_i \partial x_j}.$$

The corresponding Frobenius norm can be analogously defined as

$$\|D^2G(x)\|_{Fr} = \sqrt{\sum_{i,j=1}^n \left\| \frac{\partial^2 G(x)}{\partial x_i \partial x_j} \right\|_{Fr}^2}.$$

Hence, the following relationship holds trivially for all $j = 1, \dots, m$

$$|\lambda_j(d^\top D^2G(x)d)| \leq \|d^\top D^2G(x)d\|_{Fr} \leq \|d\|_2^2 \|D^2G(x)\|_{Fr}. \tag{9}$$

Indeed, it follows from the equality

$$\|d^\top D^2G(x)d\|_{Fr} = \sqrt{\sum_{k,l=1}^m (d^\top D^2G_{kl}(x)d)^2},$$

where $G_{kl}(x)$ is the (k, l) entry of $G(x)$.

Lemma 1 (Debreu’s Lemma). *Let $A \leq 0$ be negative semidefinite. There exists $r > 0$ such that $B + rA < 0$ if and only if $B|_{\text{Ker } A} < 0$.*

In this lemma the notation $B|_{\text{Ker } A} < 0$ means that the matrix B is negative definite on the linear space $\text{Ker } A$. This is equivalent to saying that the matrix $V^\top B V$ is negative definite, for any matrix V such that their columns build a linear basis of $\text{Ker } A$. Moreover, due to the Sylvester’s law of inertia (e.g. [18, Thm. 4.5.8]) the fact that $V^\top B V$ be negative definite does not depend on the selected basis V .

Remark 2. A simple consequence of Debreu’s Lemma is that: if $A \preceq 0$ and $B|_{\text{Ker } A} \prec 0$ then there exists a $\bar{\tau} > 0$ such that $A + \tau B \prec 0$ for all $\tau \in (0, \bar{\tau}]$.

The following useful lemma relates the fulfillment of the KKT optimality conditions and the MFCQ constraint qualifications.

Lemma 3. *Let us suppose that the point \bar{x} is feasible for (P) satisfying the MFCQ condition (6). If \bar{x} is not a critical solution of (P), then there exist a unitary vector $\bar{s} \in \mathbb{R}^p$ and some $\bar{\eta} > 0$ such that for all $\eta \in (0, \bar{\eta}]$ we have*

$$\nabla f(\bar{x})^\top \bar{s} < 0, \tag{10a}$$

$$Dh(\bar{x})\bar{s} = 0, \tag{10b}$$

$$G(\bar{x}) + \eta DG(\bar{x})\bar{s} \prec 0. \tag{10c}$$

Proof. Note that the vectors $\{\nabla h_j(\bar{x})\}_{j=1}^p$ are linearly independent due to the fact that MFCQ is fulfilled at \bar{x} .

Since KKT conditions are not satisfied at \bar{x} , we can separate the point $0 \in \mathbb{R}^n$ from the convex and closed set $\mathcal{L}(\mathbb{R}^p \times S_+^q)$, where

$$\mathcal{L}(\lambda, \Phi) := \nabla f(\bar{x}) + Dh(\bar{x})^\top \lambda + DG(\bar{x})^\top (E\Phi E^\top)$$

is a linear function in $(\lambda, \Phi) \in \mathbb{R}^p \times S_+^q$, q is the dimension of $\text{Ker } G(\bar{x})$, and $E \in \mathbb{R}^{m \times q}$ is a matrix whose columns are an orthonormal basis of $\text{Ker } G(\bar{x})$ (see (4)).

We obtain therefore the existence of a unitary vector $\bar{s} \in \mathbb{R}^n$ such that

$$\begin{aligned} 0 &> \nabla f(\bar{x})^\top \bar{s} + \lambda^\top Dh(\bar{x})\bar{s} + DG(\bar{x})\bar{s} \cdot (E\Phi E^\top) \\ &= \nabla f(\bar{x})^\top \bar{s} + \lambda^\top Dh(\bar{x})\bar{s} + (E^\top DG(\bar{x})\bar{s}E) \cdot \Phi, \quad \forall (\lambda, \Phi) \in \mathbb{R}^p \times S_+^q. \end{aligned}$$

Since the latter inequality is valid for all $(\lambda, \Phi) \in \mathbb{R}^p \times S_+^q$ and \bar{x} satisfies MFCQ, it directly follows that conditions (10a) and (10b) hold, as well as $E^\top DG(\bar{x})\bar{s}E \prec 0$ (by redefining, if necessary, \bar{s} as $\bar{s} + \delta \bar{d}$ with $\delta > 0$ small enough and \bar{d} given by (6b)). This last inequality together with Debreu’s lemma (see also remark 2) finally implies (10c) for all positive η smaller than some $\bar{\eta} > 0$. □

3 The Filter-SDP Algorithm

Let us first adapt the filter SQP algorithm of Fletcher et al. [15] to problem (P).

For a given point $x \in \mathbb{R}^n$ and a positive radius $\rho > 0$ we define the following trust region local semidefinite approximation of the problem (P):

$$\begin{aligned}
 QP(x, \rho) \quad & \min_{d \in \mathbb{R}^n} \quad \nabla f(x)^\top d + \frac{1}{2} d^\top B d \\
 & s.t. \quad h(x) + Dh(x)d = 0 \\
 & \quad \quad G(x) + DG(x)d \leq 0 \\
 & \quad \quad \|d\|_\infty \leq \rho.
 \end{aligned} \tag{11}$$

Here $B \in S^n$ is a matrix containing second order information of the problem (P) at x . Since its definition is not crucial for the analysis of the global convergence, it is not specified in detail.

Remark 4. In classical nonlinear programming the use of the Euclidian norm $\|\cdot\|_2$ would add extra difficulties in the resolution of the subproblems $QP(x, \rho)$. However, since the constraint $\|d\|_2 \leq \rho$ can be expressed as a *second-order cone constraint* (see [5] for a definition), which at the same time can be seen as a SDP constraint, its use does not increase the difficulty of the problem. Hence, this modification could be tested to obtain numerical improvements. From a theoretical point of view, the use of $\|\cdot\|_2$ only modifies slightly the proofs given onwards.

The filter algorithm deals simultaneously with the optimality and feasibility aspects of (P) using a dominance strategy typical of multiobjective problems. In order to quantify the feasibility let us define

$$\theta(x) := \|h(x)\|_2 + \lambda_1(G(x))_+. \tag{12}$$

where $\lambda_1(A)$ is the largest eigenvalue of the matrix A , and $(\alpha)_+ := \max\{0, \alpha\}$ denotes the positive part of the real number α . This function is obviously positive and vanishes exactly on the feasible points of (P).

A filter, denoted by \mathcal{F} , is a finite collection of two dimensional vectors. In the vectors of the filter the first and second component refer to the value of the feasibility function, θ , and of the objective function, f , respectively.

The new candidate $(\bar{\theta}, \bar{f})$ is called *acceptable* to the filter $\mathcal{F} = \{(\theta_j, f_j)\}_{j=1}^N$, if for each $j = 1 \dots N$ at least one of the following two conditions is fulfilled:

$$\bar{\theta} \leq \beta \theta_j, \quad (13a)$$

$$\bar{f} + \gamma \bar{\theta} \leq f_j, \quad (13b)$$

where $\beta \in (0, 1)$ and $\gamma \in (0, \beta)$ are two fixed parameters.

Let us suppose that the point $(\bar{\theta}, \bar{f})$ is acceptable for \mathcal{F} . The new filter $Add((\bar{\theta}, \bar{f}), \mathcal{F})$ is defined as

$$Add((\bar{\theta}, \bar{f}), \mathcal{F}) = (\mathcal{F} \cup \{(\bar{\theta}, \bar{f})\}) \setminus \{(\theta_j, f_j) \in \mathcal{F} \mid \bar{\theta} \leq \theta_j, \bar{f} \leq f_j\}$$

Note that a two dimensional vector is acceptable to $Add((\bar{\theta}, \bar{f}), \mathcal{F})$ if and only if it is acceptable for $\mathcal{F} \cup \{(\bar{\theta}, \bar{f})\}$. Also, acceptability considered in this way uses smaller filters because all the *dominated* pairs are removed.

Let us now fix the parameters of our filter version of the SSDP algorithm as follows: $\beta \in (0, 1)$, $\gamma \in (0, \beta)$, $\rho^{\max} > \bar{\rho} > 0$, $\sigma \in (0, 1)$ and $u > 0$.

