Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 16, No. 1, pp. 129–142, 2020

A $\mathcal{VU}$-DECOMPOSITION TECHNIQUE FOR A CLASS OF EIGENVALUE OPTIMIZATIONS*

Ming Huang, Sida Lin, Jinlong Yuan, Zhaohua Gong† and Xijun Liang

Abstract: The past decade has seen the tremendous development of a number of classes of nonsmooth functions, especially those possess smooth substructure, such as max-type functions and maximum eigenvalue functions. In this paper, we mainly consider a special class of eigenvalue optimization problems: the arbitrary eigenvalue functions. For this class of functions, we can extract this property of smooth part similarly. We give second-order expansions for this class of nonsmooth functions from the $\mathcal{VU}$-space decomposition point of view. A new $\mathcal{U}$-Lagrangian function is proposed, and expressions for the associated second-order objects are given in terms of $\mathcal{U}$-subspace Hessians. Moreover, we explore an algorithmic framework with superlinear convergence.

Key words: *nonsmooth optimization, eigenvalue optimization, $\mathcal{VU}$-decomposition, second-order derivative, superlinear convergence*

Mathematics Subject Classification: *15A18, 49J52*

---

1 Introduction

There are many situations in operations research where one has to optimize a function which fails to have derivatives for some values of the variables. This is what nonsmooth optimization (or nondifferentiable optimization) deals with. Nonsmooth analysis is essential for the understanding of nonsmooth optimization problems. In applications, nonsmoothness of functions typically does not appear in a general way, but in a structured manner. The problem of defining good structures for nonsmooth functions has been addressed by many authors in various ways. Often, the focus of this research is on functions with some sort of potential underlying smooth substructure. This substructure is then exploited for algorithmic purposes, to create calculus rules, or develop stability analysis.

Optimizations of eigenvalues of real symmetric matrices have become an independent area of research with both theoretical and practical aspects since 1980s [11, 13, 22–25, 27].

*This work is supported by the National Natural Science Foundation of China (Grant Nos. 11501080, 11626053, 11701063, 11771008, 71831002, 11901075, 51979034), the China Postdoctoral Science Foundation (Grant Nos. 2019M651091, 2019M661073), the China Scholarship Council (No. 201902575002), the Fundamental Research Funds for the Central Universities (Grant Nos. 3132018215, 3132019179, 3132019323, 3132019501, 3132019502, DUT19LK37), the Scientific Research Foundation Funds of DLMU (Grant No. 02501102), the National Science Foundation of Shandong Province of China (Grant Nos. ZR2017MA005, ZR2019MA031), the Program for Innovative Research Team in University (Grant No. IRT-17R13) and the Xinghai Project of Dalian Maritime University.

†Corresponding Author.

Early contributions owe to [5, 8, 26, 31]. The problems enjoy great importance in physics, engineering, statistics, and finance; see, e.g., composite materials [6], experimental design [29], optimal system design [1,17,19,21,30,33,34], shape optimization [7], robust optimization [9, 14–16, 18, 32], relaxations of combinatorial optimization problems [10] and so on. A good recent survey on this topic with numerous references therein is [13]. One of the main difficulties with numerical analysis of such problems is that the eigenvalues, considered as functions of a symmetric matrix, are not smooth at those points where they coalesce. This causes that the problems are typically nondifferentiable.

The algorithms for optimization of nonsmooth functions applied to eigenvalue optimization problems could be traced back to 1970's and 1980's, where the methods were mainly the first-order form. At the same time various attempts were made to develop second-order theoretical analysis for nonsmooth optimization problems. As far as the second-order analysis is concerned, Shapiro and Fan [27], as well as Overton and Womersley [23, 25], have considered the problem from an algorithmic point of view and their independent searches have led to two algorithms that use second-order information. Overton et al studied the following particular eigenvalue problem

$$(P) \qquad \min_{A\in S_n} \lambda_1(A),$$

where $\lambda_1(\cdot)$ is the maximum eigenvalue function. If the multiplicity r of $\lambda_1(A^*)$ at an optimal point A^* is known, then the approach consists of minimizing the maximum eigenvalue subject to the constraint whose multiplicity is r. A local C^2-parametrization of (P) is then used to develop a successive quadratic programming method.

The $\mathcal{VU}$-decomposition theory was firstly developed in 2000's [12]. Lemaréchal, Oustry, and Sagastizábal created $\mathcal{VU}$-decompositions, and proposed the notions of "smooth fast tracks" and "primal-dual gradient structures" which have their roots in "$\mathcal{VU}$-decompositions" and the "$\mathcal{U}$-Lagrangian". They examined methods to create second-order expansions for non-smooth functions. These decompositions separate the domain of a convex function into directions along which it behaves smoothly ($\mathcal{U}$), and directions perpendicular to smooth behaviour ($\mathcal{V}$). Using these subspaces they defined the $\mathcal{U}$-Lagrangian, a smooth estimate of the function along the $\mathcal{U}$ subspace which can be used to create the desired second-order expansion.

