Math. Program., Ser. A 89: 1–33 (2000)

Digital Object Identifier (DOI) 10.1007/s101070000166

François Oustry

A second-order bundle method to minimize the maximum eigenvalue function*

Received: February 9, 1998 / Accepted: May 2, 2000
Published online September 20, 2000 – © Springer-Verlag 2000

Abstract. In this paper we present a nonsmooth algorithm to minimize the maximum eigenvalue of matrices belonging to an affine subspace of $n \times n$ symmetric matrices. We show how a simple bundle method, the *approximate eigenvalue method* can be used to globalize the second-order method developed by M.L. Overton in the eighties and recently revisited in the framework of the $\mathcal{U}$-Lagrangian theory. With no additional assumption, the resulting algorithm generates a minimizing sequence. A geometrical and constructive proof is given. To prove that quadratic convergence is achieved asymptotically, some *strict complementarity* and *non-degeneracy* assumptions are needed. We also introduce new variants of bundle methods for semidefinite programming.

Key words. eigenvalue optimization – semidefinite programming – convex optimization – second-order bundle methods

1. Introduction

1.1. Overview

Eigenvalue optimization problems have a long history: as mentioned in [25], Lagrange had already stated in 1773 an eigenvalue optimization problem to design the shape of the strongest axially symmetric column with prescribed length, volume and boundary conditions. Yet it is only very recently that it became an independent area of research with both theoretical and practical aspects. Although the mathematical models of the underlying physical problems are generally not convex, it is notable that the area has very strong connections with convex analysis. In fact, these problems have often a composite structure with a convex component. The role of convex analysis was first emphasized by R. Bellman and K. Fan in [4]; more recently, this point of view was developed further in [15] and [25].

We consider here a basic eigenvalue optimization problem

$$(P) \qquad \inf_{x \in \mathbb{R}^m} \lambda_1(A(x))$$

F. Oustry: Inria, 655 avenue de l'Europe, 38330 Montbonnot, France.
e-mail: Francois.Oustry@inrialpes.fr

Mathematics Subject Classification (1991): 90C25, 52A41, 65K10, 15A18

* Received by the editors February 1998; revised version June 1999. This work was initiated with a scholarship from ENSTA/DGA (France) and was pursued with a post-doctoral scholarship from INRIA and the NSF grant 655302574100-F1852 at Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012 USA.

where $\lambda_1(X)$ is the largest eigenvalue of $X = A(x)$, element of $\mathcal{S}_n$ the space of $n \times n$ symmetric matrices and

$$\mathbb{R}^m \ni x \mapsto A(x) := A_0 + \mathcal{A}x \tag{1}$$

is affine: $A_0 \in \mathcal{S}_n$ and $\mathcal{A}$ is a linear operator from $\mathbb{R}^m$ to $\mathcal{S}_n$.

Existing numerical methods to solve (P) can be arranged in two classes: interior-point methods and nonsmooth optimization methods. The first interior-point methods for solving (P) (in the framework of *semidefinite programming*) were developed by Nesterov and Nemirovski [30]. With the exception of Nemirovski's projective method [31, 26], all the interior-point schemes proposed in the early 1990's (see the numerous references in [5, Chap. II, Notes and References]) were path-following or potential reduction methods. As recently explained in a clear survey by Yu. Nesterov [27], "classical" interior-point methods can be seen as a process to transform the initial problem into an equivalent one which can be solved "easily" thanks to an addition of *structure*: *self-concordance* is used to obtain the *polynomiality* of interior-point schemes [31]. A similar presentation can be done for *predictor-corrector* type methods using small neighborhoods; many variants of them can be found for semidefinite programming: to give only a sample we refer to [2, 38, 21].

Our approach is quite different but, as we will see, does not exempt us from finding a trade-off between global and local requirements, *i.e.*, between total complexity and speed of convergence. Starting directly from problem (P) itself, we will use a recent second-order theory, namely the $\mathcal{U}$-Lagrangian theory [24], to speed up the asymptotic convergence of a first-order method developed by Cullum, Donath and Wolfe [6] for a particular instance of (P) ($\mathcal{A}$ diagonal), and by Polak and Wardi [37] in a more general framework. Using the terminology of [16, Chap. XIII], the method can be seen as a *Markovian dual bundle method*: at each iteration an approximation of the ε-subdifferential is computed, via a bundling process, without using information from the previous iterations. We call it the *approximate eigenvalue method*. More recently a stabilization of the cutting planes algorithm was proposed in [39] and enriched in [19, 14, 23, 13] with *semidefinite models* of the objective function; this belongs to the class of *primal bundle methods* [16, Chap. XV] which are very efficient to solve large-scale problems with a moderate accuracy.

When high accuracy is needed, second-order information must be added in the model. Combining geometrical and Sequential Quadratic Programming approaches, a local algorithm was presented and analyzed in [10], [34], [36], [35] and [41]; in the latter two papers, a quadratic rate of convergence was obtained. However, in this SQP framework, the authors considered only local analysis; issues of global convergence were not addressed.

In this paper, we present, as in [33], the second-order analysis of the maximum eigenvalue function using the $\mathcal{U}$-Lagrangian theory [24] and we show how to use the approximate eigenvalue method to globalize the second-order algorithm while preserving asymptotically a quadratic rate of convergence.

Our paper is organized as follows. We first recall some well-known results on the first-order analysis of λ_1. Then using simple chain rules, we easily derive a first-order analysis of the composite function $f := \lambda_1 \circ A$. This enables us to simplify the approximate

eigenvalue method developed in Sect. 3: at a point x we consider the enlargement of the subdifferential of λ_1 obtained with eigenvectors associated with ε-maximal eigenvalues; this set $\delta_\varepsilon f(x)$ plays the role of an approximation of the the true ε-subdifferential $\partial_\varepsilon f(x)$. By measuring the quality of this approximation, we provide an explicit ε-strategy to ensure global convergence of the method. In Sect. 4, we present the second-order analysis of λ_1 using the $\mathcal{U}$-Lagrangian theory. We recall the main result of [33]: this theory provides us with a second-order development of λ_1 along a smooth manifold: the set $\mathcal{M}_r$ of matrices whose largest eigenvalue has multiplicity r. We derive similar results for f with a so-called transversality condition. Then, in Sect. 5, we show how to use the approximate eigenvectors (Sect. 2) to stabilize the second-order objects presented in Sect. 4. In particular we provide a constructive characterization of the projection of a matrix $X \in \mathcal{S}_n$ onto the manifold r_ε (Theorem 13). This results in a *second-order bundle method* which is globally and quadratically convergent. With no additional assumptions, a minimizing sequence is generated. Some *strict complementarity* and *non-degeneracy* assumptions are needed to guarantee the quadratic rate of convergence. In Sect. 6 using some duality, we explain how the approximate eigenvalue bundle method is related to a new generation of spectral proximal-type bundle methods in which second-order information can also be introduced. Finally we have chosen a numerical example from combinatorial optimization to illustrate a qualitative distinction between interior-point methods and second-order bundle methods: for the latter methods, superlinear convergence can be observed even when strict complementarity does not hold.

1.2. Basic notation and terminology

Our notation follows closely that of [24] and [15].
$\mathbb{R}^m$ m-dimensional Euclidean space
$x^T y$ scalar product of $x, y \in \mathbb{R}^m$
$\|x\| := \sqrt{x^T x}$ Euclidean norm of $x \in \mathbb{R}^m$
$\mathcal{U}^\perp$ orthogonal subspace of the subspace $\mathcal{U}$
$\mathrm{proj}_C : \mathbb{R}^m \to \mathcal{U}$ projection operator onto the closed convex set $C \subset \mathbb{R}^m$
aff C affine hull of the nonempty set $C \subset \mathbb{R}^m$
ri C relative interior of the convex set $C \subset \mathbb{R}^m$
span C linear subspace generated by the nonempty set $C \subset \mathbb{R}^m$
$B(x, \delta)$ open ball centered at $x \in \mathbb{R}^m$ with radius $\delta > 0$
$\mathbb{R}^m \ni d \mapsto \sigma_C(d) := \sup_{g \in C} g^T d$ the support function of the nonempty set $C \subset \mathbb{R}^m$
$F_C(d) := \mathrm{Argmax}_{c \in C} d^T c$ the face of the nonempty set $C \subset \mathbb{R}^m$ exposed by $d \in \mathbb{R}^m$
$\partial f(x)$ the subdifferential of the finite-valued convex function f at $x \in \mathbb{R}^m$

$$\partial f(x) := \{s \in \mathbb{R}^m : f(y) - f(x) \geq s^T(y - x), \quad \text{for all } y \in \mathbb{R}^m\}$$

$f'(x; d)$ the directional derivative of a convex function f at $x \in \mathbb{R}^m$ in the direction $d \in \mathbb{R}^m$

$$f'(x; d) := \inf_{t>0} \frac{f(x + td) - f(x)}{t}$$

or equivalently (see [16, Sect. VI.1]), the support function of $\partial f(x)$

$$f'(x;\cdot) = \sigma_{\partial f(x)}(\cdot) \tag{2}$$

$\partial_\varepsilon f(x)$ the ε-subdifferential of f at $x \in \mathbb{R}^m$:

$$\partial_\varepsilon f(x) := \{s \in \mathbb{R}^m : f(y) - f(x) \geq \langle s, y - x\rangle - \varepsilon \text{ for all } y \in \mathbb{R}^m\}$$

$f'_\varepsilon(x; d)$, the ε-directional derivative of f at $x \in \mathbb{R}$ in the direction $d \in \mathbb{R}$, is the support function of $\partial_\varepsilon f(x)$:

$$f'_\varepsilon(x;\cdot) := \sigma_{\partial_\varepsilon f(x)}(\cdot)$$

$\mathcal{S}_n$ space of $n \times n$ symmetric matrices
$\mathcal{S}_n^+$ cone of positive semidefinite matrices
$X \succ Y$ (resp. $X \succeq Y$) means that the matrix $X - Y \in \mathcal{S}_n$ is positive definite (resp. positive semidefinite)
$\operatorname{tr} X := \sum_{i=1}^n X_{ii}$ trace of the matrix $X \in \mathcal{S}_n$
$\langle X, Y\rangle := \operatorname{tr} XY$ Frobenius scalar product of $X, Y \in \mathcal{S}_n$
$\|X\| := \sqrt{\langle X, X\rangle}$ Frobenius norm of $X \in \mathcal{S}_n$
$X^\dagger$ Moore-Penrose inverse of X: if $X = \sum_{i=1}^n \lambda_i(X) q_i q_i^T$ is the spectral decomposition of X, $X^\dagger$ can be defined as $X^\dagger := \sum_{\lambda_i(X)\neq 0} \frac{1}{\lambda_i(X)} q_i q_i^T$
$\lambda_1(X) \geq \ldots \geq \lambda_n(X)$ eigenvalues of $X \in \mathcal{S}_n$ in decreasing order
$E_1(X)$ first eigenspace of $X \in \mathcal{S}_n$, *i.e.,* the eigenspace associated with $\lambda_1(X)$
$\mathcal{A}^* : \mathcal{S}_n \to \mathbb{R}^m$ is the adjoint operator of the linear operator $\mathcal{A} : \mathbb{R}^m \to \mathcal{S}_n$ defined by: for all $(d, D) \in \mathbb{R}^m \times \mathcal{S}_n$

$$\langle D, \mathcal{A}d\rangle = d^T \mathcal{A}^* D$$

When $\mathcal{A}d = \sum_{j=1}^m d_j A_j$, with $A_j \in \mathcal{S}_n$ for $j = 1, \ldots, m$, we have for all $D \in \mathcal{S}_n$, $\mathcal{A}^* D = [\langle A_1, D\rangle, \ldots, \langle A_m, D\rangle]^T$

2. First-order analysis

In this section we recall elementary results for the maximum eigenvalue function: the subdifferential of λ_1 can be characterized as an exposed face of a compact section of the cone of semidefinite matrices. Then we propose an enlargement of the subdifferential of λ_1 based on the computation of *approximate eigenvectors* and a *vertical development* of λ_1, *i.e.,* a development of the function $\varepsilon \mapsto f'_\varepsilon(x; d)$. We derive similar results for $f := \lambda_1 \circ A$ using a simple chain rule. This will lead us to the main result of this section: any direction d separating 0 from the chosen enlargement of $\partial f(x)$ is a "good" descent direction.

2.1. Subdifferentials and faces

In this paragraph we give explicit descriptions of the subdifferential and the approximate subdifferential of λ_1. In this analysis, a convex compact set plays a paramount role: the intersection of the cone of semidefinite matrices with the hyperplane $\{V \in \mathcal{S}_n : \operatorname{tr} V = 1\}$,

$$\mathcal{C}_n := \{V \in \mathcal{S}_n : V \succeq 0,\ \operatorname{tr} V = 1\}\,. \tag{3}$$

By analogy with the *simplex of* $\mathbb{R}^n$ in linear programming, we will refer to $\mathcal{C}_n$ as the *spectraplex* of $\mathcal{S}_n$. The following result is well-known; the proof is easy to derive via the spectral decomposition of symmetric matrices.

Lemma 1. *The spectraplex $\mathcal{C}_n$ is the convex hull of rank-one matrices:*

$$\mathcal{C}_n = \operatorname{co}\{qq^T : \|q\| = 1\}\,.$$

□

Using *Rayleigh's variational formulation*

$$\lambda_1(X) = \max_{q\in\mathbb{R}^m, \|q\|=1} q^T X q\,,$$

together with Lemma 1, a *support function formulation* is obtained (see Sect. 1.2 for a definition of a support function):

$$\lambda_1(X) = \sigma_{\mathcal{C}_n}(X)\,. \tag{4}$$

As in [15], we will favor the support function formulation since it will be our main tool in the analysis of Sect. 2.3. In order to describe the exposed faces of the spectraplex $\mathcal{C}_n$ (Theorem 1), we first establish the following technical result.

Lemma 2. *Let X and Z be in $\mathcal{S}_n$; let $Q = [q_1, \ldots, q_r]$ be an $n \times r$ matrix whose columns form an orthonormal basis of* $\ker X$.