Filter-SDP Algorithm

Step 0 Define $k = 1$, $\mathcal{F}^0 = \{(u, -\infty)\}$.

Step 1 Find some vector x_k and a corresponding trust region radius $\rho^{\max} \geq \rho \geq \bar{\rho}$ such that

$$(A1) \ (\theta(x_k), f(x_k)) \text{ is acceptable to } \mathcal{F}^{k-1}$$

$$(B1) \ QP(x_k, \rho) \text{ is feasible.}$$

Go to Step 3.

Step 2 If $QP(x_k, \rho)$ is not feasible then set $\mathcal{F}^k = Add((\theta(x_k), f(x_k)), \mathcal{F}^{k-1})$, $k = k + 1$ and go to Step 1.

Step 3 If $d = 0$ is a critical solution of $QP(x_k, \rho)$ then STOP. Otherwise, fix an optimal solution $d \neq 0$ of $QP(x_k, \rho)$.

Step 4 If $(\theta(x_k + d), f(x_k + d))$ is not acceptable to the filter $\mathcal{F}^{k-1} \cup \{(\theta(x_k), f(x_k))\}$ then set $\rho \leftarrow \rho/2$ and go to Step 2.

Step 5 If the following conditions are fulfilled

$$\nabla f(x_k)^\top d + \frac{1}{2}d^\top B d < 0, \quad (14a)$$

$$f(x_k) + \sigma \left(\nabla f(x_k)^\top d + \frac{1}{2}d^\top B d \right) < f(x_k + d), \quad (14b)$$

then set $\rho \leftarrow \rho/2$ and go to Step 2.

Step 6 If $\nabla f(x_k)^\top d + \frac{1}{2}d^\top B d \geq 0$ then $\mathcal{F}^k = \text{Add}((\theta(x_k), f(x_k)), \mathcal{F}^{k-1})$, otherwise set $\mathcal{F}^k = \mathcal{F}^{k-1}$.

Step 7 Define $\rho_k = \rho$, $d_k = d$. Set $x_{k+1} = x_k + d_k$, $k = k + 1$. Finally, initialize $\rho^{\max} \geq \rho \geq \bar{\rho}$ and go to Step 2.

In what follows, Step 1 is called *restoration phase* and the loop between the Steps 2 and 5 is defined as the *inner loop*.

Remark 5. The restoration phase in Step 1 should be seen as an isolated routine mainly dealing with the problem of approaching feasible points (for both problems: (P) and (QP)). From a theoretical point of view, no particular algorithm is explicitly considered in this phase. However, in practice, a reasonable and simple routine consists of few minimizing steps for a function measuring the feasibility of the original problem (see, for instance, [19, 45] in the NLP case). This is also the approach we follow in the numerical experiences of Section 5 where the function θ is used to measure the feasibility of the problem (P). Nevertheless, for this type of routines, the eventual convergence towards a local minimizer of the feasibility function, which is infeasible for the original problem (and consequently, perhaps not acceptable for the filter) is an important obstacle for the global convergence of the whole algorithm. This difficulty is also found in the NLP framework, as the reader can see in the references mentioned above.

In the above algorithm the iterates x_k are generated at Steps 2 and 7. Once Step 7 is achieved and a new iterate x_{k+1} is computed, the filter remains the same only if the following condition holds:

$$f(x_k) - f(x_{k+1}) \geq -\sigma \left(\nabla f(x_k)^\top d + \frac{1}{2}d^\top B d \right) > 0. \quad (15)$$

In this case x_k is called an f -type iteration, because the above condition guarantees an adequate improvement of the objective function f . Otherwise, when $(\theta(x_k), f(x_k))$ is entered to the filter, either at Step 2 or at Step 6, we say that a θ -type iteration x_k was accomplished.

Lemma 6. *Let us suppose that the MFCQ condition is satisfied at every feasible point of (P). If \mathcal{F}^k is a filter generated by the above algorithm, then*

$$\tau^k = \min_{(\theta, f) \in \mathcal{F}^k} \theta > 0.$$

Proof. It suffices to show that every point added to the Filter must be infeasible. By contradiction, suppose that $(\theta(x_k), f(x_k))$ is added to the filter and that x_k is a feasible point of (P). There are only two steps where a new member can be added to the filter: Steps 2 or 6.

Firstly, the feasibility of x_k implies that $d = 0$ is a feasible point of $QP(x_k, \rho)$ for all $\rho \geq 0$. Consequently $(\theta(x_k), f(x_k))$ could not be added to the filter at step 2.

Secondly, at Step 6 the problem $QP(x_k, \rho)$ is supposed to be feasible but with a nonnegative optimal value. According to Step 3, vector $d = 0$ is a feasible but not a critical point of $QP(x_k, \rho)$. Now since at $d = 0$ the norm inequality is not active, the KKT conditions for (P) at x_k and for $QP(x_k, \rho)$ at $d = 0$ coincide. We thus deduce that x_k is not a critical point of (P). Then Lemma 3 implies the existence of a nonzero vector $s \in \mathbb{R}^n$ such that for all $t \in (0, 1)$

$$\begin{aligned} \nabla f(x_k)^\top s &< 0, \\ h(x_k) + Dh(x_k)ts &= 0, \\ G(x_k) + DG(x_k)ts &\leq 0. \end{aligned}$$

Therefore s is a local feasible descent direction of $QP(x_k, \rho)$ at $d = 0$. This contradicts the fact that the optimal value of this auxiliary problem is nonnegative. \square

Remark 7. The above lemma ensures that Step 1 is always realizable when the original problem (P) is feasible.

The following two lemmas concerning sequences generated by the Filter can be found in [15, pp. 47–48]. Both results are independent to the SDP constraint appearing in (P). Nevertheless, we present for the sake of completeness a slightly different proof of the second result.

Lemma 8. *Consider a sequence $\{(\theta_k, f_k)\}_{k=1}^\infty$ such that $\theta_k \geq 0$, $\{f_k\}_{k=1}^\infty$ is decreasing and bounded below, and (θ_{k+1}, f_{k+1}) is acceptable to (θ_k, f_k) for all k . Then θ_k converges to zero.*

Lemma 9. *Consider a sequence of iterations $\{(\theta_k, f_k)\}_{k=1}^\infty$ entered into the filter in the above algorithm and such that $\{f_k\}_{k=1}^\infty$ is bounded below. It follows that θ_k converges to zero.*

Proof. Since the initial filter is given by $\mathcal{F}^0 = \{(u, -\infty)\}$, Lemma 6 implies that $\{\theta_k\}_{k=1}^\infty \subset (0, u]$. If we now suppose that $\{\theta_k\}_{k=1}^\infty$ does not converge to zero, then there must exist a minimal index $\bar{l} \geq 0$ such that

$$\text{card}\left(\{k \in \mathbb{N} \mid \beta^{\bar{l}+1}u < \theta_k \leq \beta^{\bar{l}}u\}\right) = \infty, \quad (16a)$$

$$\text{card}\left(\{k \in \mathbb{N} \mid \beta^{l+1}u < \theta_k \leq \beta^l u\}\right) < \infty, \quad \forall 0 \leq l < \bar{l}, \quad (16b)$$

where $\text{card}(S)$ denotes the cardinality of a given set S . Note that $(\beta^{\bar{l}+1}u, \beta^{\bar{l}}u]$ contains a subsequence of $\{\theta_k\}_{k=1}^\infty$ converging to $\limsup_k \theta_k$. As a consequence we can find an index $\bar{k} > 0$ such that

$$\theta_k \leq \beta^{\bar{l}}u, \quad \forall k \geq \bar{k}, \quad (17a)$$

$$\text{card}\left(\{k \in \mathbb{N} \mid k \geq \bar{k}, \beta^{\bar{l}+1}u < \theta_k \leq \beta^{\bar{l}}u\}\right) = \infty. \quad (17b)$$

We proceed to define recursively the following subsequence $\{\theta_{k(i)}\}_{i \in \mathbb{N}}$. First $\theta_{k(1)} = \theta_{\bar{k}}$. Then, given $\theta_{k(i)}$ we define $\theta_{k(i+1)}$ by taking

$$k(i+1) = \min\{k > k(i) \mid \theta_k > \beta\theta_{k(i)}\}.$$

Note that $k(i+1) > k(i)$ is well defined due to (17). It follows also from the above definition that

$$\theta_k \leq \beta\theta_{k(i)}, \quad \forall k(i) < k < k(i+1). \quad (18)$$

Since all iterates are entered to the filter, it follows

$$f(x_{k(i)}) - f(x_{k(i+1)}) \geq \gamma \theta_{k(i+1)} > 0, \quad \forall i = 1, 2, \dots$$

This implies that $\{f(x_{k(i)})\}_{i \in \mathbb{N}}$ is decreasing, and taking into account that $\{f_k\}_{k=1}^{\infty}$ is bounded below it is then obtained

$$\sum_{i=1}^{\infty} \theta_{k(i)} < \infty.$$

Obtaining that $\theta_{k(i)} \rightarrow 0$. The convergence of θ_k to zero follows then finally from (18). \square

4 Convergence results

In the remainder of this article we assume the following hypotheses:

- (i) The points sampled by the algorithm lie in a nonempty compact set $X \subset \mathbb{R}^n$.
- (ii) The MFCQ condition is satisfied at each feasible point of (P) lying in the set X .
- (iii) There exists a positive constant $M > 0$ such that $\|B^k\|_2 \leq M$ for all k .