Here we apply the $\mathcal{VU}$-decomposition theory to more general case: the function of arbitrary eigenvalues, This kind of functions has smooth substructure, which is worth of exploring more fast algorithm. Its form is as follows

$$(P_1) \qquad \min_{A\in S_n} \lambda_r(A), \tag{1.1}$$

where $\lambda_r(\cdot)$ is the eigenvalue function and here we denote that these eigenvalues satisfy the decreasing order of $\lambda_1(\cdot) \geq \lambda_2(\cdot) \geq \cdots \geq \lambda_n(\cdot)$. In fact, the arbitrary function $\lambda_r(A)$ can be written as

$$\lambda_r(A) = \phi_1(A) - \phi_2(A),$$

where $\phi_1(A) := \lambda_1(A) + \lambda_2(A) + \cdots + \lambda_r(A) = \sum_{i=1}^{r} \lambda_i(A)$, and $\phi_2(A) := \lambda_1(A) + \lambda_2(A) + \cdots + \lambda_{r-1}(A) = \sum_{i=1}^{r-1} \lambda_i(A)$. In Ref. [24] Overton and Womersley showed that the function of the sum of the largest eigenvalues is convex. So $\phi_1(A)$ and $\phi_2(A)$ are both convex. Their difference is called the difference of convex functions, i.e., D.C. for short. Eigenvalue programs (P_1) have been intensely studied since 1990s. They arise in many applications, such as in automatic control, finance, statistics and design engineering. The readers can see the references [2,13,28] for detail. Here we suppose the function $\phi_1(A)$ is smooth, i.e., $\lambda_r(A)$ ranks the last of the equal eigenvalues. Such problems involve many practical applications, including minimum eigenvalue (a concave programming actually), the second minimum eigenvalue, etc.

Because λ_r is not convex anymore, the corresponding $\mathcal{U}$-Lagrangian cannot be directly applied. Our contribution is mainly stated as follows: from a theoretical point of view, we will use a new $\mathcal{U}$-Lagrangian function to derive an explicit expression for a second-order operator in the case of the arbitrary eigenvalues: $\mathcal{U}$-Hessian. The $\mathcal{U}$-Lagrangian function is D.C. form instead of convex. The resulting $\mathcal{VU}$-decomposition algorithms make a step in the $\mathcal{V}$-subspace, followed by a $\mathcal{U}$-Newton move in order to obtain superlinear convergence. As far as I know, no existing algorithms in the literature is given to compute this problem. This is the first superlinear algorithm to study the arbitrary eigenvalue.

The layout of the paper is as follows. In Section 2, we outline the notation used in this paper and provide the precise definitions from $\mathcal{VU}$ space decomposition theory. Section 3 presents the $\mathcal{U}$-Lagrangian of the arbitrary eigenvalue λ_r. Meantime, the second-order derivatives can be explicitly computed. A conceptual space-decomposition algorithm is proposed, and the corresponding superlinear convergence is proven in Section 4. Finally, in Section 5, we summarize the obtained results, and discuss several questions arising from this paper.

2 Preparation and Preliminary Results

In this section, we recall some classical notions and basic results from nonsmooth analysis and $\mathcal{VU}$-decomposition needed in what follows. Our notation is basically standard.

Let S_n be the space of $n \times n$ symmetric matrices, S_n^+ stands for the cone of $n \times n$ positive semidefinite symmetric matrices. Define $\text{proj}_{\mathcal{U}} : S_n \mapsto \mathcal{U}$ is a projection operator onto the subspace $\mathcal{U}$ of S_n, $\text{proj}_{\mathcal{U}}^* : \mathcal{U} \mapsto S_n$ is the canonical injection $\mathcal{U} \ni u \mapsto u \oplus 0 \in S_n$. Define $X \cdot Y := \text{tr}XY$ as Fröbenius scalar product of $X, Y \in S_n$, and $X^\dagger$ indicates Moore-Penrose inverse of X. Suppose X lies on the submanifold $\mathcal{M}(p,q) := \{X \in S_n : \lambda_p(X) > \lambda_{p+1}(X) = \cdots = \lambda_k(X) = \cdots = \lambda_q(X) > \lambda_{q+1}(X)\}$, where $\mathcal{M}(p,q)$ is a C^∞-submanifold of S_n, then the multiplicity of k-th largest eigenvalue $\lambda_k(X)$ of X is $q - p \geq 1$. Let $E_{p,q}(X)$ be the eigenspace associated with $\lambda_{p+1}, \ldots, \lambda_q$, $Q_{p,q}(X) := Q_1(X)$ be an orthonormal basis of $E_{p,q}(X)$, $T_{\mathcal{M}}(X)$ and $N_{\mathcal{M}}(X)$ are respectively the tangent and normal spaces to the manifold $\mathcal{M}$ at $X \in \mathcal{M}$. Denote the relative interior of the set S as $\text{ri}\, S$, and the interior of the set S as int S. The trace of a matrix A is written as $\text{tr}\, A$.

We recall the definition of $\mathcal{VU}$-decomposition and the $\mathcal{U}$-Lagrangian for a convex function.