(i) Then, $XZ = 0$ if and only if there exists $Y \in \mathcal{S}_r$ such that $Z = QYQ^T$.
(ii) Assume in addition $X,\ Z \in \mathcal{S}_n^+$; then $\langle X, Z\rangle = 0$ if and only if there exists $Y \succeq 0$ such that $Z = QYQ^T$.

Proof ((i)). We have $XZ = 0$ if and only if range $Z \subset \ker X = \operatorname{span}\{q_1, \ldots, q_r\}$. This is equivalent to saying that Z belongs to the subspace

$$\operatorname{span}\left\{q_i q_j^T + q_j q_i^T : i, j = 1, \ldots, r\right\} = Q\mathcal{S}_r Q^T\,.$$

[*(ii)*] When X and Z positive semidefinite, a consequence of the Schur product theorem [17, Sect. 5.2] is that $\operatorname{tr} XZ = 0$ if and only if $XZ = 0$. Then, via (i), Z has the form QYQ^T; it is positive semidefinite if and only if $Y \succeq 0$.

□

The following theorem recalls previously known geometrical descriptions of $\partial\lambda_1(X)$ (see [10] or [34]) and, along the lines of [15], makes an explicit link with the exposed faces of $\mathcal{C}_n$.

Theorem 1. *(i) Let $X \in \mathcal{S}_n$ and let Q_1 be an $n \times r$ matrix whose columns form an orthonormal basis of $E_1(X)$. The face of the spectraplex $\mathcal{C}_n$ exposed by X is*

$$F_{\mathcal{C}_n}(X) = \left\{ Q_1 Y Q_1^T : Y \in \mathcal{C}_r \right\} = \mathrm{co}\{qq^T : \|q\| = 1, q \in E_1(X)\}\,; \tag{5}$$

it is the face exposed by any $X' \in \mathcal{S}_n$ such that $E_1(X') = E_1(X)$.
(ii) The subdifferential of λ_1 at X is an exposed face of the spectraplex:

$$\partial\lambda_1(X) = F_{\mathcal{C}_n}(X)\,. \tag{6}$$

Proof ((i)). Realize first that

$$\mathrm{co}\{qq^T : \|q\| = 1, q \in E_1(X)\} = \left\{ Q_1 Y Q_1^T : Y \in \mathcal{C}_r \right\}. \tag{7}$$

Indeed write any normalized vector of $E_1(X)$ under the form $q = Q_1 z$, with $z \in \mathbb{R}^r$ and $\|z\| = 1$. We get

$$\begin{aligned} \mathrm{co}\{qq^T : \|q\| = 1, q \in E_1(X)\} &= \mathrm{co}\left\{ Q_1 z z^T Q_1^T : z \in \mathbb{R}^r, \|z\| = 1 \right\} \\ &= Q_1\, \mathrm{co}\left\{ zz^T : z \in \mathbb{R}^r, \|z\| = 1 \right\} Q_1^T\,, \end{aligned}$$

where, in view of Lemma 1, $\mathrm{co}\{zz^T : z \in \mathbb{R}^r, \|z\| = 1\} = \mathcal{C}_r$. Now, by definition of an exposed face, $Z \in F_{\mathcal{C}_n}(X)$ means

$$Z \in \mathcal{C}_n \ \text{ and } \ \langle X, Z\rangle = \sigma_{\mathcal{C}_n}(X) = \lambda_1(X)\,,$$

or equivalently,

$$Z \in \mathcal{C}_n \ \text{ and } \ \langle \lambda_1(X) I_n - X, Z\rangle = 0\,.$$

Altogether, (5) is obtained with Lemma 2 and $F_{\mathcal{C}_n}(X') = F_{\mathcal{C}_n}(X)$ if and only if $\ker(\lambda_1(X) I_n - X) = \ker(\lambda_1(X') I_n - X')$, *i.e.*, $E_1(X) = E_1(X')$.

[*(ii)*] It is well known that the subdifferential of a support function $\sigma_{\mathcal{C}_n}$ at a point X is the exposed face of $\mathcal{C}_n$ exposed by X [16, Example VI.3.1].

□

The description of the approximate subdifferential is also obtained directly from the support function formulation of λ_1.

Theorem 2. *For all $\varepsilon \geq 0$, we have*

$$\partial_\varepsilon \lambda_1(X) = \{Z \in \mathcal{C}_n : \langle Z, X\rangle \geq \lambda_1(X) - \varepsilon\} \tag{8}$$

Proof. The approximate subdifferential of a support function is given in [16, Example XI.1.2.5]. Then (8) follows immediately. In [44] the same result is obtained via an analysis of the conjugate function of λ_1.

□

The directional derivative of λ_1 has an easy expression.

Theorem 3. *For all $D \in \mathcal{S}_n$, we have*

$$\lambda_1'(X; D) = \lambda_1\left(Q_1^T D Q_1\right).$$

Proof. Use (2) and (6) to obtain

$$\begin{aligned}\lambda_1'(X;\,D) &= \sigma_{\partial\lambda_1(X)}(D) = \max_{Y\in\mathcal{C}_r}\left\langle D,\, Q_1 Y Q_1^T\right\rangle \\ &= \max_{Y\in\mathcal{C}_r}\left\langle Q_1^T D Q_1,\, Y\right\rangle = \lambda_1\big(Q_1^T D Q_1\big)\,.\end{aligned}$$

This completes the proof. This result was already established using perturbation theory in [22] and can also be found in [15], [41].

□

It is well-known in nonsmooth optimization that the descent property $\lambda_1'(X;\,D) < 0$ of a direction D is unstable, the function $X \mapsto \lambda_1'(X;\,D)$ being discontinuous. Minimization algorithms based on this mere property are usually not convergent, because descent along such directions may be numerically insufficient. One is much more interested in ε-descent directions, for which

$$\lambda_{1,\varepsilon}'(X;\,D) := \sigma_{\partial_\varepsilon\lambda_1(X)}(D)$$

is negative. Said otherwise, these directions separate 0 from $\partial_\varepsilon\lambda_1(X)$. The continuity of $X \mapsto \lambda_{1,\varepsilon}'(X;\,D)$ [16, Theorem XI.4.1.3] guarantees the numerical efficiency of such directions. Yet the difficulty here is to get a separation algorithm, and this is the rationale for dual bundle methods [16, XIII,XIV]. This paper follows the same approach; but, instead of separating 0 from $\partial_\varepsilon\lambda_1(X)$, we use the structure of our specific problem to provide a tractable "good approximation" of the latter set.

2.2. Enlargement of the subdifferential

Since λ_1 is the support function of a convex compact set, it can be seen as an infinite-max function. Then a first idea could be to consider the enlargement proposed in [8, Chap. VI]: the convex hull of the gradients of ε-active functions. Here the functions are linear and it is easy to see, via (8), that the obtained enlargement is exactly the ε-subdifferential of λ_1. For practical reasons (see Remark 1), we will work here with a smaller set $\delta_\varepsilon\lambda_1(X)$ which satisfies

$$\partial\lambda_1(X) \subset \delta_\varepsilon\lambda_1(X) \subset \partial_\varepsilon\lambda_1(X)\,.$$

Definition 1. *For $X \in \mathcal{S}_n$ and $\varepsilon \geq 0$ we define*

- *the set of indices of ε-largest eigenvalues*

$$I_\varepsilon(X) := \{i \in \{1,\ldots,n\} : \lambda_i(X) > \lambda_1(X) - \varepsilon\}\,, \tag{9}$$

- *the ε-multiplicity of $\lambda_1(X)$: $r_\varepsilon := \max\{i : i \in I_\varepsilon(X)\}$,*
- *the ε-first eigenspace: $E_\varepsilon(X) := \oplus_{i\in I_\varepsilon(X)} E_i(X)$, where $E_i(X)$ is the eigenspace of X associated with the ith eigenvalue $\lambda_i(X)$,*
- *its orthogonal complement: $F_\varepsilon(X) := \oplus_{i\notin I_\varepsilon(X)} E_i(X)$,*
- *the "spectral separation" of ε: $\Delta_\varepsilon(X) := \lambda_{r_\varepsilon}(X) - \lambda_{r_\varepsilon+1}(X)$.*

□

Pseudospectrum. The notions of approximate eigenvalues can be connected with the recent theory of *pseudospectra of linear operators* [43,42]: this notion is mainly used to cope with the lack of regularity of nonsymmetric matrices; here, for symmetric matrices, a pseudospectrum can also be useful. Indeed it enables us to recover more than first-order regularity of the largest eigenvalue (precisely some local regularity of the set of approximate subgradients). For a normal matrix (in particular for a symmetric matrix) the ε-pseudospectrum comprises the union of the closed balls of radius ε about each eigenvalue. In fact we consider here one of these ε-balls, the one centered at $\lambda_1(X)$, and we take its intersection with the spectrum of X. The important role played by the approximate eigenvalues justifies the wording *approximate eigenvalues method*.

□

Take now an $n \times r_\varepsilon$ matrix Q_ε whose columns form an orthonormal basis of $E_\varepsilon(X)$. Then we define the following compact convex set:

$$\delta_\varepsilon\lambda_1(X) := \mathrm{co}\{ee^T : \|e\| = 1, e \in E_\varepsilon(X)\}\,. \tag{10}$$

or equivalently, via Theorem 1,

$$\delta_\varepsilon\lambda_1(X) = \left\{Q_\varepsilon Y Q_\varepsilon^T : Y \in \mathcal{C}_{r_\varepsilon}\right\} = F_{\mathcal{C}_n}\left(Q_\varepsilon Q_\varepsilon^T\right) = \partial\lambda_1\left(Q_\varepsilon Q_\varepsilon^T\right). \tag{11}$$

This set is an outer-approximation of $\partial\lambda_1(X)$ and an inner-approximation of $\partial_\varepsilon\lambda_1(X)$:

Proposition 1. *Let $X \in \mathcal{S}_n$. Then for all $\varepsilon \geq 0$, we have*

$$\partial\lambda_1(X) \subset \delta_\varepsilon\lambda_1(X) \subset \partial_\varepsilon\lambda_1(X)\,. \tag{12}$$

Proof. The inclusion $\partial\lambda_1(X) \subset \delta_\varepsilon\lambda_1(X)$ derives directly from (5) and (10). Another easy inclusion is $\delta_\varepsilon\lambda_1(X) \subset \mathcal{C}_n$. Take now $Z \in \delta_\varepsilon\lambda_1(X)$: $Z = Q_\varepsilon Y Q_\varepsilon^T$ with $Y \in \mathcal{C}_{r_\varepsilon}$ implies $\langle Z, X\rangle = \langle Y, Q_\varepsilon^T X Q_\varepsilon\rangle \geq \lambda_{r_\varepsilon}$ since

$$Q_\varepsilon^T X Q_\varepsilon = \mathrm{diag}\,(\lambda_1(X), \ldots, \lambda_{r_\varepsilon}(X)) \;\succeq \lambda_{r_\varepsilon} I_{r_\varepsilon}$$

and $\mathrm{tr}\, Y = 1$. Together with (9), we obtain $\langle Z, X\rangle \geq \lambda_1(X) - \varepsilon$; since $Z \in \mathcal{C}_n$ this means, according to (8), that $Z \in \partial_\varepsilon\lambda_1(X)$.

□

A crucial point consists now in quantifying the (Hausdorff) distance between our enlargement and the approximate subdifferential. One way to proceed is to get a vertical development of λ_1.

2.3. Vertical development

The notion of vertical development is presented in [33]. It will result in Theorem 4, which gives an explicit upper bound for the distance between the approximate subdifferential and the enlargement $\delta_\varepsilon\lambda_1(X)$. We start with some linear algebra.

Lemma 3. *Let $U \in \mathbb{R}^{n\times n}$ be such that $U^TU = I_n$. Then, there exist $n \times n$ matrices $(E_\varepsilon, F_\varepsilon, \Sigma, T)$ such that*

$$\begin{cases} \text{the columns of } E_\varepsilon \text{ are unit vectors of } E_\varepsilon(X) & (a) \\ \text{the columns of } F_\varepsilon \text{ are unit vectors of } F_\varepsilon(X) & (b) \\ \Sigma \text{ and } T \text{ are diagonal and positive semidefinite} & (c) \\ \Sigma^2 + T^2 = I & (d) \\ U = E_\varepsilon\Sigma + F_\varepsilon T & (e) \end{cases} \tag{13}$$

Proof. In [33, Lemma 5.2], we give a constructive proof of this result for $\varepsilon = 0$. There is no difficulty to see that the same proof can be applied here by decomposing each column of U on $E_\varepsilon(X) \oplus F_\varepsilon(X) = \mathbb{R}^n$.

□

The decomposition $\mathbb{R}^m = E_\varepsilon(X) \oplus F_\varepsilon(X)$ provides us with some useful relations.

Lemma 4. *Let $(E_\varepsilon, F_\varepsilon, \Sigma, T)$ be a quadruplet satisfying* (13) *and $\Theta = \text{diag}\,(\theta_1, \ldots, \theta_n) \in \mathcal{C}_n$. Then we have,*

$$\begin{cases} E_\varepsilon^T X F_\varepsilon = F_\varepsilon^T X E_\varepsilon = 0 & (a) \\ \lambda_{r_\varepsilon}(X) \le \left\langle X, \left(E_\varepsilon \Theta E_\varepsilon^T\right)\right\rangle \le \lambda_1(X) & (b) \\ \left\langle X, F_\varepsilon \Theta F_\varepsilon^T\right\rangle \le \lambda_{r_\varepsilon+1}(X) & (c) \\ \text{tr}\,(\Sigma\Theta T) \le [\text{tr}\,(T\Theta T)]^{1/2} & (d) \end{cases} \tag{14}$$

Proof.
[a]: The subspaces $E_\varepsilon(X)$ and $F_\varepsilon(X)$ are orthogonal and invariant by X. Then the columns of XF_ε are in $F_\varepsilon(X)$ and they are orthogonal to the columns of E_ε. This implies (a).