This set of hypotheses will be called (H) in the sequel. Without loss of generality we suppose that the quantities $\|\nabla f(x)\|_2$, $\|D^2 f(x)\|_2$, $\|Dh_i(x)\|_2$, $\|D^2 h_i(x)\|_2$, $\|DG(x)\|_{Fr}$ and $\|D^2 G(x)\|_{Fr}$ are all bounded by M on a sufficiently large compact convex set containing X . In what follows we just denote $B = B^k$ avoiding the index notation.

Remark 10. Let us note that all points of the form $x_k + d$ rejected by the algorithm in Steps 2 or 5 belong to a ρ^{\max} -neighborhood of points x_k , that were already accepted in Steps 1 or 7. Therefore, Assumption (i) applies, *a posteriori*, to all points generated, by the algorithm including even the trial points that are finally rejected.

The first lemma of this section provides useful bounds for the variation of the feasibility and the objective functions, when feasible points of the local approximated problem are considered.

Lemma 11. *Let the above assumptions (H) be fulfilled. If d is a feasible point of $QP(x_k, \rho)$, then the following inequalities hold true.*

$$f(x_k + d) - \left(f(x_k) + \nabla f(x)^\top d + \frac{1}{2} d^\top B d \right) \leq n\rho^2 M, \quad (19a)$$

$$|h_i(x_k + d)| \leq \frac{1}{2} n\rho^2 M, \quad i = 1, \dots, p \quad (19b)$$

$$G(x_k + d) \leq \frac{1}{2} (n\rho^2 M) I \quad (19c)$$

Proof. The first inequality is a direct consequence of a second order approximation of f at x_k

$$f(x_k + d) = f(x_k) + \nabla f(x)^\top d + \frac{1}{2} d^\top D^2 f(x_k + \xi d) d,$$

where $\xi \in (0, 1)$. Indeed,

$$\begin{aligned} f(x_k + d) - \left(f(x_k) + \nabla f(x)^\top d + \frac{1}{2} d^\top B d \right) &= \frac{1}{2} d^\top (D^2 f(x_k + \xi d) - B) d \\ &\leq \frac{1}{2} \|d\|_2^2 \|D^2 f(x_k + \xi d) - B\|_2 \leq \|d\|_2^2 M \leq n \|d\|_\infty^2 M \leq n\rho^2 M. \end{aligned}$$

The second inequality follows in the same way, since

$$|h_i(x_k + d)| = \left| \frac{1}{2} d^\top D^2 h_i(x_k + \xi d) d \right| \leq \frac{1}{2} \|d\|_2^2 \|D^2 h_i(x_k + \xi d)\|_2 \leq \frac{1}{2} n\rho^2 M.$$

The same idea used for G leads to

$$G(x_k + d) = G(x_k) + DG(x_k)d + \frac{1}{2} d^\top D^2 G(x_k + \xi d) d \leq \frac{1}{2} d^\top D^2 G(x_k + \xi d) d.$$

Using now (9) we obtain

$$\lambda_1(d^\top D^2 G(x_k + \xi d) d) \leq \|d\|_2^2 \|D^2 G(x_k + \xi d)\|_{Fr} \leq n\rho^2 M.$$

This is equivalent to the condition

$$d^\top D^2 G(x_k + \xi d) d - (n\rho^2 M) I \leq 0,$$

that trivially implies (19c). \square

The next result gives a condition on ρ that ensures the acceptability of new iterates for previous filters as required in Step 4.

Lemma 12. *Let the assumptions (H) be satisfied. If*

$$\rho^2 \leq \frac{2\beta\tau^{k-1}}{(p+1)nM}$$

and d is a feasible point of $QP(x_k, \rho)$, then $(\theta(x_k + d), f(x_k + d))$ is acceptable to the filter \mathcal{F}^{k-1} .

Proof. It suffices to prove the inequality

$$\theta(x_k + d) \leq \beta\tau^{k-1}. \tag{20}$$

In virtue of the definition of θ and inequalities (19b) and (19c), it follows that

$$\theta(x_k + d) = \|h(x_k + d)\|_2 + \lambda_1(G(x_k + d))_+ \leq (p+1)\frac{1}{2}n\rho^2M.$$

Inequality (20) is now straightforwardly obtained from the assumption

$$\rho^2 \leq \frac{2\beta\tau^{k-1}}{(p+1)nM}. \quad \square$$

Now we proceed to extend the main lemma of [15].

Lemma 13. *Let us assume the set of hypotheses (H) concerning the problem (P). Let us further suppose \bar{x} to be a feasible point that is not a critical one. Then there exist a neighborhood \mathcal{N} of \bar{x} and strictly positive constants ϵ , μ , and κ such that if $x \in \mathcal{N} \cap X$, $\rho > 0$ and*

$$\mu\theta(x) \leq \rho \leq \kappa, \tag{21}$$

then there exists a feasible point d of the problem $QP(x, \rho)$ satisfying the following inequalities

$$f(x + d) + \gamma\theta(x + d) \leq f(x), \tag{22a}$$

$$f(x + d) - \sigma(\nabla f(x))^\top d + \frac{1}{2}d^\top Bd \leq f(x), \tag{22b}$$

$$\nabla f(x)^\top d + \frac{1}{2}d^\top Bd < -\frac{1}{3}\rho\epsilon. \tag{22c}$$

Moreover, the above inequalities hold also true at any optimal solution of $QP(x, \rho)$.

Proof. Due to hypotheses (H) the MFCQ condition holds at the feasible point \bar{x} . Then, since \bar{x} is not a critical point, it follows from Lemma 3 the existence of a unitary vector \bar{s} and a real number $\eta > 0$ such that conditions (10) are satisfied.

Let $x \in \mathcal{N}$, with \mathcal{N} a neighborhood of \bar{x} . Denote by $P_x(y)$ the orthogonal projection of a vector $y \in \mathbb{R}^n$ onto $\text{Ker } Dh(x)$. Notice that, since $Dh(\cdot)$ is continuous, vectors $\nabla h_j(x)$, with $j = 1, \dots, p$, are linear independent when x is close enough to \bar{x} . Thus, we can write $P_x(y) = (I - A_x^\dagger A_x)y$ for all y , where $A_x := Dh(x)$ and A^\dagger denotes the Moore-Penrose inverse of the matrix A . Recall that when the matrix A is surjective it follows that $A^\dagger = A^\top(AA^\top)^{-1}$.