Definition 2.1. Let f be a convex function which is finite at the point $\bar{x} \in R^n$. For a subgradient $g \in \text{ri}\partial f(\bar{x})$, we define the $\mathcal{VU}$-decomposition as the subspaces

$$\mathcal{U} := N_{\partial f(\bar{x})}(g) \quad \text{and} \quad \mathcal{V} := T_{\partial f(\bar{x})}(g).$$

We have noticed that $\mathcal{U} = \mathcal{V}^\perp$. These spaces represent the directions from $\bar{x}$ for which f behaves nonsmoothly ($\mathcal{V}$) and smoothly ($\mathcal{U}$). The goal is then to find a smooth function which describes f in the directions of $\mathcal{U}$. For any point $x \in R^n$, we may express x via its projections onto $\mathcal{U}$ and $\mathcal{V}$. We use the compact notation $\oplus$ for such decomposition, and write $R^n = \mathcal{U} \oplus \mathcal{V}$, i.e.,

$$R^n \ni x = \text{proj}_{\mathcal{U}}(x) + \text{proj}_{\mathcal{V}}(x) = x_{\mathcal{U}} \oplus x_{\mathcal{V}} \in \mathcal{U} \times \mathcal{V}.$$

From Definition 2.1, the relative interior of $\partial f(\bar{x})$, denoted by $\mathrm{ri}\partial f(\bar{x})$, is the interior of $\partial f(\bar{x})$ relative to its affine hull, a manifold that is parallel to $\mathcal{V}$. Accordingly,

$$\bar{g} \in \mathrm{ri}\partial f(\bar{x}) \quad \Rightarrow \quad \bar{g} + (B(0,\eta) \cap \mathcal{V}) \subset \partial f(\bar{x}) \quad \text{for some } \eta > 0,$$

where $B(0,\eta)$ denotes a ball in R^n centered at 0, with radius η. An extremely useful theorem on alternate definitions of $\mathcal{VU}$-decompositions follows, which is stated in [12].

Proposition 2.2. *Let f be a convex function, which is finite at the point $\bar{x} \in R^n$. Then the $\mathcal{VU}$-decomposition of f is independent of the subgradient $g \in \mathrm{ri}\partial f(\bar{x})$ chosen. Moreover, the following subspaces are equal to $\mathcal{U}$:*

(i) *The subspace perpendicular to the affine plane of $\partial f(\bar{x})$,*

(ii) $\{d \in R^n : \sup_{g \in \partial f(\bar{x})} \langle g, d\rangle = \inf_{g \in \partial f(\bar{x})} \langle g, d\rangle\}$,

(iii) $\{d \in R^n : f'(\bar{x}; d) = -f'(\bar{x}; -d)\}$,

(iv) $\{d \in R^n : \langle g, d\rangle = \langle g, d\rangle \text{ } for \text{ } all \text{ } g \in \partial f(\bar{x})\}$.

Next we are ready to define the $\mathcal{U}$-Lagrangian function, whose solution mapping is key in defining smooth trajectories.

Definition 2.3. For a convex function f on R^n, which is finite at the point $\bar{x}$, given a subgradient $\bar{g} \in \partial f(\bar{x})$ with $\mathcal{V}$-component $\bar{g}_{\mathcal{V}}$, the $\mathcal{U}$-Lagrangian of f, depending on $\bar{g}_{\mathcal{V}}$, is defined by

$$\mathcal{U}(\bar{x}) \ni u \mapsto L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}}) = \min_{v \in \mathcal{V}(\bar{x})} \{f(\bar{x} + u \oplus v) - \langle \bar{g}_{\mathcal{V}}, v\rangle_{\mathcal{V}}\}, \tag{2.1}$$

where $\mathcal{V}(\bar{x})$ and $\mathcal{U}(\bar{x})$ are the $\mathcal{VU}$-decomposition subspaces, and $\langle \cdot, \cdot\rangle_{\mathcal{V}}$ denotes a scalar product induced in the subspace $\mathcal{V}$. When the infimum in (2.1) is attained, the set of corresponding $\mathcal{V}$-space minimizers is defined by

$$W(u; \bar{g}_{\mathcal{V}}) = \{v \in \mathcal{V}(\bar{x}) : L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}}) = f(\bar{x} + u \oplus v) - \langle \bar{g}_{\mathcal{V}}, v\rangle_{\mathcal{V}}\}. \tag{2.2}$$

A point to notice is that, the vector $\bar{g}_{\mathcal{V}}$ in our notation $L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}})$ plays the roles of a multiplier vector, such as one that occurs in a Lagrangian from constrained optimization, because multipliers coming from the $\mathcal{V}$-subspace minimization in (2.1) hinge on $\bar{g}_{\mathcal{V}}$.

When $W(u; \bar{g}_{\mathcal{V}})$ is nonempty, the associated $\mathcal{U}$-Lagrangian is a convex function that is differentiable at $u = 0$, with

$$\nabla L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) = \bar{g}_{\mathcal{U}} = g_{\mathcal{U}} = \mathrm{proj}_{\mathcal{U}(\bar{x})} g \quad for \text{ } all \text{ } g \in \partial f(\bar{x}). \tag{2.3}$$

For $u \neq 0$, the subdifferential of $L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}})$ can be written as

$$\begin{aligned} \partial L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}}) &= \mathrm{proj}_{\mathcal{U}(\bar{x})}[\partial f(\bar{x} + u \oplus v) \cap (g + \mathcal{U}(\bar{x}))] \\ &= \{g_{\mathcal{U}} : g_{\mathcal{U}} \oplus \bar{g}_{\mathcal{V}} \in \partial f(\bar{x} + u \oplus v)\}, \end{aligned} \tag{2.4}$$

where v is taken arbitrary in $W(u; \bar{g}_{\mathcal{V}})$; in addition, the multifunction $u \mapsto \partial L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}})$ is continuous at $u = 0$:

$$\lim_{u \to 0} \partial L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}}) = \nabla L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}). \tag{2.5}$$