[b]: Since $E_\varepsilon\Theta E_\varepsilon^T = \sum_{i=1}^p \theta_i e_i e_i^T$, where the e_i's are the columns of E_ε, we have

$$\left\langle X, E_\varepsilon\Theta E_\varepsilon^T\right\rangle = \sum_{i=1}^p \theta_i e_i^T X e_i\,.$$

Now use the fact that for all unit vectors $e \in E_\varepsilon(X)$, $\lambda_{r_\varepsilon}(X) \le e^T X e \le \lambda_1(X)$, together with $\Theta = \text{diag}\,(\theta_1, \ldots, \theta_n) \in \mathcal{C}_n$, to get ($b$).

[c]: Similarly to [b], we have

$$\left\langle X, F_\varepsilon\Theta F_\varepsilon^T\right\rangle = \sum_{i=1}^p \theta_i\, f_i^T X\, f_i \le \lambda_{r_\varepsilon+1}(X)\Big(\sum_{i=1}^n \theta_i\Big) = \lambda_{r_\varepsilon+1}(X)\,.$$

[d]: Note that $\Sigma \preceq I_n$ and ΘT is (diagonal) positive semidefinite to get

$$\text{tr}\,(\Sigma\Theta T) \le \text{tr}\,(\Theta T)\,. \tag{15}$$

Now use $\Theta = \mathrm{diag}\,(\theta_1, \ldots, \theta_n) \in \mathcal{C}_n$ and $T = \mathrm{diag}\,(t_1, \ldots, t_n) \succeq 0$, together with the concavity of the square-root function, to obtain

$$\sum_{i=1}^{n} \theta_i t_i \leq \Big(\sum_{i=1}^{n} \theta_i t_i^2\Big)^{1/2}.$$

In matrix notation, this means $\mathrm{tr}\,(\Theta T) \leq [\mathrm{tr}\,(T\Theta T)]^{1/2}$. Together with (15), this gives (d).

□

We can now give a key result toward relating the two sets of (8) and (11).

Proposition 2. *Let $X \in \mathcal{S}_n$, $\varepsilon \geq 0$ and $\eta \geq 0$. For all $Z \in \partial_\eta \lambda_1(X)$, there exists $G_\varepsilon \in \delta_\varepsilon \lambda_1(X)$ and five $n \times n$ matrices $(E_\varepsilon, F_\varepsilon, \Sigma, T, \Theta)$ such that*

$$\begin{cases} (E_\varepsilon, F_\varepsilon, \Sigma, T) \text{ satisfies } (13)\,(a, b, c, d) & (a) \\ \Theta = \mathrm{diag}\,(\theta_1, \ldots, \theta_n) \in \mathcal{C}_n & (b) \\ Z = G_\varepsilon + \big(E_\varepsilon \Sigma \Theta T F_\varepsilon^T + F_\varepsilon \Sigma \Theta T E_\varepsilon^T\big) + \big(F_\varepsilon T \Theta T F_\varepsilon^T - E_\varepsilon T \Theta T E_\varepsilon^T\big) & (c) \\ \mathrm{tr}\, T\Theta T \leq \frac{\eta}{\lambda_{r_\varepsilon}(X) - \lambda_{r_\varepsilon+1}(X)} = \frac{\eta}{\Delta_\varepsilon \lambda_1(X)} & (d) \end{cases} \tag{16}$$

Proof. Write the spectral decomposition of $Z \in \partial_\eta \lambda_1(X)$: there exists $U \in \mathbb{R}^{n \times n}$, such that $U^T U = I_n$ and a diagonal matrix Θ such that $Z = U \Theta U^T$. In view of (8), we have $\Theta \in \mathcal{C}_n$. Then, apply Lemma 3: $U = E_\varepsilon \Sigma + F_\varepsilon T$ where $(E_\varepsilon, F_\varepsilon, \Sigma, T)$ satisfies $(13)_{(a,b,c,d)}$. Substituting this in the spectral decomposition of Z, we obtain

$$Z = G_\varepsilon + \big(E_\varepsilon \Sigma \Theta T F_\varepsilon^T + F_\varepsilon T \Theta \Sigma E_\varepsilon^T\big) + \big(F_\varepsilon T \Theta T F_\varepsilon^T - E_\varepsilon T \Theta T E_\varepsilon^T\big), \tag{17}$$

where $G_\varepsilon := E_\varepsilon \Theta E_\varepsilon^T = \sum_{i=1}^{p} \theta_i e_i e_i^T$ and the e_i's are unit vectors of $E_\varepsilon(X)$. According to (11), this means $G_\varepsilon \in \delta_\varepsilon \lambda_1(X)$. Then (a, b, c) are satisfied. In order to prove (d), take the scalar product of X with the left- and right-hand side of (17) and use $(14)_{(a,b,c,d)}$ to obtain

$$\lambda_1(X) + (\lambda_{r_\varepsilon+1}(X) - \lambda_{r_\varepsilon}(X))\mathrm{tr}\,(T\Theta T) \geq \langle X, Z \rangle.$$

Since $Z \in \partial_\eta \lambda_1(X)$, we have together with (8), $\langle X, Z \rangle \geq \lambda_1(X) - \eta$. This enables us to complete the proof.

□

The following result says that $\delta_\varepsilon \lambda_1(X)$ is a good approximation of $\partial_\eta \lambda_1(X)$ for η small enough, depending on the spectral separation of Definition 1.

Theorem 4. *For all $\varepsilon \geq 0$, $\eta \geq 0$ and $D \in \mathcal{S}_n$, we have*

$$\lambda'_{1,\eta}(A;\ D) \leq \sigma_{\delta_\varepsilon \lambda_1(X)}(D) + \rho(\eta, \varepsilon)\,\|D\|, \tag{18}$$

or equivalently,

$$\partial_\eta \lambda_1(X) \subset \delta_\varepsilon \lambda_1(X) + B(0, \rho(\eta, \varepsilon)), \tag{19}$$

where $\rho(\eta, \varepsilon) := (\frac{2\eta}{\Delta_\varepsilon(X)})^{1/2} + (\frac{2\eta}{\Delta_\varepsilon(X)})$.

Proof. Let $\varepsilon \geq 0$, $\eta \geq 0$, $D \in \mathcal{S}_n$ and $Z \in \partial_\eta \lambda_1(X)$. From $(16)_{(c)}$, we obtain

$$\langle Z, D\rangle = \langle G_\varepsilon, D\rangle + \left\langle \Sigma\Theta T, E_\varepsilon^T DF_\varepsilon + F_\varepsilon^T DE_\varepsilon \right\rangle + \left\langle T\Theta T, F_\varepsilon^T DF_\varepsilon - E_\varepsilon^T DE_\varepsilon \right\rangle .$$

Let us bound each of the three terms from above. First, $G_\varepsilon \in \delta_\varepsilon \lambda_1(X)$ implies

$$\langle G_\varepsilon, D\rangle \leq \sigma_{\delta_\varepsilon \lambda_1(X)}(D) . \tag{20}$$

Then, denoting $(\Sigma\Theta T) = \operatorname{diag}(\sigma_1\theta_1 t_1, \ldots, \sigma_n\theta_n t_n)$, we have

$$\left\langle \Sigma\Theta T, E_\varepsilon^T DF_\varepsilon + F_\varepsilon^T DE_\varepsilon \right\rangle = \sum_{i=1}^n \sigma_i\theta_i t_i \left[\left\langle D, e_i f_i^T + f_i e_i^T \right\rangle \right] .$$

Now, use the Cauchy-Schwarz inequality, together with $\|e_i f_i^T + f_i e_i^T\| = \sqrt{2}$, to get

$$\left\langle \Sigma\Theta T, E_\varepsilon^T DF_\varepsilon + F_\varepsilon^T DE_\varepsilon \right\rangle \leq \sqrt{2}\|D\| \operatorname{tr}(\Sigma\Theta T) . \tag{21}$$

Similarly we have for the last term

$$\begin{aligned} \left\langle T\Theta T, F_\varepsilon^T DF_\varepsilon - E_\varepsilon^T DE_\varepsilon \right\rangle &= \textstyle\sum_{i=1}^n \theta_i t_i^2 \left[\left\langle D, f_i f_i^T - e_i e_i^T \right\rangle \right] \\ &\leq \operatorname{tr}(T\Theta T) \|D\| \left(\|f_i f_i^T\| + \|e_i e_i^T\| \right) \\ &\leq 2\operatorname{tr}(T\Theta T) \|D\| , \end{aligned} \tag{22}$$

since $\|f_i f_i^T\| = \|e_i e_i^T\| = 1$. Putting together (20), (21), (14)-(d) and (22), we get

$$\langle Z, D\rangle \leq \sigma_{\delta_\varepsilon \lambda_1(X)}(D) + \sqrt{2}\|D\| \left(\operatorname{tr}(T\Theta T)\right)^{1/2} + 2\|D\| \operatorname{tr}(T\Theta T) . \tag{23}$$

Together with $(16)_{(d)}$, we obtain (18); (19) is the geometrical form of (18).

□

We will use this result in a simplified form.

Corollary 1. *Let $X \in \mathcal{S}_n$, $\varepsilon \geq 0$ and $\eta \in [0, \frac{\Delta_\varepsilon(X)}{2}]$. Then for all $D \in \mathcal{S}_n$, we have*

$$\lambda'_{1,\eta}(X;\ D) \leq \sigma_{\delta_\varepsilon \lambda_1(X)}(D) + \left(\frac{8\eta}{\Delta_\varepsilon(X)} \right)^{1/2} \|D\| . \tag{24}$$

Proof. Since $\eta \leq \frac{\Delta_\varepsilon(X)}{2}$, we have $0 \leq \frac{2\eta}{\Delta_\varepsilon(X)} \leq (\frac{2\eta}{\Delta_\varepsilon(X)})^{1/2}$ which, together with (18), gives (24).

□

Finally, we show we have s simple expression for the support function of $\delta_\varepsilon \lambda_1(X)$: it is the largest eigenvalue of an $r_\varepsilon \times r_\varepsilon$ symmetric matrix.

Proposition 3. *Let $X \in \mathcal{S}_n$. Then for all $\varepsilon \geq 0$ and $D \in \mathcal{S}_n$, we have*

$$\sigma_{\delta_\varepsilon \lambda_1(X)}(D) = \lambda_1(Q_\varepsilon^T D Q_\varepsilon) . \tag{25}$$

Proof. The set $\delta_\varepsilon \lambda_1(X)$ has the same structure as $\partial\lambda_1(X)$; the proof is then similar to that of Theorem 3.

□

In the following paragraph we extend these results to the composite function $f := \lambda_1 \circ A$.

2.4. Composition with an affine mapping

We recall that $\mathcal{A}$ denotes the linear part of $A(\cdot)$ and $\mathcal{A}^*$ its adjoint. First-order composition is based on the following elementary chain rule.

Proposition 4. *For $x \in \mathbb{R}^m$ and $\varepsilon \geq 0$,*

$$\partial_\varepsilon f(x) = \mathcal{A}^* \partial_\varepsilon \lambda_1(A(x)) \,. \tag{26}$$

Proof. This is a straightforward application of the chain rule given in [16, Theorem XI.3.2.1].

□

Then an enlargement of $\partial f(x)$ is the following convex set:

$$\delta_\varepsilon f(x) := \mathcal{A}^* \delta_\varepsilon \lambda_1(A(x)) \,, \tag{27}$$

and its associated support function

$$\tilde{f}'_\varepsilon(x;\, d) := \sigma_{\delta_\varepsilon f(x)}(d) \,. \tag{28}$$

Here we use the notation $\tilde{f}'_\varepsilon(x;\, d)$ to emphasize the analogy with the approximate directional derivative of f at x:

$$f'_\varepsilon(x;\, d) := \sigma_{\partial_\varepsilon f(x)}(d) \,.$$

Applying the linear mapping $\mathcal{A}^*$ in (12), it comes

$$\partial f(x) \subset \delta_\varepsilon f(x) \subset \partial_\varepsilon f(x) \,, \quad \text{[geometric form]} \tag{29}$$

or equivalently

$$f'(x; d) \leq \tilde{f}'_\varepsilon(x;\, d) \leq f'_\varepsilon(x; d) \,, \quad \text{[analytic form]}$$

The quality of this approximation is derived from inequality (24): for all $\varepsilon \geq 0$, $\eta \in [0, \Delta_\varepsilon(A(x))]$ and $d \in \mathbb{R}^m$,

$$f'_\eta(x;\, d) \leq \tilde{f}'_\varepsilon(x;\, d) + \left(\frac{8\eta}{\Delta_\varepsilon(A(x))}\right)^{1/2} \|\mathcal{A}\, d\| \leq \tilde{f}'_\varepsilon(x;\, d) + \left(\frac{8\eta}{\Delta_\varepsilon(A(x))}\right)^{1/2} \kappa \|d\| \,, \tag{30}$$

where $\kappa := \sup_{\|x\|=1} \|\mathcal{A}(x)\|$ is the largest singular value of $\mathcal{A}$.

Furthermore, it is straightforward from (25) and (27) that $\tilde{f}'_\varepsilon(x;\, d)$ is also a maximum eigenvalue:

$$\tilde{f}'_\varepsilon(x;\, d) = \lambda_1\big(Q_\varepsilon^T (\mathcal{A}\, d) Q_\varepsilon\big) \,. \tag{31}$$

In particular for $\varepsilon = 0$, we have

$$\tilde{f}'_0(x;\, d) = f'(x; d) = \lambda_1\big(Q_1 (\mathcal{A}\, d) Q_1^T\big) \,. \tag{32}$$

Ideally we would like to choose $\varepsilon > 0$ and find a so-called direction d of ε-descent.