Denote by $q = q(x) := -A_x^\dagger h(x)$ and $\phi = \phi(x) := \|q\|_2$. In order to directly deal with conditions related to equality constraints $h(x) = 0$, we project the vector \bar{s} onto $\text{Ker } Dh(x)$, normalize it, and obtain then $s = s(x) := P_x(\bar{s})/\|P_x(\bar{s})\|_2$. Note that s depends continuously on $x \in \mathcal{N}$, and $s = \bar{s}$ when $x = \bar{x}$. Thus, it follows from conditions (10a) and (10c) that there exist a constant $\epsilon > 0$ and a (smaller) bounded neighborhood \mathcal{N} of \bar{x} such that

$$\nabla f(x)^\top s < -\epsilon \quad \text{and} \quad G(x) + \eta DG(x)s < -\epsilon I, \quad \forall x \in \mathcal{N}. \quad (23)$$

Let us now fix some $\kappa > 0$ satisfying the following inequality

$$\kappa < \min \left\{ \eta, \frac{\epsilon}{M}, \frac{1}{3} \frac{(1-\sigma)\epsilon}{nM}, \frac{2}{3} \frac{\sigma\epsilon}{\gamma(p+1)nM} \right\}. \quad (24)$$

The continuity properties of functions f , h , and G imply the existence of a constant $C > 0$ such that the following three inequalities hold on the (eventually reduced) neighborhood \mathcal{N} :

$$\begin{aligned} \phi &\leq C\theta(x), \quad \lambda_1(DG(x)q) \leq C\theta(x), \quad \text{and} \\ \left(\|\nabla f(x)\|_2 + \epsilon + \frac{1}{2}M\phi + \kappa M \right) \phi &\leq C\theta(x), \quad \forall x \in \mathcal{N}. \end{aligned} \quad (25)$$

Let us finally fix some $\mu > 0$ so that

$$\mu > \max \left\{ (1+C)(1+\eta/\epsilon), \frac{6C}{\epsilon} \right\}. \quad (26)$$

In what follows, we consider a fixed value $\rho > 0$ satisfying (21). Note that such a ρ exists, since the neighborhood \mathcal{N} can be taken small enough in order to ensure the inequality

$$\mu\theta(x) < \kappa, \quad \forall x \in \mathcal{N}.$$

Inequality (26) in particular implies that $\mu > C$, which together with (21) yields

$$\phi < \rho \quad \text{and} \quad 0 < \frac{\rho - \phi}{\eta} < 1. \quad (27)$$

We proceed to prove that $d_1 = q + (\rho - \phi)s$ is a feasible vector of problem $QP(x, \rho)$. Due to the definition of q and s the equality constraint $h(x) + Dh(x)d_1 = 0$ is satisfied. Also, since q and s are orthogonal, we get

$$\|d_1\|_2 = \sqrt{\phi^2 + (\rho - \phi)^2} = \sqrt{\rho^2 - 2\rho\phi + 2\phi^2} \leq \rho.$$

Hence, d_1 satisfies the trust region constraint $\|d_1\|_\infty \leq \rho$ of $QP(x, \rho)$. It is just left to verify the inequality constraint $G(x) + DG(x)d_1 \leq 0$. Condition (23) and the first inequality in (27) imply

$$\begin{aligned} G(x) + DG(x)d_1 &= G(x) + DG(x)q + (\rho - \phi)DG(x)s \\ &= \left(1 - \frac{\rho - \phi}{\eta}\right)G(x) + DG(x)q + \frac{\rho - \phi}{\eta}[G(x) + \eta DG(x)s] \\ &\preceq \left(1 - \frac{\rho - \phi}{\eta}\right)G(x) + DG(x)q - \epsilon \left(\frac{\rho - \phi}{\eta}\right)I. \end{aligned}$$

By applying the largest eigenvalue function $\lambda_1(\cdot)$ to this last matrix inequality, Weyl's theorem (e.g. [18]), and by using inequalities (27) and (25), respectively, we obtain

$$\begin{aligned} \lambda_1(G(x) + DG(x)d_1) &\leq \lambda_1\left(\left(1 - \frac{\rho - \phi}{\eta}\right)G(x) + DG(x)q\right) - \epsilon \left(\frac{\rho - \phi}{\eta}\right) \\ &\leq \left(1 - \frac{\rho - \phi}{\eta}\right)\lambda_1(G(x)) + \lambda_1(DG(x)q) - \epsilon \left(\frac{\rho - \phi}{\eta}\right) \\ &\leq \left(1 - \frac{\rho - \phi}{\eta}\right)\lambda_1(G(x))_+ + \lambda_1(DG(x)q) - \epsilon \left(\frac{\rho - \phi}{\eta}\right) \\ &\leq \lambda_1(G(x))_+ + \lambda_1(DG(x)q) - \epsilon \left(\frac{\rho - \phi}{\eta}\right) \\ &\leq (1 + C)\theta(x) + \frac{\epsilon}{\eta}\phi - \frac{\epsilon}{\eta}\rho \leq \left(1 + C + \frac{\epsilon}{\eta}C\right)\theta(x) - \frac{\epsilon}{\eta}\rho \\ &\leq \frac{\epsilon}{\eta}\mu\theta(x) - \frac{\epsilon}{\eta}\rho \leq 0. \end{aligned}$$

The last two inequalities follow directly from (26) and (21), respectively. We thus conclude that d_1 is feasible for problem $QP(x, \rho)$.

It is left to prove that inequalities (22) hold at $d = d_1$. Let us start by showing (22c). By definition of d_1 we have

$$\begin{aligned} \nabla f(x)^\top d_1 + \frac{1}{2} d_1^\top B d_1 &= \nabla f(x)^\top q + (\rho - \phi) \nabla f(x)^\top s + \frac{1}{2} q^\top B q \\ &\quad + (\rho - \phi) q^\top B s + \frac{1}{2} (\rho - \phi)^2 s^\top B s. \end{aligned}$$

Thanks to the upper bounds given in (H), and inequalities (23) and (27), it follows that

$$\nabla f(x)^\top d_1 + \frac{1}{2} d_1^\top B d_1 < \|\nabla f(x)\|_2 \phi - (\rho - \phi) \epsilon + \frac{1}{2} M \phi^2 + \rho M \phi + \frac{1}{2} \rho^2 M.$$

Using (21), (24), (25) and (26) leads then to

$$\begin{aligned} \nabla f(x)^\top d_1 + \frac{1}{2} d_1^\top B d_1 &< \|\nabla f(x)\|_2 \phi - \rho \epsilon + \phi \epsilon + \frac{1}{2} M \phi^2 + \kappa M \phi + \frac{1}{2} \rho \epsilon \\ &\leq -\frac{1}{2} \rho \epsilon + C \theta(x) \\ &\leq -\frac{1}{2} \rho \epsilon + \frac{1}{6} \epsilon \rho = -\frac{1}{3} \rho \epsilon. \end{aligned}$$

The relation (22c) is then proved. Let us finally show (22a) and (22b) for any feasible point d of $QP(x, \rho)$ satisfying the inequalities

$$\nabla f(x)^\top d + \frac{1}{2} d^\top B d \leq \nabla f(x)^\top d_1 + \frac{1}{2} d_1^\top B d_1 < -\frac{1}{3} \rho \epsilon, \quad (28)$$

which in particular implies (22c) for d . This analysis obviously includes the point d_1 and any optimal solution of $QP(x, \rho)$.

The relation (28) together with inequality (19a) of Lemma 11 provides

$$\begin{aligned} f(x+d) - f(x) - \sigma \left(\nabla f(x)^\top d + \frac{1}{2} d^\top B d \right) \\ &\leq n \rho^2 M + (1 - \sigma) \left(\nabla f(x)^\top d_1 + \frac{1}{2} d_1^\top B d_1 \right) \\ &\leq n \rho^2 M - (1 - \sigma) \frac{1}{3} \rho \epsilon. \end{aligned}$$

Applying (21) and (24) to the right hand side of the above inequality leads easily to (22b).