Furthermore, if $g \in \mathrm{ri}\,\partial f(\bar{x})$, then $W(u; \bar{g}_{\mathcal{V}})$ is nonempty, with each $v \in W(u; \bar{g}_{\mathcal{V}})$ is $o(\|u\|)$, i.e.,

$$\sup_{v \in W(u; \bar{g}_{\mathcal{V}})} \|v\| = o(\|u\|), \tag{2.6}$$

and the multifunction $u \mapsto W(u; \bar{g}_{\mathcal{V}})$ is continuous at $u = 0$:

$$\lim_{u \to 0} W(u; \bar{g}_{\mathcal{V}}) = \{0\}. \tag{2.7}$$

When $u = 0$, we have $W(0; \bar{g}_{\mathcal{V}}) = \{0\}$ and $L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) = f(\bar{x})$.

Let $f : R^n \to R$ be a lower semicontinuous function so that its epigraph, denoted and defined by $\text{epi}f := \{(x, \beta) \in R^n \times R : \beta \geq f(x)\}$, is a closed set in R^{n+1}. Take $x \in R^n$, where f is finite-valued, and consider the Clarke cone normal to epigraph of f at $(x, f(x))$. The set of Clarke subgradients of f at x is denoted and defined by

$$\partial_C f(x) := \{g : (g, -1) \in \bar{N}_{\text{epi}f}(x, f(x))\},$$

where $\bar{N}_{\text{epi}f}(x, f(x))$ is the Clarke normal cone of $\text{epi}f$ at $(x, f(x))$. When f is Lipschitz around x, from [3], $\partial_C f(x)$ is the convex hull of all possible limits of gradients at points of differentiability of f in sequences converging to x.

3 $\mathcal{VU}$-Space Decomposition

This section is devoted to the main results in this paper. Here we compute $\mathcal{VU}$-space decomposition based on $\mathcal{VU}$ theory. Next, a new $\mathcal{U}$-Lagrangian function is introduced for deducing the corresponding second-order terms. Furthermore, the second-order expansion of $\lambda_r(A)$ along some smooth trajectory is presented.

In order to give the explicit structure of our $\mathcal{VU}$-space decomposition, we need the following subdifferential results.

Proposition 3.1. *For the function λ_r defined in (P_1) and $\bar{A} \in S_n$, we have*

(i) $\lambda_r'(\bar{A}; H) = \phi_1'(\bar{A}; H) - \phi_2'(\bar{A}; H) = \langle \nabla \phi_1(\bar{A}), H \rangle - \phi_2'(\bar{A}; H) = \lambda_{\min}(Q_1^T H Q_1), \quad \forall H \in S_n$;

(ii) $G \in \partial_C \lambda_r(\bar{A}) \Leftrightarrow G \in \nabla \phi_1(\bar{A}) - \partial \phi_2(\bar{A}) \Leftrightarrow \nabla \phi_1(\bar{A}) - G \in \partial \phi_2(\bar{A})$; *where*

$$\begin{aligned}
&\partial_C \lambda_r(\bar{A}) \\
&= \nabla \phi_1(\bar{A}) - \partial \phi_2(\bar{A}) \\
&= P_1 P_1^T + Q_1 Q_1^T - \{U \in S_n : 0 \preceq \tilde{U} \preceq I, \text{tr}\tilde{U} = r - 1 - l, U = P_1 P_1^T + Q_1 \tilde{U} Q_1^T\} \\
&= \{Q_1 \tilde{U} Q_1^T : 0 \preceq \tilde{U} \preceq I, \text{tr}\tilde{U} = 1\}.
\end{aligned}$$

Proof. (i) It is obvious from the structure of the function in (P_1).

(ii) In view of Proposition 2.11 of [4], we have

$$\partial_C \lambda_r(\bar{A}) \subseteq \nabla h_1(\bar{A}) - \partial h_2(\bar{A}).$$

To complete the proof, it suffices to show that the opposite inclusion relation is true. Taking any $\Xi \in \nabla \phi_1(\bar{A}) - \partial \phi_2(\bar{A})$, one has $\nabla \phi_1(\bar{A}) - \Xi \in \partial \phi_2(\bar{A})$. For the subdifferential of λ_r, according to [24], we can easily deduce the corresponding results. □

Remark 3.2. Because λ_r ranks last in a group of equal eigenvalues, $\lambda_r'(\bar{A}; H)$ is the smallest eigenvalue of $Q_1^T H Q_1$; one can say that λ_r imitates the smallest eigenvalue λ_n. This is a situation where $\lambda_r'(\bar{A}; \cdot)$ is concave.

In Ref. [20], Mifflin and Sagastizábal studied explicitly the $\mathcal{VU}$-decomposition of a special class of the nonconvex lower semi-continuous functions, whose definition is stated as follows:

Definition 3.3. Given an lower semi-continuous function f, a point $\bar{x} \in R^n$, where $f(\bar{x})$ is finite and $\partial_C f(\bar{x})$ is nonempty, and arbitrary subgradient $g \in \partial_C f(\bar{x})$, the orthogonal subspaces

$$\mathcal{V} := lin(\partial_C f(\bar{x}) - g) \qquad and \quad \mathcal{U} := \mathcal{V}^{\perp} \tag{3.1}$$

define the $\mathcal{VU}$-space decomposition; in other words, $\mathcal{V}$ is the subspace parallel to the affine hull of $\partial_C f(\bar{x})$.