Lemma 5 ([16, Lemma XIII.1.2.3]). *Suppose that $d \in \mathbb{R}^m$ is a direction of ε-descent: $f'_\varepsilon(x; d) < 0$. Then the set of all $t > 0$ such that $f(x + t\,d) < f(x) - \varepsilon$ forms a non empty interval.*

Remark 1. The difficulty here is that $\partial_\varepsilon f(x)$ is so rich that computing its support function $f'_\varepsilon(x; d)$ for a given direction d or *a fortiori* looking for the best ε-descent direction $\mathrm{Argmin}_{\|d\|=1} f'_\varepsilon(x; d)$ seems to be as expensive as the original problem (P). Therefore instead of working with $\partial_\varepsilon f(x)$ we deal with its inner approximation $\delta_\varepsilon f(x)$. Equality (31) *quantifies* the effort needed to evaluate $\tilde{f}'_\varepsilon(x;\, d)$: it requires the computation of the largest eigenvalue of an $r_\varepsilon \times r_\varepsilon$ matrix. We can now specify to what extent a direction $d \in \mathbb{R}^m$ satisfying at $x \in \mathbb{R}^m$

$$\tilde{f}'_\varepsilon(x;\, d) < 0\,, \tag{33}$$

is a "good" descent direction. The following result is an improvement of [32, Theorem 5.5] where the distance between the exact subdifferential of λ_1 at $X \in \mathcal{S}_n$ and its ε-subdifferential was considered.

□

Theorem 5. *Let $x \in \mathbb{R}^m$, $\varepsilon \geq 0$ and $d \in \mathbb{R}^m$ be such that* (33) *holds. Then*

(i) d is a direction of $\eta(x, \varepsilon)$-descent, where

$$\eta(x, \varepsilon) := \left[\frac{\tilde{f}'_\varepsilon(x;\, d)}{4\kappa\|d\|}\right]^2 \Delta_\varepsilon(A(x))\,.$$

(ii) In addition, assume there exist $\omega \in [0, 1]$, $\delta > 0$ and $\mu > 0$ such that

$$\tilde{f}'_\varepsilon(x;\, d) \leq -\omega\|d\|^2\,, \tag{34}$$

$\|d\| \geq \delta$ and $\Delta_\varepsilon(A(x)) \geq \mu$; then

$$\eta(x, \varepsilon) \geq \left[\frac{\omega\delta}{4\kappa}\right]^2 \mu\,. \tag{35}$$

Proof. It is straightforward from (30).

□

Remark 2. Let $g \in \delta_\varepsilon f(x)$ be such that $d = -g$ satisfies (34) for $\omega = 1$; then $g = \mathrm{proj}_{\delta_\varepsilon f(x)} 0$. When $\omega \in]0, 1]$, g is an approximation of this projection. In Sect. 3.1, we present a separation algorithm to obtain such directions.

□

3. First-order algorithm

We describe here an iterative process to compute η-descent directions using the information stored in $\delta_\varepsilon f(x)$, the parameters η and ε being related as in Theorem 5. Then the step-length is determined with a (finite) dichotomous line-search for η-descent. The approximate eigenvalue algorithm and its convergence analysis complete the section.

3.1. Projection problem

The problem we want to solve is

$$\min_{g \in \delta_\varepsilon f(x)} \|g\|^2 . \tag{36}$$

This is a quadratic optimization problem over the cone of positive semidefinite matrices. Indeed, in view of (27) and (11), program (36) is equivalent to

$$\begin{cases} \min \left\| \mathcal{A}^*\left(Q_\varepsilon Y Q_\varepsilon^T\right) \right\|^2 \\ Y \succeq 0 \\ \operatorname{tr} Y = 1 . \end{cases} \tag{37}$$

In [35], a similar projection problem is encountered. The authors adopt the following approach: instead of projecting onto $\delta_\varepsilon f(x)$, they project onto aff $\delta_\varepsilon f(x)$, *i.e.*, the constraint $Y \succeq 0$ is replaced by $Y \in \mathcal{S}_{r_\varepsilon}$. This leads them to a quadratic problem with linear equality constraints which can be solved with classical (and efficient) techniques. Yet this approach has a major drawback: when the minimizer is not positive semidefinite, we have no interpretation for the resulting projection; one has to escape from the current iteration and to change the multiplicity r_ε.

Here we compute an approximation g of $\operatorname{proj}_{\delta_\varepsilon f(x)} 0$: we require

$$\tilde{f}_\varepsilon'(x; -g) \leq -\omega \|g\|^2 ,$$

where $\omega \in]0, 1]$ is a tolerance that controls the proximity of g from the projection $\operatorname{proj}_{\delta_\varepsilon f(x)} 0$ (see Remark 2). The algorithm we present here is essentially the one proposed by J. Cullum, W.E. Donath and P. Wolfe in [6]. It is an instance of a general scheme, called the *support-black box method* in [16, Sect. IX.3] to minimize a quadratic form on a convex set. This algorithm is in fact a separation algorithm that was first given in [12].

SUPPORT-BLACK BOX METHOD
Step 0. Set $l = 1$ and $s = s_1 \in \delta_\varepsilon f(x)$.
Step 1. Compute $d_l = -\operatorname{proj}_{P_l} 0$, where $P_l = \operatorname{co}\{s_1, \ldots, s_l\}$.
Step 2. Compute $s_{l+1} \in \delta_\varepsilon f(x)$ such that

$$s_{l+1}^T d_l = \sigma_{\delta_\varepsilon f(x)}(d_l) .$$

Step 3. (Stopping criterion). If $\sigma_{\delta_\varepsilon f(x)}(d_l) \leq -\omega \|d_l\|^2$ then STOP.
Replace l by $l + 1$ and return to Step 1.

□

Note that s_{l+1} in Step 2 has the form $\mathcal{A}^* uu^T$ where u is a unit eigenvector associated with $\lambda_1(Q_\varepsilon^T(\mathcal{A} d_l) Q_\varepsilon)$. During the separation process a *bundle* of ε-subgradients $\{s_1, \ldots, s_k\}$ is generated; in that sense the first order method presented in this section is a *bundle method*.

The convergence of the support-black box method is investigated in detail in [16, Sect. IX.3]. In particular, we have the following properties.

Proposition 5. *The support-black box method at $x \in \mathbb{R}^m$ with $\omega \in]0,1[$ converges in a finite number of steps. Assume that the ε-Strict Complementarity condition holds at $x \in \mathbb{R}^m$,*

$$(SC)_\varepsilon \qquad \mathrm{proj}_{\delta_\varepsilon f(x)}\, 0 \in \mathrm{ri}\, \delta_\varepsilon f(x)\,.$$

Then finite convergence is obtained also if $\omega = 1$.

Proof. Set $g_\varepsilon := \mathrm{proj}_{\delta_\varepsilon f(x)}\, 0$. Then the results are directly derived from Theorem IX.3.3.3 and Proposition IX.3.3.4 in [16].

□

Remark 3. The complexity of the support-black box method is not explicitly known. It seems that the number of steps will depend on the condition number of the linear operator

$$\mathcal{S}_{r_\varepsilon} \ni Y \mapsto \mathcal{K}_\varepsilon(Y) := \mathcal{A}^*(Q_\varepsilon Y Q_\varepsilon^T)\,. \tag{38}$$

Requiring that $\mathcal{K}_\varepsilon$ is not singular was already a condition introduced in [35] and was the first appearance of the notion of *transversality* in semidefinite programming although it was not named as such. The connection is clearly established in [41].

□

3.2. Line-search

Let $x \in \mathbb{R}^m$ and $d \in \mathbb{R}^m$ be produced by the support-black box method of Sect. 3.1, so that $\tilde{f}'_\varepsilon(x;\, d) < 0$. Then, according to Theorem 5, the objective function can be decreased by a positive number $\eta(x, \varepsilon)$.

The problem we consider is now: find $t > 0$ such that

$$f(x + td) \leq f(x) - \eta(x, \varepsilon)\,. \tag{39}$$

When assumptions of Theorem 5 (ii) hold, a simple line-search for η-descent can be presented. We expose here a line-search based on a dichotomous scheme controlled by a η-descent stopping criterion. For an advanced implementation we refer to [16, Remark XIII.2.1.2].

LINE-SEARCH
Step 0. Set $t_L = 0$, $t_R = +\infty$ and $t_0 = 1$.
Step 1 (work). Obtain $q(t) := f(x + td)$ and $q'_+(t) := f'(x + td, d)$ using (32).
Step 2 (η-descent test). If (39) holds stop.
Step 3 (Dichotomous search). If $q'_+(t) > 0$ set $t_R := t$; else set $t_L = t$. Compute $t = \frac{t_L + t_R}{2}$.

□

Theorem 6. *Let $x \in \mathbb{R}^m$, $\varepsilon > 0$ and $d \in \mathbb{R}^m$ be satisfying the assumptions of Theorem 5 (ii). Then the line-search for $\eta(x, \varepsilon)$-descent stops after a finite number of steps.*

Proof. From Theorem 5 (ii), the quantity $\eta(x,\varepsilon)$ is strictly positive. Use now Theorem 5 (i), together with Lemma 5: $\{t :\in \mathbb{R}_+ : f(x+td) \leq f(x) - \eta(x,\varepsilon)\}$ is a non empty interval. Therefore the dichotomous scheme will detect one of its elements after a finite number of iterations.

□

3.3. The approximate eigenvalues algorithm

To get a simple convergence analysis, the first-order algorithm we present here will use the following ε-strategy (see Fig. 1).

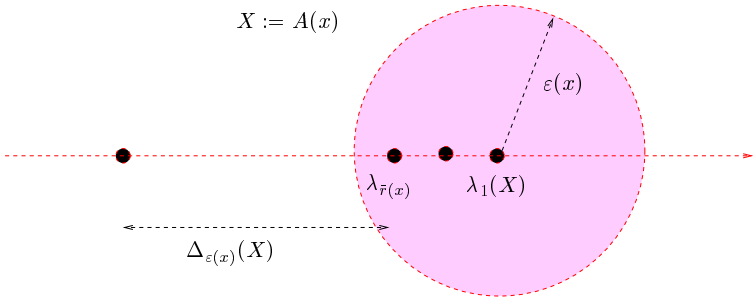

Fig. 1. ε-strategy

Choose first a tolerance $\bar{\varepsilon} > 0$. Then define at $x \in \mathbb{R}^m$

$$\left\{\begin{array}{l} R_{\bar{\varepsilon}}(x) := \left\{r \in \{1,\ldots,n-1\} : \lambda_r(A(x)) - \lambda_{r+1}(A(x)) \geq \frac{\bar{\varepsilon}}{n}\right\}, \\ \\ \bar{r}(x) := \left\{\begin{array}{ll} \min\{r : r \in R_{\bar{\varepsilon}}(x)\} & \text{if } \ R_{\bar{\varepsilon}}(x) \neq \emptyset \\ n & \text{otherwise,} \end{array}\right. \\ \\ \varepsilon(x) := \left\{\begin{array}{ll} f(x) - \lambda_{\bar{r}(x)}(A(x)) + \frac{\bar{\varepsilon}}{2n} & \text{if } \ R_{\bar{\varepsilon}}(x) \neq \emptyset \\ \\ \bar{\varepsilon} & \text{otherwise.} \end{array}\right. \end{array}\right. \tag{40}$$

Here, $\bar{\varepsilon}$ is the final tolerance: Algorithm 1 below is aimed at minimizing f within a tolerance $\bar{\varepsilon}$. With reference to Definition 1, observe that $\bar{r}(x)$ is the $\varepsilon(x)$-multiplicity of $\lambda_1(A(x))$. If $R_{\bar{\varepsilon}}(x) \neq \emptyset$, the spectral separation of $\varepsilon(x)$ is larger than $\frac{\bar{\varepsilon}}{n}$. Moreover we have $\lambda_r(A(x)) - \lambda_{r+1}(A(x)) < \frac{\bar{\varepsilon}}{n}$, for $r = 1,\ldots,\bar{r}(x)-1$, when $\bar{r}(x) > 1$.

Algorithm 1.
Step 0 (Initialization). Choose the tolerances $\delta > 0$, $\bar{\varepsilon} > 0$, $\omega \in]0,1[$; initialize $x := x_0 \in \mathbb{R}^m$.
Step 1 (Separation). Set $\varepsilon := \varepsilon(x)$ and compute $d \in -\delta_\varepsilon f(x)$ satisfying (34) with the support-black box method.

Step 2 (Stopping criterion). If $\|d\| \leq \delta$ stop.
Step 3 (Line-search). Compute t such that

$$f(x + td) \leq f(x) - \eta(x, \varepsilon)\,,$$

with the Line-Search of Sect. 3.2.
Step 5 (Update). Replace x by $x + td$; return to Step 1.

□

To ensure that our problem (P) makes sense, we assume that f is bounded from below. We have:

Lemma 6. *The function f is bounded from below if and only if*

$$\begin{array}{ll} (i) & 0 \in \mathcal{A}^*(\mathcal{C}_n)\,, \\ \textit{or equivalently,} & \\ (ii) & (\text{range}\,\mathcal{A})^{\perp} \cap \mathcal{S}_n^+ \neq \{0\}\,. \end{array}$$

Proof. From convex duality [16, Chap. X], $-f^*(0) = \inf_{x\in\mathbb{R}^m} f(x)$, where f^* is the conjugate function of f; as shown in [44], we have

$$-f^*(0) = \sup\{\langle Z, A_0\rangle : Z \in \mathcal{C}_n \cap \ker \mathcal{A}^*\}\,.$$

Then f is bounded from below if and only if $\mathcal{C}_n \cap \ker A^* \neq \emptyset$, which is clearly equivalent to (i) and (ii). Note that (ii) can be directly obtained from [26].

□

This leads us to the following result.

Lemma 7. *For all $x \in \mathbb{R}^m$, the parameter $\varepsilon(x)$ of* (40) *is not greater than $\bar{\varepsilon}$. Therefore*

$$\delta_{\varepsilon(x)} f(x) \subset \partial_{\bar{\varepsilon}} f(x)\,. \tag{41}$$

As a result, if f is bounded from below, we have

$$\begin{array}{ll} (i) & \Delta_{\varepsilon(x)}(A(x)) \geq \frac{\bar{\varepsilon}}{n}\,, \\ (ii)\ \textit{or} & 0 \in \delta_{\varepsilon(x)} f(x)\,. \end{array} \tag{42}$$

Proof. Take $x \in \mathbb{R}^m$ and consider the two following cases.
$[R_{\bar{\varepsilon}}(x) \neq \emptyset]$: by construction of $\varepsilon(x)$ (40), (i) holds: $\Delta_{\varepsilon(x)}(A(x)) \geq \frac{\bar{\varepsilon}}{n}$. Furthermore

$$\varepsilon(x) = f(x) - \lambda_{\bar{r}(x)}(A(x)) + \frac{\bar{\varepsilon}}{2n} < \left(\bar{r}(x) - \frac{1}{2}\right)\frac{\bar{\varepsilon}}{n} \leq \bar{\varepsilon}\,;$$

together with (29) and the fact that the set-valued function $\varepsilon \mapsto \partial_\varepsilon f(x)$ is not decreasing, this gives (41).