In order to prove (22a) note that the already proved relations (22b) and (28) imply

$$f(x+d) - f(x) \leq -\sigma \frac{1}{3} \rho \epsilon.$$

On the other hand, due to (19b) and (19c) of Lemma 11 we have

$$\gamma \theta(x+d) \leq \gamma(p+1) \frac{1}{2} n \rho^2 M.$$

Adding up both inequalities, it is obtained

$$f(x+d) - f(x) + \gamma \theta(x+d) \leq \rho \left(-\sigma \frac{1}{3} \epsilon + \rho \gamma(p+1) \frac{1}{2} n M \right).$$

Then (22a) follows easily from (21) and (24). \square

Corollary 14. *Let us consider an iteration point x_k of the algorithm lying in the neighborhood \mathcal{N} defined in Lemma 13 above. If the inequality*

$$\mu \theta(x_k) < \frac{1}{2} \min \left\{ \sqrt{\frac{2\beta\tau^{k-1}}{(p+1)nM}}, \bar{\rho}, \kappa \right\} \quad (29)$$

is satisfied, then x_k is an iteration of f -type satisfying

$$\rho_k \geq \frac{1}{2} \min \left\{ \sqrt{\frac{2\beta\tau^{k-1}}{(p+1)nM}}, \bar{\rho}, \kappa \right\}.$$

Proof. In the algorithm a sequence of auxiliary problem $QP(x_k, \rho_l)$ is solved for the decreasing values $\rho_l = \frac{\rho_0}{2^l}$, where $\rho_0 > \bar{\rho}$. By inequality (29) there must be an index \bar{l} such that

$$\mu \theta(x_k) < \frac{1}{2} \min \left\{ \sqrt{\frac{2\beta\tau^{k-1}}{(p+1)nM}}, \bar{\rho}, \kappa \right\} \leq \rho_{\bar{l}} < \min \left\{ \sqrt{\frac{2\beta\tau^{k-1}}{(p+1)nM}}, \bar{\rho}, \kappa \right\}. \quad (30)$$

By definition, x_k is a θ -type iteration if the inner loop ends adding a new point to the filter at Steps 2 or 6. Otherwise, the inner loop terminates at Step 7 with the same filter and x_k is of f -type. It suffices to show twofold. 1) First, that for

the values $\rho_l > \rho_{\bar{l}}$ the loop cannot end defining x_k as of θ -type. 2) Second, if the value $\rho_{\bar{l}}$ is achieved, then the inner loop ends with x_k as f -type iteration.

1) We proceed to prove the first case. From the definition of $QP(x_k, \rho)$, it is obvious that, when ρ decreases to $\rho_{\bar{l}}$, the feasible set of $QP(x_k, \rho)$ shrinks and its optimal value (denoted by $\text{val}(QP(x_k, \rho))$) increases. Since $QP(x_k, \rho_{\bar{l}})$ is feasible (which is a consequence of Lemma 13 and (30)), it follows that the inner loop will not stop at Step 2 for any $\rho_l > \rho_{\bar{l}}$.

Note that the solution selected at Step 3, denoted by d^l , provides also the optimal value $\text{val}(QP(x_k, \rho_l)) = \nabla f(x_k)^\top d^l + \frac{1}{2} d^{l\top} B d^l$.

As mentioned before, the sequence $\text{val}(QP(x_k, \rho_l))$ increases when ρ_l decreases to $\rho_{\bar{l}}$. Taking now into account (30) we can conclude from (22c) in Lemma 13, that $\text{val}(QP(x_k, \rho_l)) < 0$. Consequently the filter will not be enlarged at Step 6 for any $\rho_l > \rho_{\bar{l}}$, since in fact this happens only for nonnegative optimal values.

2) Let us finally see that when $\rho_{\bar{l}}$ is achieved the inner loop stops without enlarging the filter. Since $d^{\bar{l}}$ is an optimal solution of $QP(x_k, \rho_{\bar{l}})$, Lemma 12 together with (22a) and the third inequality in (30) implies the acceptability of $(\theta(x_k + d^{\bar{l}}), f(x_k + d^{\bar{l}}))$ to the filter $\mathcal{F}^{k-1} \cup \{(\theta(x_k), f(x_k))\}$ at Step 4. Note also that the inner loop is not restarted at Step 5, due to the condition (22b). Furthermore, (22c) ensures that $(\theta(x_k), f(x_k))$ does not enter to the filter \mathcal{F}^{k-1} at Step 6.

Finally, cases 1) and 2) imply that $\rho_k \geq \rho_{\bar{l}} \geq \frac{1}{2} \min \left\{ \sqrt{\frac{2\beta\tau^{k-1}}{(\rho+1)nM}}, \bar{\rho}, \kappa \right\}$. \square

Let us study in the last partial result the finite termination of the inner loop.

Lemma 15. *Let the assumptions (H) be satisfied. Given a fixed iterate x_k , the inner loop generated between Steps 2 and 5 finishes after finitely many steps.*

Proof. If x_k is a critical point of (P), then $d = 0$ is also a critical point for $QP(x_k, \rho)$ and the inner loop stops at Step 3. We argue by contraction, that is, we suppose that x_k is not a critical point and that the loop is executed infinitely many times. In this case ρ converges to zero, because it is divided by two at those Steps (4 and 5) where the inner loop restarts.

In the case when x_k is not feasible, one of the following two inequalities occurs: either $\lambda_1(G(x_k)) > 0$ or $\|h(x_k)\|_2 > 0$. The second one implies that there exists some $i \in \{1, \dots, p\}$ with $|h_i(x_k)| > 0$. Thus, in virtue of the inequality

$$|h_i(x_k) + \nabla h_i(x_k)d| \geq |h_i(x_k)| - \rho \|\nabla h_i(x_k)\|_2,$$

it follows that $QP(x_k, \rho)$ becomes infeasible when ρ is sufficiently close to zero. In this case the inner loop is finished at Step 2.

On the other hand, by using Weyl's theorem [18], and inequalities (7) and (8), it is obtained

$$\begin{aligned} \lambda_1(G(x_k) + DG(x_k)d) &\geq \lambda_1(G(x_k)) + \lambda_n(DG(x_k)d) \\ &\geq \lambda_1(G(x_k)) - \|DG(x_k)d\|_{Fr} \\ &\geq \lambda_1(G(x_k)) - n\rho \|DG(x_k)\|_{Fr}. \end{aligned}$$

Therefore also in the case when $\lambda_1(G(x_k)) > 0$, the problem $QP(x_k, \rho)$ becomes infeasible for any sufficiently small ρ .

It remains to consider the case of a feasible x_k . Due to assumptions (H), the MFCQ condition holds at x_k . From Lemma 3 we obtain the existence of a unitary vector \bar{s} and a positive number $\bar{\eta}$ such that the following conditions are satisfied for all $\eta \in (0, \bar{\eta}]$

$$\nabla f(x_k)^\top \bar{s} < 0, \tag{31a}$$

$$Dh(x_k)\bar{s} = 0, \tag{31b}$$

$$G(x_k) + \eta DG(x_k)\bar{s} < 0. \tag{31c}$$

Let us denote $\delta = -\nabla f(x_k)\bar{s}$. We now proceed to show that the inner loop finishes for any $\rho > 0$ satisfying the following relation

$$\rho \leq \min \left\{ \bar{\eta}, \frac{\delta}{M}, \frac{(1 - \sigma)\delta}{2nM}, \frac{\sigma\delta}{\gamma(p + 1)nM}, \sqrt{\frac{2\beta\tau^{k-1}}{(p + 1)nM}} \right\}. \tag{32}$$

Recall that Lemma 6 provides the positivity of τ^{k-1} .

Let us define $\bar{d} = \rho\bar{s}$. Since x_k is feasible, it follows from (31b), (31c) and (32) that \bar{d} is a feasible point of $QP(x_k, \rho)$. Let us now estimate the value

of the objective function of the auxiliary problem $QP(x_k, \rho)$ at \bar{d} . From the assumption (H) we obtain

$$\nabla f(x_k)^\top \bar{d} + \frac{1}{2} \bar{d}^\top B \bar{d} = -\rho\delta + \frac{1}{2} \rho^2 \bar{s}^\top B \bar{s} \leq -\rho\delta + \frac{1}{2} \rho^2 M.$$

It follows from (32) that

$$\nabla f(x_k)^\top \bar{d} + \frac{1}{2} \bar{d}^\top B \bar{d} \leq -\frac{1}{2} \rho\delta.$$

Now let us suppose that d is the solution of $QP(x_k, \rho)$ selected at Step 3. Due to the feasibility of \bar{d} we get

$$\nabla f(x_k)^\top d + \frac{1}{2} d^\top B d \leq -\frac{1}{2} \rho\delta < 0. \quad (33)$$

According to Lemma 11, the following inequality holds

$$f(x_k) - f(x_k + d) \geq -\left(\nabla f(x_k)^\top d + \frac{1}{2} d^\top B d\right) - n\rho^2 M.$$

Then, relations (32) and (33) yield to the following inequalities

$$\begin{aligned} f(x_k) - f(x_k + d) &\geq -\sigma \left(\nabla f(x_k)^\top d + \frac{1}{2} d^\top B d\right) + \frac{1}{2} (1 - \sigma) \rho\delta - n\rho^2 M \\ &\geq -\sigma \left(\nabla f(x_k)^\top d + \frac{1}{2} d^\top B d\right), \end{aligned}$$

where the parameter $\sigma \in (0, 1)$ was used in Step 5 of the algorithm. Therefore, we have already deduced that

$$f(x_k) + \sigma \left(\nabla f(x_k)^\top d + \frac{1}{2} d^\top B d\right) \geq f(x_k + d), \quad (34a)$$

$$f(x_k) - f(x_k + d) \geq \sigma \frac{1}{2} \rho\delta. \quad (34b)$$

Consequently, (34a) implies that the inner loop is not restarted at Step 5.