Evidently, the structure is similar to (i) of Proposition 2.2. It is not hard to obtain that the results are still satisfied when the function is lower semi-continuous.

Now we are ready to establish our $\mathcal{VU}$-decomposition structure for λ_r.

Proposition 3.4. (i) *Suppose the function λ_r is defined in (P_1). The $\mathcal{VU}$-space decompositions $\mathcal{U}(\bar{A})$ and $\mathcal{V}(\bar{A})$ are, respectively, written as in the following way:*

$$\mathcal{U}(\bar{A}) = \{H \in S_n : Q_1^T H Q_1 = \frac{1}{r-l}\mathrm{tr}(Q_1^T H Q_1) I_{r-l}\}, \tag{3.2}$$

$$\mathcal{V}(\bar{A}) = \{Q_1 Z Q_1^T : Z \in S_{r-l}^{+},\ \mathrm{tr}Z = 0\}. \tag{3.3}$$

(ii) *If $p+1 = q$, then $\mathcal{U}(\bar{A}) = S_n$, and $\mathcal{V}(\bar{A}) = \{0\}$.*

(iii) *If $\mathrm{int}\,\partial_C \lambda_r(\bar{A}) \neq \emptyset$, then we have $\mathcal{U}(\bar{A}) = \{0\}$, and $\mathcal{V}(\bar{A}) = S_n$.*

Proof. According to the item (ii) of Proposition 2.2, we have $\mathcal{U}(x) = \{d \in R^n : f'(x; d) = -f'(x; -d)\}$, $\mathcal{V}(x) = \mathcal{U}(x)^{\perp}$, so we only compute $D \in S_n$ which satisfies

$$\begin{aligned}
\mathcal{U}(\bar{A}) &= \{D : \lambda_r'(\bar{A}; D) = -\lambda_r'(\bar{A}; -D)\} \\
&= \{D : \max_{G \in \partial_C \lambda_r(\bar{A})} \langle G, D\rangle = \min_{G \in \partial_C \lambda_r(\bar{A})} \langle G, D\rangle\} \\
&= \{D : \max_{\tilde{U}} \langle \tilde{U}, Q_1^T D Q_1\rangle = \min_{\tilde{U}} \langle \tilde{U}, Q_1^T D Q_1\rangle\} \\
&= \{D : Q_1^T D Q_1 = \frac{1}{r-l}\mathrm{tr}(Q_1^T D Q_1) I_{r-l}\}.
\end{aligned}$$

Therefore the above equation gives (3.2). The remaining equality follows directly from the definition of $\mathcal{V}(\bar{A})$, $\mathcal{V}(\bar{A}) = \mathcal{U}(\bar{A})^{\perp}$.

For the item (ii), when $p+1 = q$, the multiplicity of λ_r is single. The function λ_r is a differentiable function at $\bar{A}$, and $\mathcal{U}(\bar{A}) = S_n$, and $\mathcal{V}(\bar{A}) = \{0\}$.

If $\mathrm{int}\,\partial_C \lambda_r(\bar{A}) \neq \emptyset$, then $\partial \lambda_r(\bar{A})$ has full dimension. So $\mathcal{U}(\bar{A}) = \{0\}$, and $\mathcal{V}(\bar{A}) = S_n$. This gives item (iii). □

Next we present the $\mathcal{U}$-Lagrangian function of the function λ_r.

Let $\nabla \phi_1(\bar{A}) := \tilde{G} = \tilde{G}_{\mathcal{U}} \oplus \tilde{G}_{\mathcal{V}} = \mathrm{proj}_{\mathcal{U}(\bar{A})}\tilde{G} \oplus \mathrm{proj}_{\mathcal{V}(\bar{A})}\tilde{G}$ and $H_1 := \nabla^2 \phi_1(\bar{A})$. Because $\phi_1(\cdot)$ is a convex function, H_1 is actually positive semi-definite. Denote $\bar{\mathbb{U}}$ and $\bar{\mathbb{V}}$ respectively as the basis matrix of the subspace $\mathcal{U}$ and $\mathcal{V}$. Associated with $\bar{G} = \bar{G}_{\mathcal{U}} \oplus \bar{G}_{\mathcal{V}} \in \partial \phi_2(\bar{A})$, define the $\mathcal{U}$-Lagrangian of λ_r as

$$L_{\mathcal{U}}(U; \bar{G}) = (\phi_1(\bar{A}) + \langle \tilde{G}_{\mathcal{U}}, U\rangle_{\mathcal{U}} + \frac{1}{2}\langle \bar{\mathbb{U}}^T H_1 \bar{\mathbb{U}}\, U, U\rangle_{\mathcal{U}}) - (\inf_{V \in \mathcal{V}}\{\phi_2(\bar{A} + U \oplus V) - \langle \bar{G}_{\mathcal{V}}, V\rangle_{\mathcal{V}}\}), \tag{3.4}$$

and the set of minimizers in $\mathcal{V}$-spaces is stated as

$$W(U) = \arg\min_{V\in\mathcal{V}}\{\phi_2(\bar{A}+U\oplus V) - \langle \bar{G}_{\mathcal{V}}, V\rangle_{\mathcal{V}}\}. \tag{3.5}$$

Denote $L^2_{\mathcal{U}}(U;\bar{G}) := \inf_{V\in\mathcal{V}}\{\phi_2(\bar{A}+U\oplus V) - \langle \bar{G}_{\mathcal{V}}, V\rangle_{\mathcal{V}}\}$.