$[R_{\bar{\varepsilon}} = \emptyset]$: we have $\bar{r}(x) = n$ and $\varepsilon(x) = \bar{\varepsilon}$. Hence (41) still holds. Then from (11) and (27), we have $\delta_{\varepsilon(x)} f(x) = \mathcal{A}^*(\mathcal{C}_n)$. On the other hand, we know from Lemma 6 that f bounded means $0 \in \mathcal{A}^*(\mathcal{C}_n)$. This implies (ii).

□

Remark 4. *In practice* we will use ε bigger than $\varepsilon(x)$: at each iteration k we choose ε_k in $[0,\ \max\{\varepsilon(x_k), \theta^k\varepsilon_0\}[$ such that $\Delta_{\varepsilon_k}(x_k)$ is maximized, where $\theta \in]\frac{1}{2}, 1[$ is a given parameter which forces the convergence of ε_k towards $\bar{\varepsilon}$.

We are now in a position to prove finite convergence of Algorithm 1 towards an approximate solution of (P).

Theorem 7. *Assume that f is bounded from below. Then Algorithm 1 stops after a finite number of iterations, yielding $\bar{x}$ satisfying the approximate minimality condition*

$$f(y) \geq f(\bar{x}) - \bar{\varepsilon} - \delta\|y - \bar{x}\| \quad \text{for all} \quad y \in \mathbb{R}^m.$$

Proof. While $\|d\| > \delta$, we have $0 \notin \delta_{\varepsilon(x)} f(x)$ and according to Lemma 7, $\Delta_{\varepsilon(x)}(A(x)) \geq \frac{\bar{\varepsilon}}{n}$. Together with (35), this gives us $f(x) - f(x+td) \geq \eta(x,\varepsilon) \geq [\frac{m\delta}{4\kappa}]^2 \frac{\bar{\varepsilon}}{n}$. Then Algorithm 1 must stop before N steps, where N is the first integer satisfying

$$f(x_0) - N\left[\frac{m\delta}{4\kappa}\right]^2 \frac{\bar{\varepsilon}}{n} \leq -f^*(0)\,.$$

□

4. Second-order analysis

As with the first-order analysis, a convenient approach consists in studying the second-order behavior of λ_1 and then deriving that for the composite function $f = \lambda_1 \circ A$ by chain rules. Yet, even though their formulation is simple, chain rules are not easy to obtain here; as explained in [33], we need to introduce a geometrical condition to get them.

4.1. The $\mathcal{U}$-Lagrangian of λ_1

For a presentation of the $\mathcal{U}$-Lagrangian theory in a more general framework we refer to [24]. The second-order analysis we present here starts with the following idea: consider at $X \in \mathcal{S}_n$, the largest subspace where $\lambda_1'(X;\cdot)$ is linear.

Definition 2. *At $X \in \mathcal{S}_n$, we define*

$$\mathcal{U}(X) := \left\{U \in \mathcal{S}_n : \lambda_1'(X;U) + \lambda_1'(X;-U) = 0\right\},$$

and $\mathcal{V}(X) := \mathcal{U}(X)^\perp$.

□

The subspaces $\mathcal{U}(X)$ and $\mathcal{V}(X)$ are also characterized as follows.

Proposition 6 ([24]). *Let $X \in \mathcal{S}_n$.*

(i) For any $G \in \operatorname{ri}\partial\lambda_1(X)$, $\mathcal{U}(X)$ and $\mathcal{V}(X)$ are respectively the normal and tangent cones to $\partial\lambda_1(X)$ at G.

(ii) $\mathcal{U}(X)$ *and* $\mathcal{V}(X)$ *are respectively the subspaces orthogonal and parallel to* aff $\partial\lambda_1(X)$.

□

A first attempt to reach second order could consist in introducing the function induced by λ_1 in $\mathcal{U}(X)$. Yet, proceeding this way we would miss a major fact: a good model of λ_1 must consider the local behavior of all "active constraints" at X. At this stage geometry can help: it suggests fixing the multiplicity (*i.e.*, the "activity") of λ_1; this point of view is the one adopted in [36]. The "surface of activity" is defined as

$$\mathcal{M}_r := \{M \in \mathcal{S}_n : \lambda_1(M) = \ldots = \lambda_r(M) > \lambda_{r+1}(M)\}\,.$$

Reference [3] is usually proposed to prove the smoothness of $\mathcal{M}_r$. In fact, it can be obtained as a simple consequence of the *Constant Rank Theorem* [40, Chap. III Sect. 9]. This gives us a geometrical interpretation of the subspaces $\mathcal{U}(X)$ and $\mathcal{V}(X)$.

Theorem 8 ([33, Corollary 4.8]). *Let* $X \in \mathcal{M}_r$. *The subspaces* $\mathcal{U}(X)$ *and* $\mathcal{V}(X)$ *are respectively the tangent and normal spaces to* $\mathcal{M}_r$ *at* X*:*

$$\begin{aligned} &\mathcal{U}(X) = \left\{U \in \mathcal{S}_n : Q_1^T U Q_1 - \tfrac{1}{r}\big(\operatorname{tr} Q_1^T U Q_1\big) I_r = 0\right\}, \\ &\text{and } \ \mathcal{V}(X) = \left\{Q_1 Y Q_1^T : Y \in \mathcal{S}_r\,,\ \langle Y, I_r\rangle = 0\right\}. \end{aligned} \tag{43}$$

□

The $\mathcal{U}$-Lagrangian is a convex function which identifies locally the "ridge" of λ_1.

Definition 3. *Let* $X \in \mathcal{S}_n$ *and* $G \in \partial\lambda_1(X)$. *The* $\mathcal{U}$*-Lagrangian of* λ_1 *at the primal-dual pair* (X, G) *is the function*

$$\mathcal{U}(X) \ni U \mapsto L_{\mathcal{U}}(X, G; U) := \min_{V \in \mathcal{V}(X)} \lambda_1(X + U + V) - \langle G, V\rangle\,.$$

We define also the associated set of minimizers

$$V(X, G; U) = \operatorname{Argmin}_{V \in \mathcal{V}(X)} \lambda_1(X + U + V) - \langle G, V\rangle\,.$$

□

The following theorem is established in [24] for any finite valued convex function, we express it in our specific context.

Theorem 9. *Let* $(X, G) \in \mathcal{S}_n \times \partial\lambda_1(X)$. *Then, we have*

(i) the function $L_{\mathcal{U}}(X, G; \cdot)$ *is well-defined and convex over* $\mathcal{U}(X)$.

Assume, in addition, that $G \in \operatorname{ri} \partial\lambda_1(X)$. *Then,*

(ii) for all $U \in \mathcal{U}(X)$, $V(X, G; U)$ *is a nonempty compact convex set which satisfies*

$$\sup_{V \in V(X,G;U)} \|V\| = o(\|U\|)\,. \tag{44}$$

(iii) In particular, at $U = 0$, we have $V(U) = \{0\}$, $L_{\mathcal{U}}(X, G; 0) = \lambda_1(X)$ and $\nabla L_{\mathcal{U}}(X, G; 0) = \mathrm{proj}_{\mathcal{U}(X)} G$ exists.

□

A geometrical interpretation of (44) is that $\mathcal{U}(X)$ is tangent at X to the "ridge":

$$\{X + U + V(X, G; U) : U \in \mathcal{U}(X)\}\,.$$

In our context we can prove that this geometrical set coincides in a neighborhood of X with $\mathcal{M}_r$ when $G \in \mathrm{ri}\,\partial\lambda_1(X)$.

Theorem 10 ([33, Theorem 4.11]). *Assume $(X, G) \in \mathcal{S}_n \times \mathrm{ri}\,\partial\lambda_1(X)$. Then there exists $\delta > 0$ such that, for all $U \in \mathcal{U}(X) \cap B(0, \delta)$, $V(X, G; U)$ is a singleton. Moreover the map*

$$\mathcal{U}(X) \cap B(0, \delta) \ni U \mapsto X + U + V(U) \tag{45}$$

is a C^∞-parameterization of the sub-manifold $\mathcal{M}_r$.

□

This gives us a second-order development of λ_1 along $\mathcal{M}_r$.

Theorem 11 ([33, Theorem 4.12–Corollary 4.13]). *Take $(X, G) \in \mathcal{S}_n \times \mathrm{ri}\,\partial\lambda_1(X)$. Then $L_{\mathcal{U}}(X, G; \cdot)$ is C^∞ in a neighborhood of $U = 0$ and we have the following second-order development of λ_1*

$$\begin{aligned}\lambda_1(X + U + V(U)) = {} & \lambda_1(X) + \langle G, U + V(U)\rangle \\ & + \frac{1}{2}\langle \nabla^2 L_{\mathcal{U}}(X, G; 0)\cdot U, U\rangle + o(\|U\|^2)\,,\end{aligned} \tag{46}$$

where $\nabla^2 L_{\mathcal{U}}(X, G; 0)$ is known explicitly:

$$\nabla^2 L_{\mathcal{U}}(X, G; 0) = \mathrm{proj}^*_{\mathcal{U}(X)}\, H(X, G)\, \mathrm{proj}_{\mathcal{U}(X)}\,,$$

and $H(X, G)$ is the symmetric positive semidefinite operator

$$\mathcal{S}_n \ni Y \mapsto H(X, G)\, Y := GY[\lambda_1(X) I_n - X]^\dagger + [\lambda_1(X) I_n - X]^\dagger Y G\,.$$

□

The operator $\nabla^2 L_{\mathcal{U}}(X, G; 0)$ is called the $\mathcal{U}$-Hessian of λ_1 at (X, G).

4.2. Composition with affine operator

When composing with the affine operator $A(\cdot)$, we expect the same type of results as for λ_1 and we would like to have similar geometrical interpretations. It is obvious that the subspace where $f'(x; \cdot)$ is linear and its orthogonal complement can be written:

$$\mathcal{U}^f(x) = \mathcal{A}^{-1}(\mathcal{U}(A(x))) \quad \text{and} \quad \mathcal{V}^f(x) = \mathcal{A}^* \mathcal{V}(A(x))\,.$$

Yet, when concentrating on second-order, the first difficulty encountered is that the inverse image of $\mathcal{M}_r$, *i.e.,*

$$\{x \in \mathbb{R}^m : A(x) \in \mathcal{M}_r\} =: \mathcal{W}_r\,,$$

may be nonsmooth. Then to simplify the analysis, a transversality condition is relevant.

Definition 4. TRANSVERSALITY *We say that $A(\cdot)$ is transversal to $\mathcal{M}_r$ at $x \in \mathcal{W}_r$ if and only if*

$$(T) \qquad \mathcal{U}(A(x)) + \operatorname{range} \mathcal{A} = \mathcal{S}_n \,.$$

□

The transversality condition guarantees a one to one relation between $\partial f(x)$ and $\partial\lambda_1(A(x))$.

Lemma 8. *Assume that transversality holds at $x \in \mathbb{R}^m$ and take $g \in \partial f(x)$. Then there exists a unique $G \in \partial\lambda_1(X)$ such that $g = \mathcal{A}^*(G)$; if g is in $\operatorname{ri}\partial f(x)$, G is also in $\operatorname{ri}\partial\lambda_1(A(x))$. Moreover*

$$\begin{aligned} &\dim \mathcal{V}^f(x) = \tfrac{r(r+1)}{2} - 1\,, \\ &\textit{and} \quad \dim \mathcal{U}^f(x) = m + 1 - \tfrac{r(r+1)}{2}\,. \end{aligned} \tag{47}$$

Proof. Apply Theorem 5.5 in [33]: let Q_1 be a $n \times r$ matrix whose columns form an orthonormal basis of $E_1(A(x))$; then the mapping

$$\mathcal{V}(A(x)) \ni V \mapsto \mathcal{A}^*(V)$$

is nonsingular. Together with (43), this implies (47). The one-to-one relation between the subdifferentials is easily derived from the fact that $G - G' \in \mathcal{V}(A(x))$ when G and G' are in $\partial\lambda_1(A(x))$.

□

Finally transversality is a sufficient condition to obtain the differentiability of the $\mathcal{U}$-Lagrangian of f and to compute its Hessian via simple chain rules.

Theorem 12 ([33, Theorem 5.10]). *Assume that the transversality condition holds at $x \in \mathbb{R}^m$ and take $g \in \operatorname{ri}\partial f(x)$. Then the $\mathcal{U}$-Lagrangian of f at x,*

$$\mathcal{U}^f(x) \ni u \mapsto L^f_{\mathcal{U}}(x, g; u) := \min_{v \in \mathcal{V}^f(x)} f(x + u + v) - g^T v\,,$$

is C^∞ in a neighborhood of $u = 0$. In particular,

$$\nabla L^f_{\mathcal{U}}(x, g; 0) = \operatorname{proj}_{\mathcal{U}^f(x)} g\,,$$

and

$$\nabla^2 L^f_{\mathcal{U}}(x, g; 0) = \operatorname{proj}^*_{\mathcal{U}^f(x)} H^f(x, G) \operatorname{proj}_{\mathcal{U}^f(x)}\,,$$

where $H^f(x, G) := \mathcal{A}^*\, H(A(x), G)\, \mathcal{A}$ and G is the unique (via Lemma 8) vector of $\operatorname{ri}\partial\lambda_1(A(x))$ such that $g = \mathcal{A}^*(G)$.