It remains only to prove that the algorithm goes from Step 4 to Step 5 (i.e. the inner loop does not restart at Step 4). This situation happens when $(\theta(x_k + d), f(x_k + d))$ is acceptable to the filter $\mathcal{F}^{k-1} \cup \{(\theta(x_k), f(x_k))\}$. By Lemma 12

and condition (32), it follows that $(\theta(x_k + d), f(x_k + d))$ is acceptable for \mathcal{F}^{k-1} . On the other hand, it can be seen from Lemma 11 that

$$\theta(x_k + d) \leq p \left(\frac{1}{2} n \rho^2 M \right) + \frac{1}{2} n \rho^2 M = \frac{1}{2} (p + 1) n \rho^2 M.$$

This together with (34b) leads to

$$f(x_k) - f(x_k + d) - \gamma \theta(x_k + d) \geq \sigma \frac{1}{2} \rho \delta - \gamma \frac{1}{2} (p + 1) n \rho^2 M.$$

The right hand side of the last expression is positive due to the definition of ρ in (32), obtaining that $f(x_k + d) + \gamma \theta(x_k + d) \leq f(x_k)$. This actually means that $(\theta(x_k + d), f(x_k + d))$ is also acceptable for $\{(\theta(x_k), f(x_k))\}$, which completes the proof. \square

We can now state and prove the global convergence theorem for the algorithm proposed.

Theorem 16. *Suppose that assumptions (H) holds true. Consider the sequence $\{x_k\}$ generated by the Filter-SDP algorithm defined above. Then one of the following situations occurs:*

1. *The restoration phase (Step 1) fails to find a vector x_k satisfying (A1) and (B1).*
2. *A critical point of (P) is found, that is, $d = 0$ solves the auxiliary problem $QP(x_k, \rho)$ for some iteration k .*
3. *There exists an accumulation point of $\{x_k\}$ that is a critical point of (P).*

Proof. We suppose that situations 1. and 2. do not happen, and proceed to prove that situation 3. occurs.

Due to hypothesis (i) in (H) the sequence $\{x_k\}_{k=1}^{\infty}$ sampled by the algorithm lies in a compact set $X \subset \mathbb{R}^n$ and has therefore at least one accumulation point.

Firstly, consider the case when the algorithm generates infinitely many θ -type iterations. Let us denote by $\{x_{k(j)}^{\theta}\}_{j=1}^{\infty}$ the respective subsequence of θ -type iterates. From Lemma 9 it follows $\theta(x_{k(j)}^{\theta}) \rightarrow 0$, and consequently $\tau^k \rightarrow 0$.

Since the decreasing sequence τ^k changes only at iterations of θ -type, we can suppose passing to a subsequence if necessary that for all j

$$\theta(x_{k(j)}^\theta) = \tau^{k(j)} < \tau^{k(j)-1}, \tag{35}$$

and that $\{x_{k(j)}^\theta\}$ converges to a feasible point \bar{x} . Let us argue by contradiction and assume that \bar{x} is not a critical point of (P).

From hypothesis (ii) in (H) we know that MFCQ is satisfied at \bar{x} . Hence, Lemma 13 can be applied, obtaining the existence of the corresponding neighborhood \mathcal{N} of \bar{x} , and of the strictly positive constants ϵ , μ , and κ . Note that the iterate $x_{k(j)}^\theta$ will belong to \mathcal{N} for sufficiently large j . It suffices then to show that the condition

$$\mu\theta(x_{k(j)}^\theta) < \frac{1}{2} \min \left\{ \sqrt{\frac{2\beta\tau^{k(j)-1}}{(p+1)nM}}, \bar{\rho}, \kappa \right\} \tag{36}$$

is also satisfied for sufficiently large j . Indeed, condition (36) and Corollary 14 imply that $x_{k(j)}^\theta$ is an iteration of f -type, which is a contradiction.

Since $\theta(x_{k(j)}^\theta)$ and $\tau^{k(j)-1}$ converge both to zero, for all j large enough, condition (36) actually reduces to the following one

$$\theta(x_{k(j)}^\theta) < \frac{1}{2\mu} \sqrt{\frac{2\beta\tau^{k(j)-1}}{(p+1)nM}}. \tag{37}$$

We thus conclude by noting that (37) is obtained from (35) when j is taken large enough such that

$$\tau^{k(j)-1} \leq \frac{\beta}{2\mu^2(p+1)nM}.$$

This proves the desired result for the first case.

Now, consider the case when only a finite number of θ -type iterations happen. Therefore, there exists an index K so that all iterations $k \geq K$ are of f -type.

In particular, it follows from (15) that $\{f(x_k)\}_{k \geq K}$ is strictly decreasing. This together with the fact that $(\theta(x_{k+1}), f(x_{k+1}))$ is acceptable to $(\theta(x_k), f(x_k))$ (cf. Step 4) implies, via Lemma 8, that $\theta(x_k) \rightarrow 0$. Let us suppose, without loss of generality and arguing by contradiction, that $\{x_k\}$ converges to a feasible vector \bar{x} , which is not a critical point of (P).

Analogously to the first case, hypotheses of Lemma 13 hold, providing the existence of the corresponding neighborhood \mathcal{N} of \bar{x} , and suitable constants ϵ , μ , and κ . Since $x_k \rightarrow \bar{x}$, it can be assumed that $x_k \in \mathcal{N}$ for all $k \geq K$.

By definition of an f -type iteration, no filter updates are made for $k \geq K$. Consequently, $\tau^k = \tau^K$ is constant, for all $k \geq K$. Notice also that condition (29) is satisfied for sufficiently large k , since its right hand side is constant and $\theta(x_k)$ converges to zero. Then, due to Corollary 14, it holds that

$$\rho_k \geq \hat{\rho} = \frac{1}{2} \min \left\{ \sqrt{\frac{2\beta\tau^K}{(p+1)nM}}, \bar{\rho}, \kappa \right\},$$

for all $k \geq K$.

By using (15) and the decreasing monotonicity on ρ of $\text{val}(QP(x, \rho))$, it yields to

$$\begin{aligned} f(x_k) - f(x_{k+1}) &\geq -\sigma \left(\nabla f(x_k)^\top d_k + \frac{1}{2} d_k^\top B d_k \right) \\ &= -\sigma \text{val}(QP(x_k, \rho_k)) \\ &\geq -\sigma \text{val}(QP(x_k, \hat{\rho})). \end{aligned} \tag{38}$$

Taking now into account that $\hat{\rho}$ satisfies (21) and using (22c) of Lemma 13, we obtain for all k large enough,

$$\begin{aligned} f(x_k) - f(x_{k+1}) &\geq -\sigma \text{val}(QP(x_k, \hat{\rho})) \\ &= -\sigma \left(\nabla f(x_k)^\top \hat{d} + \frac{1}{2} \hat{d}^\top B \hat{d} \right) \\ &\geq \frac{1}{3} \sigma \epsilon \hat{\rho}, \end{aligned} \tag{39}$$

where \hat{d} is some optimal solution of $QP(x_k, \hat{\rho})$.

On the other hand, since $\{f(x_k)\}_{k \geq K}$ is strictly decreasing, and since $\{x_k\}$ lies in a compact X , it follows that $f(x_k)$ converges. This is a contradiction because the right hand side of (39) is a strictly positive constant. The theorem is concluded. □

Remark 17. At Step 3 it is supposed that the auxiliary problem $QP(x_k, \rho)$ can solve exactly. When B is positive semidefinite, this assumption is theoretically not strong due to the convexity of $QP(x_k, \rho)$ and its equivalence to a linear

SDP problem. If the matrix B is not positive semidefinite, there is no theoretical or practical approach to identify global optimizers of the auxiliary problem.