Remark 3.5. In fact, the function $L^2_{\mathcal{U}}(U;\bar{G})$ is exactly the same as the classical $\mathcal{U}$-Lagrangian (2.1). But $L_{\mathcal{U}}(U;\bar{G})$ is different from (2.1), because the function $L_{\mathcal{U}}(U;\bar{G})$ is usually not convex (the following item (i) of Proposition 3.6). In order to be formally consistent with (2.1) in the appellation, we still call (3.4) the $\mathcal{U}$-Lagrangian function of λ_r. When $l+1=r$, i.e., the multiplicity of λ_r is single, λ_r reduces to the smooth function, and (3.4) coincides with the second-order expansion of λ_r.

The first part of $L_{\mathcal{U}}(U;\bar{G})$, $\phi_1(\bar{A}) + \langle \tilde{G}_{\mathcal{U}}, U\rangle_{\mathcal{U}} + \frac{1}{2}\langle \bar{\mathbb{U}}^T H_1 \bar{\mathbb{U}}\, U, U\rangle_{\mathcal{U}}$, is actually the second-order expansion of $\phi_1(\cdot)$ along the $\mathcal{U}$-subspace.

We see presently how the following proposition formulates the properties for $L_{\mathcal{U}}(U;\bar{G})$ and the set of minimizers $W(U)$.

Proposition 3.6. *Let the function λ_r is defined in (P_1). We have the following assertions:*

(i) *The function $L_{\mathcal{U}}(U;\bar{G})$ defined in (3.4) is a finite-valued D.C. function.*

(ii) *A minimum point $W\in W(U)$ in (3.5) is characterized by the existence of some $G\in \partial_C\lambda_r(\bar{A}+U\oplus W)$ such that $G_{\mathcal{V}} = \tilde{G}_{\mathcal{V}} - \bar{G}_{\mathcal{V}}$.*

(iii) *In particular, $0\in W(0)$ and $L_{\mathcal{U}}(0;\bar{G}) = \lambda_r(\bar{A})$.*

(iv) *If $\bar{G}\in \mathrm{ri}\,\partial\phi_2(\bar{A})$, then $W(U)$ is nonempty for each $U\in\mathcal{U}(\bar{A})$ and $W(0)=\{0\}$.*

Proof. (i) By Theorem 3.2 (i) in [12], the function $L^2_{\mathcal{U}}(U;\bar{G})$ defined in (3.4) is convex and finite everywhere. Because $\phi_1(\bar{A})$ is convex, $\phi_1(\bar{A}) + \langle \tilde{G}_{\mathcal{U}}, U\rangle_{\mathcal{U}} + \frac{1}{2}\langle \bar{\mathbb{U}}^T H_1 \bar{\mathbb{U}}\, U, U\rangle_{\mathcal{U}}$ as the second-order expansion of $\phi_1(\cdot)$ along the $\mathcal{U}$-subspace is actually also convex. Therefore, $L_{\mathcal{U}}(U;\bar{G})$ which is difference of two convex functions is finite everywhere, i.e., it is just a D.C. function.

(ii) Denote the inner function as $h(\cdot) = \phi_2(\bar{A}+U\oplus\cdot) - \langle \bar{G}_{\mathcal{V}}, \cdot\rangle_{\mathcal{V}}$ for infimum of (3.4). By convex analysis we have

$$\begin{aligned} W\in W(U) \quad &\Leftrightarrow \quad 0\in\partial h(W) = \partial\phi_2(\bar{A}+U\oplus W)\cap\mathcal{V} - \bar{G}_{\mathcal{V}} \\ &\Leftrightarrow \quad \bar{G}_{\mathcal{V}}\in\partial\phi_2(\bar{A}+U\oplus W)\cap\mathcal{V} \\ &\Leftrightarrow \quad \bar{G}_{\mathcal{V}} - \tilde{G}_{\mathcal{V}}\in\partial\phi_2(\bar{A}+U\oplus W)\cap\mathcal{V} - \tilde{G}_{\mathcal{V}} \\ &\Leftrightarrow \quad \exists\, G = G_{\mathcal{U}}\oplus G_{\mathcal{V}}\in\partial_C\lambda_r(\bar{A}+U\oplus W),\ \text{s.t. } G_{\mathcal{V}} = \tilde{G}_{\mathcal{V}} - \bar{G}_{\mathcal{V}}. \end{aligned}$$

So item (ii) follows.

(iii) For $U=0$, take $W=0$ and $G=\tilde{G}-\bar{G}\in\partial_C\lambda_r(\bar{A}) = \partial_C\lambda_r(\bar{A}+0\oplus 0)$ in (ii), so we have $0\in W(0)$ and $L_{\mathcal{U}}(0;\bar{G}) = \lambda_r(\bar{A})$.

(iv) By Theorem 2.4 in [12], it is easy to know that the function $h(\cdot)$ which is inner infimum function is inf-compact in $\mathcal{V}$ and the set $W(U)$ is nonempty. When $U=0$,

$$\phi_2(\bar{A}+0\oplus V) - \langle \bar{G}_{\mathcal{V}}, V\rangle \geq \phi_2(\bar{A}) + \eta\|V\|_{\mathcal{V}},$$

for $V\neq 0$, that is to say, $\|V\|\neq 0$, the above inequality become

$$\phi_2(\bar{A}+0\oplus V) - \langle \bar{G}_{\mathcal{V}}, V\rangle > \phi_2(\bar{A}) = \phi_2(\bar{A}+0\oplus 0) - \langle \bar{G}_{\mathcal{V}}, 0\rangle,$$

which implies that $V=0$ is the unique minimizer.