□

The operator $\nabla^2 L^f_{\mathcal{U}}(x, g; 0)$ is called the $\mathcal{U}$-Hessian of f at (x, g). The projection operator $\mathrm{proj}_{\mathcal{U}^f(x)}$ is given by the expression [33, Corollary 5.7]:

$$\mathrm{proj}_{\mathcal{U}^f(x)} = I_m - \mathcal{K}_0(\mathcal{K}_0^* \mathcal{K}_0)^{-1} \mathcal{K}_0^* ,$$

where $\mathcal{K}_0$ is obtained by taking $\varepsilon = 0$ in (38).

The transversality condition enabled M. L. Overton and R.S. Womersley, in [35], to parameterize $\mathcal{M}_r$ using exponentials of matrices and to transform the original unconstrained minimization problem into a smooth constrained one. This led them essentially to the same second derivative formula. Going along the lines of [35] and [9], we can prove that $H(A(x), G)$ is the *second covariant derivative* in the Euclidean metric of the function

$$\hat{\lambda}_1(M) := \frac{1}{r} \sum_{i=1}^{r} \lambda_i(M) ,$$

which is smooth near $A(x) \in \mathcal{M}_r$ and coincides with λ_1 on $\mathcal{M}_r$.

Yet, note that the transversality condition is only sufficient to prove the smoothness of $L^f_{\mathcal{U}}(x, g; \cdot)$. The following example shows us that it is not necessary either to get the differentiability of $L^f_{\mathcal{U}}(x, g; \cdot)$ or to guarantee the smoothness of $\mathcal{W}_r$.

Example 1. Consider the mapping from $\mathbb{R}^2$ to $\mathcal{S}_3$ defined by

$$A(x_1, x_2) := \begin{bmatrix} x_1 - x_2 & 0 & 0 \\ 0 & x_2 - x_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

We have $f(x) = \max\{|x_1 - x_2|, 1\}$ and at $\hat{x} := \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ the transversality condition does not hold. We have obviously $\partial f(\hat{x}) = \{\alpha \begin{bmatrix} -1 \\ 1 \end{bmatrix} : \alpha \in [0, 1]\}$. Then $\mathcal{V}^f(0, 1) = \mathbb{R} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is unidimensional, whereas the transversality condition would imply, via Lemma 8, $\dim \mathcal{V}^f(0, 1) = 2$. Yet, $\mathcal{W}_2$ is linear and therefore smooth in a neighborhood of $\hat{x}$:

$$\mathcal{W}_2 = \hat{x} + \mathcal{U}^f(x) = \hat{x} + \mathbb{R} \begin{bmatrix} 1 \\ 1 \end{bmatrix} .$$

Moreover, for all $g \in \partial f(\hat{x})$ and $u \in \mathcal{U}^f(0, 1)$, $v(x, g; u) = \{0\}$ and $L^f_{\mathcal{U}}(x, g; u) = 1$ is trivially twice differentiable.

□

5. Global second-order algorithm

We first explain how to "stabilize" the $\mathcal{U}$-objects. Then we present the three steps of the global $\mathcal{U}$-Newton algorithm: dual, vertical and tangent steps.

5.1. Enlargement of $\mathcal{U}$

From the enlargement of $\partial\lambda_1(X)$ introduced in Sect. 2.2, we derive easily an enlargement of $\mathcal{U}(X)$. We give here the ε-version of Definition 2; note that $\sigma_{\delta_\varepsilon\lambda_1(X)}(\cdot)$ plays the role of $\lambda_1'(X;\cdot) = \sigma_{\partial\lambda_1(X)}(\cdot)$.

Definition 5. *Let $\varepsilon \geq 0$ and $X \in \mathcal{S}_n$. We define $\mathcal{U}_\varepsilon(X)$ as the largest subspace where $\sigma_{\delta_\varepsilon\lambda_1(X)}$ is linear:*

$$\mathcal{U}_\varepsilon(X) := \{U \in \mathcal{S}_n : \sigma_{\delta_\varepsilon\lambda_1(X)}(U) + \sigma_{\delta_\varepsilon\lambda_1(X)}(-U) = 0\}\,.$$

and $\mathcal{V}_\varepsilon(X) := \mathcal{U}_\varepsilon(X)^\perp$.

Proposition 6 can also be extended in this context.

Proposition 7. *Let $\varepsilon \geq 0$ and $X \in \mathcal{S}_n$. The subspaces $\mathcal{U}_\varepsilon$ and $\mathcal{V}_\varepsilon$ are equivalently characterized by:*

(i) For any $G_\varepsilon \in \operatorname{ri}\delta_\varepsilon\lambda_1(X)$, $\mathcal{U}_\varepsilon(X)$ and $\mathcal{V}_\varepsilon(X)$ are respectively the normal and tangent cones to $\delta_\varepsilon\lambda_1(X)$ at G_ε.
(ii) $\mathcal{U}_\varepsilon(X)$ and $\mathcal{V}_\varepsilon(X)$ are the subspaces respectively orthogonal and parallel to $\operatorname{aff}\delta_\varepsilon\lambda_1(X)$.

Proposition 9 below will show that $\mathcal{U}_\varepsilon(X)$ is just the usual $\mathcal{U}$ at an appropriately shifted matrix X_ε. As a result, all the geometrical interpretations of Sect. 4.1 can be reproduced.

5.2. Dual step

The dual step is essentially done in Sect. 3.1: we compute an approximation of the projection of 0 onto $\delta_\varepsilon f(x)$ with the support black-box method and obtain $g_\varepsilon \in \delta_\varepsilon f(x)$ such that

$$\tilde{f}_\varepsilon'(x;-g_\varepsilon) \leq -\omega\|g_\varepsilon\|^2\,. \tag{48}$$

In fact the support-black box also produces a matrix G_ε in $\delta_\varepsilon\lambda_1(A(x))$ such that $g_\varepsilon = \mathcal{A}^*(G_\varepsilon)$. The subgradient g_ε will be used to guarantee an efficient descent and the dual variable G_ε will be needed to compute the Hessian of the $\mathcal{U}$-Lagrangian.

Actually, it is important for quadratic convergence to obtain the exact projection. The following result shows that this is possible.

Proposition 8. *Assume (T) of Definition 4 of and $(SC)_0$ of Proposition 5 hold at $x^* \in \mathbb{R}^m$. Then for $0 < \varepsilon < \Delta_0(A(x^*))$ (Definition 1), the ε-strict-complementarity condition $(SC)_\varepsilon$ holds in a whole neighborhood of x^*.*

Proof. When $0 < \varepsilon < \Delta_0(A(x^*))$, by continuity of eigenvalues, we have $r_\varepsilon(x) = \dim E_\varepsilon(A(x)) = r^*$ when $x \in B(x^*, \rho)$ for some $\rho > 0$. In fact, for ρ small enough, we have $E_\varepsilon(A(x)) = E_{tot}(A(x))$, where $E_{tot}(A(x))$ is the *total eigenspace* for the $\lambda_1(A(x^*))$-group at $A(x)$; this notion is defined in [18] and used in [41,33]. Then introducing, as in [33, Theorem 4.5], the C^∞ map $X \mapsto Q_{tot}(X)$ defined in a neighborhood of $A(x^*)$, the projection problem involved in $(SC)_\varepsilon$ can be written together with (11) and (27):

$$\min_{Y \in \mathcal{C}_{r^*}} \left\| \mathcal{A}^* \left(Q_{tot}(A(x)) Y Q_{tot}(A(x))^T \right) \right\|^2 . \tag{49}$$

Then the proof becomes similar to the one of [33, Proposition 6.9].

- The first-order optimality conditions of (49) with the transversality condition at x^*, enable us to use the Implicit Function Theorem and to get a C^∞ map $B(x^*, \delta) \ni x \mapsto Y(x)$ solution of (49).
- The strict complementarity condition at x^* tells us that $Y(x^*)$ is positive definite.
- The continuity of $Y(\cdot)$ implies that $Y(x)$ is positive definite for x in a neighborhood of x^*. This means that the ε-strict complementarity condition is satisfied in the latter neighborhood.

□

Corollary 2. *Assume that* (T) *and* $(SC)_0$ *hold at* x^* *and take* $0 < \varepsilon < \Delta_0(A(x^*))$. *Then there exists* $\rho_1 > 0$ *such that for all* $x \in B(x^*, \rho_1)$, *the support-black box method with* $m = 1$ *produces* $\mathrm{proj}_{\delta_\varepsilon f(x)}\, 0$ *in a finite number of steps.*

Proof. Combine Proposition 5 with Proposition 8.

□

5.3. Vertical step

It would be nice to project the current point x onto the manifold $\mathcal{W}_r := \{x \in \mathbb{R}^m : A(x) \in \mathcal{M}_r\}$. Yet, even in the case where transversality holds, such a projection is very hard to obtain. Nevertheless, in the space of matrices it is easy to compute a point of $\mathcal{M}_r$ which satisfies the first-order optimality conditions associated with the projection problem. Consider a spectral decomposition of $X \in \mathcal{M}_r$:

$$X = Q_\varepsilon \Lambda_\varepsilon Q_\varepsilon^T + R_\varepsilon \Sigma_\varepsilon R_\varepsilon^T ,$$

where Λ_ε and Σ_ε are respectively $r_\varepsilon \times r_\varepsilon$ and $(n - r_\varepsilon) \times (n - r_\varepsilon)$ diagonal matrices, and R_ε is a $n \times (n - r_\varepsilon)$ matrix whose columns form an orthonormal basis of $E_\varepsilon(X)^\perp$. Here the components of Λ_ε are greater than the components of Σ_ε. Then define for all $X \in \mathcal{S}_n$,

$$\hat{\lambda}_{1,\varepsilon}(X) := \frac{1}{r_\varepsilon} \sum_{i=1}^{r_\varepsilon} \lambda_i(X) ,$$

and

$$X_\varepsilon := \hat{\lambda}_{1,\varepsilon}(X) Q_\varepsilon Q_\varepsilon^T + R_\varepsilon \Sigma_\varepsilon R_\varepsilon^T .$$

We have the following result.

Theorem 13. *The matrix X_ε satisfies the first-order optimality conditions associated with the projection problem*

$$\min_{M\in\mathcal{M}_{r_\varepsilon}} \|M - X\|^2 . \tag{50}$$

Proof. Let M be a solution of (50). From Theorem 8, the manifold $\mathcal{M}_r$ is smooth in a neighborhood of M and its normal space at M is $\mathcal{V}(M)$. Then the (necessary) optimality conditions associated to (50) at M are

$$M \in \mathcal{M}_{r_\varepsilon} \ \text{ and } \ X - M \in \mathcal{V}(M) .$$

It is obvious that $X_\varepsilon \in \mathcal{M}_{r_\varepsilon}$, that

$$X - X_\varepsilon = Q_\varepsilon[\Lambda_\varepsilon - \hat{\lambda}_{1,\varepsilon}(X) I_{r_\varepsilon}] Q_\varepsilon^T ,$$

and that $\operatorname{tr}(\Lambda_\varepsilon - \hat{\lambda}_{1,\varepsilon}(X) I_{r_\varepsilon}) = 0$. This proves $X - X_\varepsilon \in \mathcal{V}(X_\varepsilon)$ and completes the proof.

□

We will also use the following property.

Proposition 9. *Let $\varepsilon \geq 0$ and $X \in \mathcal{S}_n$. We have*

$$\begin{gathered} \delta_\varepsilon\lambda_1(X) = \partial\lambda_1(X_\varepsilon) \\ \mathcal{U}_\varepsilon(X) = \mathcal{U}(X_\varepsilon) \ \text{ and } \ \mathcal{V}_\varepsilon(X) = \mathcal{V}(X_\varepsilon) . \end{gathered}$$

Proof. By construction, $\lambda_1(X_\varepsilon) = \hat{\lambda}_{1,\varepsilon}(X)$ and $E_1(X_\varepsilon) = E_\varepsilon(X)$. Then, together with (11) and (6), we have $\delta_\varepsilon\lambda_1(X) = \partial\lambda_1(X_\varepsilon)$. The rest is straightforward.

□

The step (in the space of matrices) from $A(x)$ to $A_\varepsilon(x) := [A(x)]_\varepsilon$ will be called a vertical step or $\mathcal{V}_\varepsilon$-step.

5.4. Tangent step

We assume that, at $x \in \mathbb{R}^m$, the dual and vertical steps have been previously computed: we have $g_\varepsilon(x) = \mathcal{A}^*(G_\varepsilon(x)) \in \delta_\varepsilon f(x)$ and $A_\varepsilon(x) \in \mathcal{M}_{r_\varepsilon}$. Then we define the following quadratic program,

$$\begin{cases} \min \langle G_\varepsilon(x), U\rangle + \frac{1}{2}\langle H(A_\varepsilon(x), G_\varepsilon(x))\, U, U\rangle \\ U \in \mathcal{U}(A_\varepsilon(x)) \\ A_\varepsilon(x) + U \in A_0 + \operatorname{range}(\mathcal{A}) , \end{cases} \tag{51}$$

where H is defined in Theorem 11. When $G_\varepsilon \in \operatorname{ri}\delta_\varepsilon\lambda_1(A(x))$, *i.e.,* (using Proposition 9) $G_\varepsilon \in \operatorname{ri}\partial\lambda_1(A_\varepsilon(x))$, (51) is equivalent to minimizing the second-order approximation of $L_{\mathcal{U}}(A_\varepsilon(x), G_\varepsilon; \cdot)$ subject to $A_\varepsilon(x) + U$ lying in the image of the affine mapping

$A(\cdot)$: the existence of a corresponding step in the space of variables is guaranteed. Then, program (51) takes the following form in the space of variables:

$$\begin{cases} \min\ g_\varepsilon(x)^T d + \frac{1}{2} d^T H_\varepsilon^f(x, G_\varepsilon(x))\, d \\ A(x) - A_\varepsilon(x) + \mathcal{A}\, d \in \mathcal{U}_\varepsilon(A(x)) \end{cases} \tag{52}$$

To solve (52), we assume nonemptiness of the feasible domain: there exists $d_0 \in \mathbb{R}^m$ such that

$$A(x) - A_\varepsilon(x) + \mathcal{A}\, d_0 \in \mathcal{U}_\varepsilon(A(x))\,. \tag{53}$$

When (53) is feasible, we set $u := d - d_0$. Then program (52) amounts to

$$\begin{cases} \min\ b_\varepsilon(x)^T u + \frac{1}{2} u^T H^f(x, G_\varepsilon(x)) u \\ u \in \mathcal{U}_\varepsilon^f(x)\,, \end{cases} \tag{54}$$

where H^f is defined in Theorem 12, $b_\varepsilon(x) := g_\varepsilon(x) + H^f(x, G_\varepsilon(x))\, d_0$, and $\mathcal{U}_\varepsilon^f(x) := \{u \in \mathbb{R}^m : \mathcal{A}\, u \in \mathcal{U}_\varepsilon(A(x))\}$. To guarantee that u is well-defined, we assume that $H^f(x, G_\varepsilon(x))$ is positive definite. Then, to ensure global convergence, we check whether the direction $d := d_0 + u$ satisfies

$$\tilde{f}_\varepsilon'(x;\ d) \le -\omega' \|d^2\|\,, \tag{55}$$

where ω' is a given number in $]0, \omega[$.