In practice, the solution of $QP(x_k, \rho)$ can be obtained, even in the convex case, only up to certain degree of exactness. This fact can be taken into account just reformulating the algorithm such that only an inexact solution of the auxiliary problems is required in Step 3. The proof of convergence can be adapted as long as the suboptimal solution d provided in Step 3 satisfy, when the optimal value $\text{val}(QP(x_k, \rho))$ is negative, a uniform degree of exactness like

$$\nabla f(x_k)^\top d + \frac{1}{2}d^\top B d \leq \gamma \text{val}(QP(x_k, \rho)),$$

for a sufficient small $\gamma \in (0, 1)$.

5 Numerical Experiments

A MATLAB[®] code (version 6.0) was written for the filter algorithm presented in Section 3. We used Jos Sturm' SeDuMi code (cf. [46]) for testing the feasibility of problem QP and for solving it at Steps 1 and 3 of the algorithm. The link between the MATLAB[®] code and the SeDuMi was provided by the parser YALMIP (see [36]).

The restoration phase described at Step 1 has been implemented in a very simple way. It is just tried to obtain an (almost) feasible point by minimizing the nondifferentiable merit function θ . The routine employed for this purpose neither uses derivative information nor exploits in any way the structure of the problems solved. Consequently in this implementation the restoration step is not efficient. We expect to improve this issue in future works.

In order to make a preliminary test of the algorithm we selected a set of twenty small scale examples of the publicly available benchmark collection COMPluib [32]. For details about the library and a quick introduction see also [33], and the references therein.

With the data contained in COMPluib it is possible to construct particular nonlinear semidefinite optimization problems arising in feedback control design (see [32]). We considered in our numerical tests only the basic *Static (or reduced order) Output Feedback*, SOF- \mathcal{H}_2 problem. The reader can find more details on the motivation of this problem, for instance, in [22, 32, 34, 35].

The following NLSDP formulation of the SOF- \mathcal{H}_2 problem was used (see, for instance, [32, 33]).

$$\min_{F,L} \{Tr(LQ_F); A_FL + LA_F^\top + P = 0, A_FL + LA_F^\top < 0, L > 0\} \text{ (SOF)}$$

Here

$$A_F = A + BFC, \quad Q_F = C^\top F^\top RFC + Q \quad \text{and} \quad C_F = C_1 + D_{12}FC$$

and all the matrices involved are real with the appropriate dimension. The data A, B_1, B, C_1, C and D_{12} were extracted from COMpleib. For P, Q, R we took always the identity matrix. In the problem SOFP the variables are the matrices L and F . The L is symmetric and real, but the F , associated with SOF control law, is in general not square.

For all these numerical tests, we employed a common tolerance value of $\varepsilon_{tol} = 10^{-3}$. At each iteration the matrix B was selected as the identity in the trust region problem (11). This implies that the algorithm is not using second order information in this implementation. Thus, a linear convergence behavior can be expected.

In order to calculate a start point x_0 we implemented and run algorithm (SLPMM) stated in [30, 31]. The initial matrices used therefore satisfy the SDP-inequalities of (SOF). We observed that this kind of interior property was not preserved along the whole optimization process, i.e. some iterates not satisfying the SDP-inequalities were obtained.

Each example was run twice: First using the infinity norm in (11) and, then using the euclidean norm. In the next table the numerical results with better cpu-time are given. In particular, the fifth column provides information about, which norm resulted with the better performance for each example (the rows). In some of the examples, running the algorithm with a different norm than the given on the fifth column caused not only a larger cpu time, but also a failure of the restoration step. As can be seen, the euclidean norm performed better in more examples than the infinity norm, but in most of the cases the difference in cpu time was not very significant.

Data set	n	p	m	TR norm	f -iter	θ -iter	Rest.	Filter	cpu time	$f(x_*)$	$f(x_0)$	$\theta(x_0)$
AC1	24	15	5	Eucl.	359	43	14	20	224.56	20.03	1768.36	282.51
AC2	24	15	5	Eucl.	42	49	9	22	97.68	20.03	5.56	1.06
AC3	23	15	5	Eucl.	10	42	1	31	36.37	21.84	18.45	1.01
AC4	12	10	4	Eucl.	60	10	6	6	32.90	11.99	10.31	3.45
AC15	16	10	4	Inf.	226	144	13	7	561.89	159.07	159.69	0.73
AC17	12	10	4	Eucl.	9	40	5	5	19.52	14.63	9.31	0.75
DIS1	52	36	8	Eucl.	46	2	0	2	56.20	15.36	337.65	61.25
DIS2	10	6	3	Eucl.	26	4	0	2	8.03	8.60	16.89	0.66
DIS3	37	21	6	Eucl.	26	2	0	3	10.76	5.99	9.98	1.22
HE1	12	10	4	Eucl.	281	1	0	2	74.43	13.31	4.52	0.99
HF2D13	23	15	5	Eucl.	630	5	0	3	194.59	0.51	1.39	1.01
HF2D15	23	15	5	Inf.	770	6	4	3	302.16	1.49	2.32	0.63
HF2D17	23	15	5	Eucl.	122	4	0	3	39.27	0.76	1.49	1.04
HF2D_CD4	32	28	7	Inf.	5	10	2	11	36.61	0.80	1.19	0.86
HF2D_CD5	32	28	7	Inf.	14	18	5	14	334.93	2.31	3.19	0.77
HF2D_IS7	19	15	5	Eucl.	4	3	0	4	9.21	0.37	0.69	0.63
IH	320	210	21	Inf.	15	11	0	10	52.11	42.30	80.46	1.76
NN2	4	3	2	Inf.	126	5	0	4	26.89	3.46	136.97	33.72
NN4	16	10	4	Eucl.	32	3	0	4	11.69	5.41	5.11	0.84
NN8	16	6	3	Eucl.	37	1	0	2	11.54	4.44	173.34	78.14

where

- Data set = name of the example in COMPLEIB
n = dimension of the variable $x = (F, L)$, (symmetry of L considered!)
p = number of equality constraints
m = dimension of the matrices involved into the SDP constraints
TR norm = norm (Euclidean or Infinity) used in the subproblems (11).
 f -iter = number of f -iterations
 θ -iter = number of θ -iterations
Rest. = number of times the restoration phase was used
Filter = final number of vectors in the Filter
cpu time = total cpu time (sec.) including restoration and the inner loops
 $f(x_*)$ = value of f at the optimum
 $f(x_0)$ = value of f at the initial point (F_0, L_0)
 $\theta(x_0)$ = value of θ at the initial point

6 Conclusions

The filter SQP algorithm has been extended from [15] to the case of nonlinear semidefinite constraints. The global convergence of the filter method proposed was obtained under quite mild assumptions, like MFCQ, boundedness, etc.

The trust region subproblems at each step of the algorithm are actually linear semidefinite programming problems. In this sense the approach selected is also related to the SSDP algorithm presented in [12].

We have performed some numerical experiments applied to optimal SOF problems. In a future work we plan to study local convergence properties and to report new numerical results.

Acknowledgments. The authors would like to thank the editor Prof. José Mario Martínez and to an anonymous referee for the valuable comments that helped us to improve the presentation of the paper.

REFERENCES

- [1] F. Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*. SIAM J. Optim., **5**(1) (1995), 13–51.
- [2] J. Amaya and J.A. Gómez, *Strong duality for inexact linear programming problems*. Optimization, **49**(3) (2001), 243–269.
- [3] P. Apkarian, D. Noll and D. Tuan, *Fixed-order H_1 control design via a partially augmented lagrangian method*. Int. J. Robust Nonlinear Control, **13**(12) (2003), 1137–1148.
- [4] A. Ben-Tal, F. Jarre, M. Kočvara, A. Nemirovski and J. Zowe, *Optimal design of trusses under a nonconvex global buckling constraint*. Optimization and Engineering, **1**(2) (2000), 189–213.
- [5] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization*. MPS-SIAM, Series on Optimization. SIAM Philadelphia (2001).
- [6] H.Y. Benson, D.F. Shanno and R.J. Vanderbei, *Interior-point methods for nonconvex nonlinear programming: Filter methods and merit functions*. Comput. Optim. Appl., **23**(2) (2002), 257–272.
- [7] J.F. Bonnans and A. Shapiro, *Perturbation analysis of optimization problems*. Springer-Verlag, New York (2000).
- [8] S. Boyd, L.E. Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in system and control theory*. SIAM Studies in Applied Mathematics. Vol. 15., SIAM, Philadelphia (1994).
- [9] R. Correa and H. Ramírez C, *A global algorithm for nonlinear semidefinite programming*. SIAM J. on Optimization, **15**(1) (2004), 303–318.
- [10] M. Diehl, F. Jarre and C.H. Vogelbusch, *Loss of superlinear convergence for an SQP-type method with conic constraints*. SIAM Journal of Optimization, **16**(4) (2006), 1201–1210.