□

Property 1. (i) If $W(U) \neq \emptyset$, then

$$\partial_C L_{\mathcal{U}}(U;\bar{G}) = \{G_{\mathcal{U}} : G_{\mathcal{U}} \oplus (\tilde{G}_{\mathcal{V}} - \bar{G}_{\mathcal{V}}) \in \partial_C \lambda_r(\bar{A} + U \oplus W), W \in W(U)\}. \tag{3.6}$$

(ii) If $\bar{G} \in \mathrm{ri}\partial\phi_2(\bar{A})$, then $W(U) = o(\|U\|_{\mathcal{U}})$.

(iii) For $U \in \mathcal{U}$ satisfying $W(U) \neq \emptyset$, we have

$$\lambda_r(\bar{A} + U \oplus W) = \lambda_r(\bar{A}) + \langle \tilde{G} - \bar{G}, U \oplus W\rangle + o(\|U\|_{\mathcal{U}}), \quad \forall W \in W(U). \tag{3.7}$$

Proposition 3.7. *The convex function $L^2_{\mathcal{U}}(U,\bar{G})$ is differentiable at $U = 0$ and its gradient is written as*

$$\nabla L^2_{\mathcal{U}}(0,\bar{G}) = \mathrm{proj}_{\mathcal{U}(\bar{A})}\bar{G} = \bar{G}_{\mathcal{U}}. \tag{3.8}$$

Moreover, the D.C. function $L_{\mathcal{U}}(U;\bar{G})$ is also differentiable at $U = 0$,

$$\nabla L_{\mathcal{U}}(0,\bar{G}) = \tilde{G}_{\mathcal{U}} - \bar{G}_{\mathcal{U}}. \tag{3.9}$$

Proof. We can directly apply (2.3) in matrix form for the first development. The second development can be obtained by adopting the composite chain rule. □

We have the following second-order expansion of λ_r:

Theorem 3.8. *Let $D \in S_n$, $\bar{G} \in \mathrm{ri}\partial\phi_2(\bar{A})$ and $\|D\| \to 0$, then*

$$\lambda_r(\bar{A}+D) = \lambda_r(\bar{A}) + \langle \tilde{G} - \bar{G}, D\rangle + \frac{1}{2}\mathrm{proj}_{\mathcal{U}} D \cdot (\bar{\mathbb{U}}^T H_1 \bar{\mathbb{U}} - H_2)(\mathrm{proj}_{\mathcal{U}} D) + o(\|D\|^2), \tag{3.10}$$

where $H_2 = \nabla^2 L^2_{\mathcal{U}}(0,\bar{G})$ is defined by

$$\nabla^2 L^2_{\mathcal{U}}(0,\bar{G}) = \mathrm{proj}_{\mathcal{U}(\bar{A})} \circ H(\bar{A},\bar{G}) \circ \mathrm{proj}_{\mathcal{U}(\bar{A})}{}^*, \tag{3.11}$$

and $H(\bar{A},\bar{G})$ is the symmetric operator defined by

$$\begin{aligned} H(\bar{A},\bar{G}) \cdot Y \quad = \quad & (\bar{G} - P_1(\bar{A})P_1(\bar{A})^T)Y[\lambda^*_{l+1}I_n - \bar{A}]^\dagger + [\lambda^*_{l+1}I_n - \bar{A}]^\dagger Y(\bar{G} - P_1(\bar{A})P_1 \\ & (\bar{A})^T) + P_1(\bar{A})P_1(\bar{A})^T Y[\lambda^*_1 I_n - \bar{A}]^\dagger + [\lambda^*_1 I_n - \bar{A}]^\dagger Y P_1(\bar{A})P_1(\bar{A})^T, \end{aligned} \tag{3.12}$$

and here we assume at $\bar{A}$, $\lambda_1(\bar{A}) = \cdots = \lambda_l(\bar{A})$.

Proof. Denote $D = U \oplus V \in S_n$. In view of (3.4) and the definition of $L^2_{\mathcal{U}}(U;\bar{G})$, $L_{\mathcal{U}}(U;\bar{G})$ can be written as $L_{\mathcal{U}}(U;\bar{G}) = (\phi_1(\bar{A}) + \langle \tilde{G}_{\mathcal{U}}, U\rangle_{\mathcal{U}} + \frac{1}{2}\langle \bar{\mathbb{U}}^T H_1 \bar{\mathbb{U}}\, U, U\rangle_{\mathcal{U}}) - L^2_{\mathcal{U}}(U;\bar{G})$. According to Theorem 4.11 in [22], $L^2_{\mathcal{U}}(U;\bar{G})$ is C^∞, so is $L_{\mathcal{U}}(U;\bar{G})$. On the basis of Theorem 4.12 therein, we obtain

$$\nabla^2 L^2_{\mathcal{U}}(0,\bar{G}) = \mathrm{proj}_{\mathcal{U}(\bar{A})} \circ H(\bar{A},\bar{G}) \circ \mathrm{proj}_{\mathcal{U}(\bar{A})}{}^*,$$

where $H(\bar{A},\bar{G})$ is defined by (3.12), so (3.11) holds.