Remark 5. Even when applied to a smooth function, global convergence of the Newton algorithm is an open question when the Hessian has an unbounded condition number. This explains the need for an anti-zigzag mechanism, and we found that (55) is useful to prove convergence.

5.5. Global algorithm

The global $\mathcal{U}$-Newton algorithm is organized as follows.

Algorithm 2.
Step 0 (Initialization). Choose the tolerances $\delta > 0$, $\bar{\varepsilon} > 0$, $\omega \in]0, 1]$ and $\omega' \in]0, \omega[$; initialize $x := x_0 \in \mathbb{R}^m$ and set $\varepsilon := \varepsilon(x)$.
Step 1 (Dual step). Compute $g_\varepsilon(x) \in \delta_\varepsilon f(x)$ and $G_\varepsilon(A(x)) \in \delta_\varepsilon \lambda_1(A(x))$ satisfying (48) using the support-black box method.
Step 2 (stopping criterion). If $\|g_\varepsilon(x)\| \le \delta$ stop.
Step 3 (Vertical Step). Compute $A_\varepsilon(x)$.
Step 4 (Horizontal Step). If (53) is feasible and $H^f(x, G_\varepsilon(x))$ is positive definite, set d to the solution of (52). If (55) holds and $\|d\| > \delta$ go to Step 5.
If any of these conditions is not satisfied, set $d = -g_\varepsilon(x)$.
Step 5 (Line-search). Compute t such that

$$f(x + td) \le f(x) - \eta(x, \varepsilon)\,,$$

with the Line-Search of Sect. 3.2.
Step 6 (Update). Replace x by $x + td$ and ε by $\varepsilon(x + td)$; return to Step 1.

□

Theorem 14. *Assume that f is bounded from below. Then Algorithm 2 (with $\omega < 1$) stops after a finite number of iterations, yielding $\bar{x}$ satisfying the approximate minimality condition*

$$f(y) \geq f(\bar{x}) - \bar{\varepsilon} - \delta\|y - \bar{x}\| \quad \text{for all} \quad y \in \mathbb{R}^m . \tag{56}$$

Proof. The proof is as in Theorem 7, since (55) is guaranteed to hold at each iteration. □

In order to obtain quadratic convergence, we introduce a condition which can be seen as the generalization of the regularity assumption needed in all Newton-type methods.

Definition 6. *Let $x^* \in \mathbb{R}^m$ be a solution of (P). We say that the Strict Second-Order Condition (SSOC) holds at x^* if (T) of Definition 4 and $(SC)_0$ of Proposition 5 hold at x^* and the $\mathcal{U}$-Hessian of f at $(x^*, 0)$ is positive definite.*

We first give consequences of $(SSOC)$.

Proposition 10. *Assume that x^* is a solution of (P) and that $(SSOC)$ holds at x^*. Then*

(i) *x^* is the unique solution of (P),*

(ii) *for any $\rho > 0$ there exists $\alpha > 0$ such that*

$$f(x) \leq f(x^*) + \alpha \quad \Rightarrow \quad x \in B(x^*, \rho) ,$$

(iii) *for any $\rho > 0$ there are $\bar{\varepsilon}$ and δ small enough, such that Algorithm 2 yields at least one iterate in $B(x^*, \rho)$, and all the subsequent iterates remain in $B(x^*, \rho)$ as well.*

Proof ((i)). . Decompose an arbitrary $d \in \mathbb{R}^m$ as $d = u + v$ with $u \in \mathcal{U}^f(x^*)$ and $v \in \mathcal{V}^f(x^*)$. By definition of the $\mathcal{U}$-Lagrangian (see Theorem 12), $f(x^*+d) \geq L^f_{\mathcal{U}}(x^*, 0; u)$. Hence

$$f(x^* + d) \geq L^f_{\mathcal{U}}(x^*, 0; u) = f(x^*) + \frac{1}{2} u^T \nabla^2 L^f_{\mathcal{U}}(x^*, 0; 0)\, u + o(\|u\|^2) .$$

If d is small, u is small and clearly $f(x^* + d) > f(x^*)$.

[(ii)]. Then $\operatorname{argmin}_{x \in \mathbb{R}^m} f(x)$ is bounded; f has bounded level sets ([16, Proposition IV.3.2.5]): say $f(x) \leq f(x_0) + 1$ implies $\|x - x^*\| \leq M$. Assume for contradiction that there exist $\rho > 0$ and a sequence $\{x_k\}_{k \in \mathbb{N}}$ such that for $k \geq 1$,

$$f(x_k) \leq f(x^*) + \frac{1}{k} \leq f(x_0) + 1 \quad \text{and} \quad \|x_k - x^*\| > \rho .$$

Then x_k is bounded: $\|x_k - x^*\| \leq M$. Extract a subsequence converging to some $\hat{x}$ and pass to the limit (using continuity of f) to obtain the desired contradiction:

$$f(\hat{x}) \leq f(x^*) \quad \text{and} \quad \|\hat{x} - x^*\| > \rho ;$$

$\hat{x} \neq x^*$ is a minimizer of f, which contradicts (i).

[(iii)]. Observe that Algorithm 2 produces a decreasing sequence of f-values: every iterate satisfies $\|x - x^*\| \leq M$. Given $\rho > 0$, take $\alpha > 0$ as in (ii), set $\bar{\varepsilon}$ and δ such that $\bar{\varepsilon} + \delta M \leq \alpha$: from (56), at least the final iterate $\bar{x}$ satisfies $f(\bar{x}) \leq f(x^*) + \alpha$, hence $\bar{x} \in B(x^*, \rho)$; if this occurs before stopping, it occurs at each subsequent iteration. □

The following lemma will enable us to guarantee that, close enough to a solution x^*, the exact multiplicity r^* is identified.

Lemma 9. *Assume that* $0 < \bar{\varepsilon} < \Delta_0(A(x^*))$. *Then, there exists* $\rho_2 > 0$ *such that for all* $x \in B(x^*, \rho_2)$,

$$r_{\bar{\varepsilon}}(x) = \bar{r}(x) = r^*\,,$$

where $\bar{r}(x)$ *is defined in* (40) *and* $r_{\bar{\varepsilon}}(x) := \dim E_{\bar{\varepsilon}}(x)$.

Proof. When $0 < \bar{\varepsilon} < \Delta_0(A(x^*))$, it is clear in (40) that $r_{\bar{\varepsilon}}(x^*) = \bar{r}(x^*) = r^*$. Then, the result derives directly from continuity of all eigenvalues.

□

In the following theorem, we combine Corollary 2, Proposition 10 and Lemma 9 to show that after some iteration Algorithm 2 coincides with the $\mathcal{U}$-Newton algorithm presented in [33].

Theorem 15. *Assume* (P) *has a solution* $x^* \in \mathbb{R}^m$ *at which* $(SSOC)$ *holds. Let Algorithm 2 be applied with* $\omega = 1$ *and with* $\bar{\varepsilon} > 0$, $\delta > 0$ *and* $\|x_0 - x^*\|$ *all sufficiently small. Assume also that the solution of* (52) *is accepted by Step 4 and that Step 5 produces* $t = 1$. *Set* $x^+ = x + t\,d$. *Then, there exists* $C > 0$ *such that*

$$\|x^+ - x^*\| \leq C\|x - x^*\|^2\,. \tag{57}$$

Proof. Consider $\rho_1 > 0$ of Corollary 2 (with $\varepsilon = \bar{\varepsilon}$) and $\rho_2 > 0$ of Lemma 9. Suppose $x_0 \in B(x^*, \rho)$, with $0 < \rho < \min\{\rho_1, \rho_2\}$. By Proposition 10 (iii), all subsequent iterates are in $B(x^*, \rho)$. Then, from Lemma 9, $r_{\bar{\varepsilon}}(x) = \bar{r}(x) = r^*$ and from Corollary 2 the support-black box method with $m = 1$ converges in a finite number of steps. If in addition, (55) holds and the step length $t = 1$ is accepted in the line-search, Algorithm 2 coincides with the local second order algorithm described in [33]. This algorithm has a quadratic rate of convergence [33, Theorem 6.13]: for ρ small enough, there exists $C > 0$ such that (57) holds when $x \in B(x^*, \rho)$.

□

In the following section we connect this work with recent results on bundle methods for semidefinite programming.

6. From dual to proximal-type bundle methods

When presenting the first-order method in Sect. 3, our main objective was to provide an algorithm with a simple geometrical description to globalize second-order schemes "à la" Fletcher-Overton. Yet, for practical purposes, the following must be noted:

1. the ratio (use of the information)/(computational cost) in Algorithm 1 is quite low: subgradients and ε-subgradients computed during the global process (Step 1) or during the local line-search are used only once,

2. the algorithm is very sensitive to the choice of the ε-strategy; choosing at each iteration $\varepsilon = \varepsilon(x)$ (40) ensures global convergence but may be not the best policy in practice.

Therefore there are advantages to move from a *dual Markovian bundle method* to a *polyhedral-semidefinite proximal bundle method* [23]. The obtained algorithm, in its first-order version, is close to the one described by K.C. Kiwiel in [19] where the precision to compute the largest eigenvalue is also controlled by the global scheme.

Transporting old subgradients. A first step consists in using old $\bar{\varepsilon}$-subgradients as ε-subgradients at the current point for some $\varepsilon \geq \bar{\varepsilon}$. This can be expressed as follows.

Proposition 11. *Let $k > 1$. Assume that, at each point x_i, we have computed $g_i = \mathcal{A}^* G_i \in \delta_{\bar{\varepsilon}} f(x_i)$ with $G_i \in \mathcal{C}_n$ for $i = 1, \ldots, k-1$. Let Q_k be an $n \times r_k$ matrix whose columns form an orthonormal basis of $E_{\bar{\varepsilon}}(A(x_k))$. Set $\tilde{\alpha}_k := [\alpha_1, \ldots, \alpha_{k-1}]^T \in \mathbb{R}^{k-1}$, $\mathbf{1}_{k-1} \in \mathbb{R}^{k-1}$ the vector of all ones and for $\varepsilon \geq 0$ consider the set*

$$\mathcal{G}_{k,\varepsilon} := \Big\{ \textstyle\sum_{i=1}^{k-1} \alpha_i G_i + Q_k Y Q_k^T : \tilde{\alpha}_k \in \mathbb{R}^{k-1}_+ ,\ Y \in \mathcal{S}^+_{r_k} \\ \mathbf{1}_{k-1}^T \tilde{\alpha}_k + \langle I_{r_k}, Y \rangle = 1, \\ \big\langle \textstyle\sum_{i=1}^{k-1} \alpha_i G_i + Q_k Y Q_k^T, A(x_k) \big\rangle \geq f(x_k) - \varepsilon \Big\} .$$

Then for all $\varepsilon \geq \bar{\varepsilon}$, we have

$$\delta_{\bar{\varepsilon}} f(x_k) \subset \mathcal{A}^* \mathcal{G}_{k,\varepsilon} \subset \partial_\varepsilon f(x_k) . \tag{58}$$

Proof. This is a straightforward consequence of (8) and (26).

□

Inclusion (58) suggests that the convex set $\mathcal{A}^* \mathcal{G}_{k,\varepsilon}$ could be taken as our new approximation of $\partial_\varepsilon f(x_k)$ at x_k.

Using some duality. This leads us to a new projection problem for (36):

$$\min \|g\|^2 , \quad g \in \mathcal{A}^* \mathcal{G}_{k,\varepsilon} .$$

Setting $\mathcal{G}_k := \mathcal{G}_{k,+\infty}$ and denoting by $G(\tilde{\alpha}, Y) := \sum_{i=1}^{k-1} \alpha_i G_i + Q_k Y Q_k^T$ an element of $\mathcal{G}_k$, the projection problem amounts to

$$\begin{cases} \min \|\mathcal{A}^* G(\tilde{\alpha}, Y)\|^2 , \quad G(\tilde{\alpha}, Y) \in \mathcal{G}_k \\ f(x_k) - \big\langle \sum_{i=1}^{k-1} \alpha_i G_i + Q_k Y Q_k^T, A(x_k) \big\rangle - \varepsilon \leq 0 . \end{cases} \tag{59}$$

Penalizing the linear inequality constraint by introducing a multiplier $2\mu_k > 0$, we obtain the new problem

$$\begin{cases} \min \|\mathcal{A}^* G(\tilde{\alpha}, Y)\|^2 + 2\mu_k \big(f(x_k) - (\mathcal{A}^* G(\tilde{\alpha}, Y))^T x_k - \langle A_0, G(\tilde{\alpha}, Y) \rangle - \varepsilon \big) \\ G(\tilde{\alpha}, Y) \in \mathcal{G}_k . \end{cases} \tag{60}$$

Now, multiplying the objective function by $1/(2\mu_k)$ and dropping the constant terms, it is proved in [13, Theorem 11.3.1] that (60) is exactly the dual (in the sense of Lagrangian duality [16, Chap. XII]) of the primal problem

$$\min_{y\in\mathbb{R}^m} \varphi_k(y) + \frac{\mu_k}{2}\|y - x_k\|^2 ,$$

where φ_k is the mixed polyhedral-semidefinite model of f at x_k :

$$\varphi_k(y) := \max_{G\in\mathcal{G}_k} \langle A(x_k), G\rangle .$$

Recalling the variational formulation of λ_1 and observing that $\mathcal{G}_k \supset \delta_\varepsilon\lambda_1(A(x_k))$ contains at least one matrix of the form qq^T where q is a unit vector in $E_1(A(x_k))$, it is easy to verify that this model satisfies

$$\begin{aligned} \varphi_k(y) &\le f(y) \quad \text{for all } y \in \mathbb{R}^m , \\ \text{and } \varphi_k(y_k) &\ge f(y_k) - \varepsilon . \end{aligned}$$

In this approach the control parameter is μ_k which is updated with a simple strategy [20].