- [11] B. Fares, P. Apkarian and D. Noll, *An augmented Lagrangian method for a class of lmiconstrained problems in robust control theory*. Int. J. Control, **74**(4) (2001), 348–360.
- [12] B. Fares, D. Noll and P. Apkarian, *Robust control via sequential semidefinite programming*. SIAM J. Control Optim., **40**(6) (2002), 1791–1820.
- [13] R. Fletcher, N.I.M. Gould, S. Leyffer, Ph.L. Toint and A. Wächter, *Global convergence of a trust-region SQP-filter algorithms for nonlinear programming*. SIAM J. Optim., **13**(3) (2002), 635–659.
- [14] R. Fletcher and S. Leyffer, *Nonlinear programming without a penalty function*. Mathematical Programming, **91**(2, Ser. A) (2002), 239–269.
- [15] R. Fletcher, S. Leyffer and Ph.-L. Toint, *On the global convergence of a filter-SQP algorithm*. SIAM J. Optim., **13**(1) (2002), 44–59.
- [16] R.W. Freund and F. Jarre, *A sensitivity analysis and a convergence result for a sequential semidefinite programming method*. Technical report, No. 03-4-09, Bell Laboratories, Murray Hill, New Jersey (2003).
- [17] R.W. Freund, F. Jarre and C.H. Vogelbusch, *Nonlinear semidefinite programming: sensitivity, convergence, and an application in passive reduced-order modeling*. Mathematical Programming, **109**(2-3) (2007), 581–611.
- [18] G.H. Golub and C.F. Van Loan, editors. *Matrix Computations*, Third Edition. The Johns Hopkins University Press, Baltimore (1996).
- [19] F.A.M. Gomes, *A sequential quadratic programming algorithm that combines merit function and filter ideas*. Computational and Applied Mathematics, **26**(3) (2007), 337–379.
- [20] W. Gómez and J.A. Gómez, *Cutting plane algorithms for robust conic convex optimization problems*. Optimization Methods and Software, **21**(5) (2006), 779–803.
- [21] C. Helmberg and F. Rendl, *A spectral bundle method for semidefinite programming*. SIAM J. Optimization, **10**(3) (2000), 673–696.
- [22] Y.S. Hung and A.G.J. MacFralane, *Multivariable Feedback: A quasi-classical approach*. Lectures Notes in Control and Inform., Sci. 40, Springer-Verlag, Berlin, Heidelberg, New York (1982).
- [23] F. Jarre, *An interior method for nonconvex semidefinite programs*. Optimization and Engineering, **1**(4) (2000), 347–372.
- [24] Y. Kanno and I. Takewaki, *Sequential semidefinite program for maximum robustness design of structures under load uncertainty*. Journal of Optimization Theory and Applications, **130**(2) (2006), 265–287.
- [25] M. Kočvara and M. Stingl, *Pennon: a code for convex nonlinear and semidefinite programming*. Optimization Methods and Software, **18**(3) (2003), 317–333.

- [26] M. Kočvara and M. Stingl, *Pennon: a generalized augmented Lagrangian method for semidefinite programming*. In Di Pillo, Gianni (ed.) et al., High performance algorithms and software for nonlinear optimization, Boston, MA: Kluwer Academic Publishers. Appl. Optim., **82** (2003), 303–321.
- [27] M. Kočvara and M. Stingl, *Solving nonconvex SDP problems of structural optimization with stability control*. Optimization Methods and Software, **19**(5) (2004), 595–609.
- [28] K. Krishnan and J.E. Mitchell, *Semi-infinite linear programming approaches to semidefinite programming problems*. In Novel approaches to hard discrete optimization problems (P. Pardalos and H. Wolkowicz eds.), Fields Institute Communications Series, **37** (2003), 123–142.
- [29] S. Kurcyusz, *On the existence and nonexistence of Lagrange multipliers in banach spaces*. Journal of Optimization Theory and Applications, **20**(1) (1976), 81–110.
- [30] F. Leibfritz, *Static Output Feedback Design Problems*. Shaker Verlag, Aachen, Germany (1998).
- [31] F. Leibfritz, *Computational design of stabilizing static output feedback controllers*. Technical report, 99-01, Universität Trier, Trier, Germany (1999).
- [32] F. Leibfritz, *Compleib: Constrained matrix-optimization problem library- a collection of test examples for nonlinear semidefinite programming, control system design and related problems*. Technical report, Universität Trier, Trier, Germany (2004).
- [33] F. Leibfritz and W. Lipinski, *Compleib 1.0 – user manual and quick reference*. Technical report, Universität Trier, Trier, Germany (2004).
- [34] F. Leibfritz and M.E. Mostafa, *An interior point constrained trust region method for a special class of nonlinear semidefinite programming problems*. SIAM J. Optimization, **12**(4) (2002), 1048–1074.
- [35] F. Leibfritz and M.E. Mostafa, *Trust region methods for solving the optimal output feedback design problem*. Int. J. Control, **76**(5) (2003), 501–519.
- [36] J. Löfberg, *YALMIP: A toolbox for modeling and optimization in MATLAB*. In CCA/ISIC/CACSD, September (2004).
- [37] L. Mosheyev and M. Zibulevsky, *Penalty/barrier multiplier algorithm for semidefinite programming*. Optimization Methods and Software, **13**(4) (2000), 235–261.
- [38] A. Nemirovskii, *Several np-hard problems arising in robust stability analysis*. Math. Control, Signals Syst., **6**(2) (1993), 99–105.
- [39] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*. SIAM Stud. Appl. Math. 13, SIAM Philadelphia (1994).
- [40] D. Noll and P. Apkarian, *Spectral bundle methods for non-convex maximum eigenvalue functions: first-order methods*. Mathematical Programming, **104**(2-3, Ser. B) (2005), 701–727.

- [41] D. Noll and P. Apkarian, *Spectral bundle methods for non-convex maximum eigenvalue functions: second-order methods*. Mathematical Programming, **104**(2-3, Ser. B) (2005), 729–747.
- [42] D. Noll, M. Torki and P. Apkarian, *Partially augmented lagrangian method for matrix inequality constraints*. SIAM J. Optimization, **15**(1) (2004), 181–206.
- [43] M. Overton and H. Wolkowicz (eds.), *Semidefinite programming*. Dedicated to the memory of Svatopluk Polyak. Mathematical Programming, **77**(2, Ser. B) (1997), 97–320.
- [44] S.M. Robinson, *Stability theorems for systems of inequalities*. Part II: differentiable nonlinear systems. SIAM J. Numerical Analysis, **13** (1976), 497–513.
- [45] C. Shen, W. Xue and D. Pu, *A filter SQP algorithm without a feasibility restoration phase*. Computational and Applied Mathematics, **28**(2) (2009), 167–194.
- [46] Jos F. Sturm, *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones*. Optim. Methods Softw., **11/12**(1-4) (1999), 625–653. Interior point methods.
- [47] M.J. Todd, *Semidefinite optimization*. Acta Numer., **10** (2001), 515–560.
- [48] M. Ulbrich, S. Ulbrich and L.N. Vicente, *A globally convergent primal-dual interior-point filter method for nonconvex nonlinear programming*. Mathematical Programming, **100**(2, Ser. A) (2004), 379–410.
- [49] S. Ulbrich, *On the superlinear local convergence of a filter-SQP method*. Math. Program., **100**(1, Ser. B) (2004), 217–245.
- [50] L. Vandenberghe and S. Boyd, *Semidefinite programming*. SIAM Review, **38**(1) (1996), 49–95.
- [51] A. Wächter and L.T. Biegler, *Line search filter methods for nonlinear programming: Local convergence*. SIAM J. Optimization, **16**(1) (2005), 32–48.
- [52] A. Wächter and L.T. Biegler, *Line search filter methods for nonlinear programming: Motivation and global convergence*. SIAM J. Optimization, **16**(1) (2005), 1–31.
- [53] H. Wolkowitz, R. Saigal and L. Vandenberghe (eds.), *Handbook of Semidefinite Programming: Theory, Algorithms and Applications*. Kluwer’s International Series in Operations Research and Management Science (2000).