A contribution for large-scale problems. The resulting first-order polyhedral-semidefinite proximal bundle method seems to be promising to solve large-scale eigenvalue problems: in [23], we present applications from control theory which involve 1000×1000 matrices depending on a large number of decision variables ($m = 500000$).

7. Dropping strict complementarity: an example from combinatorial optimization

In this paragraph we present an example from combinatorial optimization, suggested by Yu. Nesterov [29], to illustrate a qualitative distinction between interior-point methods and second-order bundle methods: for the latter methods, superlinear convergence can be observed even when sctrict complementarity does not hold. This fact seems to be corroborated by the recent theoretical work of A. Forsgren [11].

Let us consider the Boolean quadratic maximization problem

$$(BQM) \quad \begin{cases} \max \ x^T Q x \\ x \in \{-1, 1\}^n , \end{cases}$$

where $Q = -cc^T$ and $c = [n-1, -1, \ldots, -1]$. A primal semidefinite relaxation [28] for (BQM) is

$$(SDP) \quad \begin{cases} \max \ \langle Q, X\rangle , \quad X \in \mathcal{S}_n \\ d(X) = \mathbf{1}_n, \ X \succeq 0 , \end{cases}$$

where $d(X)$ is the diagonal of X. The dual of this problem is then

$$(SDP)^* \quad \begin{cases} \min \ \mathbf{1}_n^T u , \quad u \in \mathbb{R}^n \\ D(u) - Q \succeq 0 , \end{cases}$$

where $D(u)$ is the matrix with $u_1, \dots, u_n$ on its diagonal. It is easy to check that the three problems (BQM), (SDP) and $(SDP)^*$ have the same zero value. The solutions of (BQM) and $(SDP)^*$ are unique: $x^* = \mathbf{1}_n$ and $u^* = 0$; $X^* = \mathbf{1}_n\mathbf{1}_n^T$ is a solution of (SDP). Strict complementarity for (SDP) [1] does not hold: the matrix $D(u^*) - Q + X^* = cc^T + \mathbf{1}_n\mathbf{1}_n^T$ is not positive definite. Consider now the equivalent eigenvalue formulation of $(SDP)^*$ [7,14]: making the change of variable $u = \alpha\mathbf{1}_n + v$ with $v^T\mathbf{1}_n = 0$, we obtain

$$\begin{cases} \min n\,\alpha\,, \quad (\alpha, v) \in \mathbb{R}^{n+1} \\ \alpha I_n \succeq Q - D(v) \\ v^T\mathbf{1}_n = 0\,, \end{cases}$$

which is in turn equivalent to the unconstrained eigenvalue problem

$$(EVP) \qquad \min_{\tilde{v}\in\mathbb{R}^{n-1}} f(\tilde{v})\,,$$

with $\mathbb{R}^{n-1} \ni \tilde{v} = (v_1, \dots, v_{n-1}) \mapsto f(\tilde{v}) = n\,\lambda_1\big(Q - D(v_1, \dots, v_{n-1}, -\sum_{i=1}^{n-1} v_i)\big)$. We know that the notion of strict complementarity $(SC)_0$ in Proposition 5 of an eigenvalue problem coincides with the corresponding one in semidefinite programing [33, Remark 6.6]. We can verify here by hand that $\tilde{v}^* = 0$ and $0 \notin \mathrm{ri}\,\partial f(0)$: $0 = \mathcal{A}^*(\frac{\mathbf{1}_n\mathbf{1}_n^T}{n})$ where $\mathcal{A}^*$ is the linear operator

$$\mathcal{S}_n \ni X \mapsto \mathcal{A}^*X = -n \begin{bmatrix} X_{11} - X_{nn} \\ \vdots \\ X_{(n-1)(n-1)} - X_{nn} \end{bmatrix} \in \mathbb{R}^{n-1}\,,$$

and $\frac{\mathbf{1}_n\mathbf{1}_n^T}{n}$ is in the boundary of $\partial\lambda_1(Q)$. We run Algorithm 2 for $n = 10$ and $\tilde{v}_0 = \mathbf{1}_{n-1}$ with $\delta = 10^{-6}$, $\bar{\varepsilon} = 10^{-4}$, $\omega = 0.1$ and $\omega' = 0.001$. Figure 2 can be interpreted as follows: At the initial point the multiplicity is 1 and the distance to the second eigenvalue is large ($\Delta_{\bar{\varepsilon}}(\tilde{v}_0) = 98$); therefore the first-line search is very efficient (recall Theorem 5).

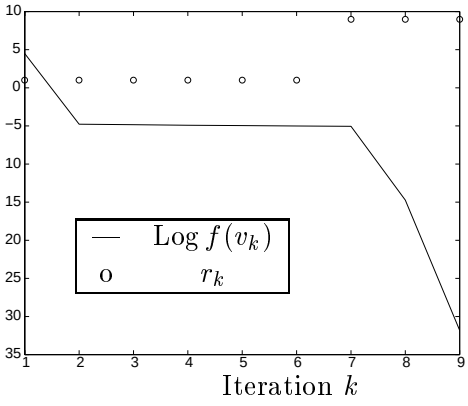

Fig. 2. Superlinear convergence

Yet the descent provokes a clustering of the 9 first eigenvalues: $\Delta_{\bar{\varepsilon}}(\tilde{v}_k)$ becomes small but not enough to increase $r_k := \bar{r}(\tilde{v}_k)$ of (40) until iteration 7 where they are counted as $\frac{\bar{\varepsilon}}{n}$-eigenvalues. Then second-order steps start to be efficient and superlinear convergence is observed.

8. Conclusion

In this paper, we have shown how to use a second-order theory, the $\mathcal{U}$-Lagrangian theory, to speed up the convergence of a first-order scheme to minimize the maximum eigenvalue function. The introduction of second-order information in a Markovian dual bundle method (the approximate eigenvalue method of [6,37]) enabled us to obtain the quadratic convergence of the resulting second-order bundle method when some regularity conditions hold. We also have made a connection with a new generation of bundle methods for semidefinite programming.

Acknowledgements. I wish to thank Claude Lemaréchal for the numerous and fruitful discussions we had together. I am also greatful to M. Overton for his careful reading and his numerous suggestions to improve the paper. Finally I thank an anonymous referee for his detailed and constructive remarks.

References

1. Alizadeh, F., Haeberly, J.-P.A., Overton, M.L. (1997): Complementarity and nondegeneracy in semidefinite programming. Math. Program. **77**, 111–128
2. Alizadeh, F., Haeberly, J.-P.A., Overton, M.L. (1998): Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results. SIAM J. Optim. **8**, 746–768
3. Arnold, V.I. (1971): On matrices depending on parameters. Russ. Math. Surveys **26**, 29–43
4. Bellman, R., Fan, K. (1963): On systems of linear inequalities in Hermitian matrix variables. In: Klee, V.L., ed., Convexity, volume 7 of Proceedings of Symposia in Pure Mathematics, pp. 1–11. American Mathematical Society
5. Boyd, S., El Ghaoui, L., Feron, E., Balakrishnan, V. (1994): Linear Matrix Inequalities in System and Control Theory, volume 15 of Studies in Applied Mathematics. SIAM, Philadelphia, PA, June 1994
6. Cullum, J., Donath, W.E., Wolfe, P. (1975): The minimization of certain nondifferentiable sums of eigenvalues of symmetric matrices. Math. Program. Study **3**, 35–55
7. Delorme, C., Poljak, S. (1993): Laplacian eigenvalues and the maximum cut problem. Math. Program. **62**, 557–574
8. Demjanov, V.F., Malozenov, V.N. (1974): Introduction to Minimax. Wiley & Sons
9. Edelman, A., Arias, T., Smith, S.T. (1998): The geometry of algorithms with orthogonality constraints. SIAM J. Matrix. Anal. Appl. **20**, 303–353
10. Fletcher, R. (1985): Semi-definite matrix constraints in optimization. SIAM J. Control Optim. **23**, 493–523
11. Forsgren, A. (2000): Optimality conditions for nonconvex semidefinite programming. Math. Program. **88**, 105–128
12. Gilbert, E.G. (1966): An iterative procedure for computing the minimum of a quadratic form on a convex set. SIAM J. Control **4**, 61–80
13. Helmberg, C., Oustry, F. (2000): Bundle methods to minimize the maximum eigenvalue function. In: Lieven, Vandenberghe, Saigal, R., Wolkovicz, H., eds., Handbook on Semidefinite Programming. Theory, Algorithms and Applications, volume 27 of International Series in Operations Research and Management Science. Kluwer Academic Publisher, March 2000
14. Helmberg, C., Rendl, F. (2000): A spectral bundle method for semidefinite programming. SIAM J. Optim. **10**, 673–696
15. Hiriart-Urruty, J.-B., Ye, D. (1995): Sensivity analysis of all eigenvalues of a symmetric matrix. Numer. Math. **70**, 45–72

16. Hiriart-Urruty, J.B., Lemaréchal, C. (1993): Convex Analysis and Minimization Algorithms. Springer-Verlag. Two volumes
17. Horn, R.A., Johnson, C.R. (1991): Topics in Matrix Analysis. Cambridge University Press
18. Kato, T. (1980): Perturbation Theory for Linear Operators. Springer-Verlag, New York
19. Kiwiel, K.C. (1986): A linearization algorithm for optimizing control systems subject to singular value inequalities. IEEE Trans. Autom. Control **AC-31**, 595–602
20. Kiwiel, K.C. (1990): Proximity control in bundle methods for convex nondifferentiable minimization. Math. Program. **46**, 105–122
21. Kojima, M., Shida, M., Shindoh, S. (1998): Local convergence of predictor-corrector infeasible-interior-point algorithms for sdps and sdlcps. Math. Program. **80**, 129–161
22. Lancaster, P. (1964): On eigenvalues of matrices dependent on a parameter. Numer. Math. **6**, 377–387
23. Lemaréchal, C., Oustry, F. (1999): Nonsmooth algorithms to solve semidefinite programs. In: El Ghaoui, L., Niculescu, S.-I., eds., Advances in LMI methods in Control, Advances in Design and Control series. SIAM
24. Lemaréchal, C., Oustry, F., Sagastizábal, C. (2000): The $\mathcal{U}$-Lagrangian of a convex function. Transact. Amer. Math. Soc. **352**
25. Lewis, A.S., Overton, M.L. (1996): Eigenvalue optimization. Acta Numer. **5**, 149–190
26. Nemirovsky, A., Gahinet, P. (1997): The projective method for solving linear matrix inequalities. Math. Program. **77**, 163–190
27. Nesterov, Yu. (1997): Interior-point methods: An old and new approach to nonlinear programming. Math. Program. **79**, 285–297, October 1997
28. Nesterov, Yu. (1997): Quality of semidefinite relaxation for nonconvex quadratic optimization. CORE Discussion, Paper # 9719
29. Nesterov, Yu. (1998): Private communication. Center of Operations Research and Econometrics, Université Catholique de Louvain, Belgium, May 1998
30. Nesterov, Yu., Nemirovsky, A. (1988): A general approach to polynomial-time algorithms design for convex programming. Technical report, Centr. Econ. & Math. Inst., USSR Academy of Sciences, Moscow, USSR, 1988
31. Nesterov, Yu., Nemirovsky, A. (1994): Interior-point polynomial methods in convex programming: Theory and applications, volume 13 of Studies in Applied Mathematics. SIAM, Philadelphia, PA
32. Oustry, F. (1998): Vertical developments of a convex function. J. Convex Anal. **5**, 153–170
33. Oustry, F. (1999): The $\mathcal{U}$-Lagrangian of the maximum eigenvalue function. SIAM J. Optim. **9**, 526–549
34. Overton, M.L. (1992): Large-scale optimization of eigenvalues. SIAM J. Optim. **2**, 88–120
35. Overton, M.L., Womersley, R.S. (1995): Second derivatives for optimizing eigenvalues of symmetric matrices. SIAM J. Matrix Anal. Appl. **16**, 667–718, July 1995
36. Overton, M.L., Ye., X. (1994): Toward second-order methods for structured nonsmooth optimization. In: Gomez, S., Hennart, J.-P., eds., Advances in Optimization and Numerical Analysis, pp. 97–109. Kluwer Academic Publishers
37. Polak, E., Wardi, Y. (1982): Nondifferentiable optimization algorithm for designing control systems having singular value inequalities. Automatica **18**, 267–283
38. Potra, F.A., Sheng, R. (1998): A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. SIAM J. Optim. **8**, 1007–1028
39. Schramm, H., Zowe, J. (1992): A version of the bundle idea for minimizing a nonsmooth function: conceptual idea, convergence analysis, numerical results. SIAM J. Optim. **2**, 121–152
40. Schwartz, L. (1967): Cours d'analyse, Vol. 1. Hermann, Paris
41. Shapiro, A., Fan, M.K.H. (1995): On eigenvalue optimization. SIAM J. Optim. **5**, 552–568
42. Trefethen, L.N. (1997): Pseudospectra of linear operators. SIAM Rev. **39**, 383–406
43. Trefethen, L.N., Bau, III D. (1997): Numerical Linear Algebra. SIAM, April 1997
44. Ye, D.Y. (1993): Sensitivity analysis of the greatest eigenvalue of a symmetric matrix via the ε-subdifferential of the associated convex quadratic form. J. Optim. Theory Appl. **76**, February 1993