Let D small enough such that $\bar{A} + D \in \mathcal{M}_{r-l}$, and set $U = \mathrm{proj}_{\mathcal{U}(\bar{A})}D$, $V = V(U) = \mathrm{proj}_{\mathcal{V}(\bar{A})}D$, apply the second-order Taylor expansion with the item (iii) of Proposition 3.6:

$$\begin{aligned} L_{\mathcal{U}}(U;G) \quad &= \quad (\phi_1(\bar{A}) + \langle \tilde{G}_{\mathcal{U}}, U\rangle_{\mathcal{U}} + \frac{1}{2}\langle \bar{\mathbb{U}}^T H_1 \bar{\mathbb{U}} U, U\rangle_{\mathcal{U}}) - L^2_{\mathcal{U}}(U;\bar{G}) \\ &= \quad (\phi_1(\bar{A}) + \langle \tilde{G}_{\mathcal{U}}, U\rangle_{\mathcal{U}} + \frac{1}{2}\langle \bar{\mathbb{U}}^T H_2 \bar{\mathbb{U}} U, U\rangle_{\mathcal{U}}) \end{aligned}$$

$$\begin{aligned} & -L^2_{\mathcal{U}}(0;\bar{G}) + \langle \nabla L^2_{\mathcal{U}}(0;G);U\rangle_{\mathcal{U}} + \frac{1}{2}\langle U, H_1 U\rangle_{\mathcal{U}} + o(\|U\|^2) \\ = \; & \lambda_r(\bar{A}) + \langle \tilde{G}_{\mathcal{U}} - \bar{G}_{\mathcal{U}}, U\rangle_{\mathcal{U}} + \frac{1}{2}\langle U, (\bar{\mathbb{U}}^T H_1 \bar{\mathbb{U}} - H_2)U\rangle_{\mathcal{U}} + o(\|U\|^2), \end{aligned}$$

on the other hand, because $W(U) = V(U)$, so we get

$$\begin{aligned} L_{\mathcal{U}}(U;G) &= (\phi_1(\bar{A}) + \langle \tilde{G}_{\mathcal{U}}, U\rangle_{\mathcal{U}} + \frac{1}{2}\langle \bar{\mathbb{U}}^T H_1 \bar{\mathbb{U}} U, U\rangle_{\mathcal{U}}) - (\phi_2(\bar{A} + U \oplus V) - \langle \bar{G}_{\mathcal{V}}, V\rangle_{\mathcal{V}}) \\ &= \lambda_r(\bar{A} + U \oplus V) - \phi_1(\bar{A} + U \oplus V) + \phi_1(\bar{A}) \\ &\quad + \frac{1}{2}\langle \bar{\mathbb{U}}^T H_2 \bar{\mathbb{U}} U, U\rangle_{\mathcal{U}} + \langle \bar{G}_{\mathcal{U}} \oplus \bar{G}_{\mathcal{V}}, U \oplus V\rangle \\ &= \lambda_r(\bar{A} + U \oplus V) - \{\phi_1(\bar{A}) + \langle \tilde{G}, U \oplus V\rangle + o(\|U \oplus V\|)\} + \phi_1(\bar{A}) \\ &\quad + \frac{1}{2}\langle \bar{\mathbb{U}}^T H_1 \bar{\mathbb{U}} U, U\rangle_{\mathcal{U}} + \langle \tilde{G}_{\mathcal{U}} \oplus \bar{G}_{\mathcal{V}}, U \oplus V\rangle \\ &= \lambda_r(\bar{A} + U \oplus V) - \langle \tilde{G}_{\mathcal{V}} - \bar{G}_{\mathcal{V}}, U \oplus V\rangle + o(\|U\|^2). \end{aligned}$$

Finally, recall Theorem 3.4 in [22], we have $V = O(\|U\|^2) = O(\|D\|^2)$, Combine the above two formulas, we obtain (3.10). □

4 A Fast $\mathcal{VU}$-Decomposition Algorithm

Suppose $\bar{A}$ is the minimizer of (P_1), and the initial point is close to the minimizer $\bar{A}$. the $\mathcal{VU}$-decomposition of the function $\lambda_r(\bar{A})$ at $\bar{A}$ has been completed, $\mathcal{U}$-Hessian $H_{\mathcal{U}}\lambda_r(\bar{A})$ exists. Along these lines the conceptual D.C. $\mathcal{VU}$-decomposition algorithm for solving (P_1) given below contains the elements essential for obtaining superlinear convergence.

Algorithm 1. D.C. $\mathcal{VU}$-decomposition algorithm: solving (P_1).
Step 0. (Initialization)
Given the initial point $A_0 \in S_n$ close to the minimizer $\bar{A}$, set $k = 0$.
Step 1. ($\mathcal{V}$-Step)

$$V_k \in \text{Argmin}_{V\in\mathcal{V}} \lambda_r(A_k + 0 \oplus V).$$

Compute $A^c_k = A_k + 0 \oplus V_k$, $\bar{G}^c \in \partial\phi_2(A^c_k)$ satisfies $\bar{G}^c_{\mathcal{V}} = \tilde{G}^c_{\mathcal{V}}$, where $\tilde{G}^c = \nabla\phi_1(A^c_k)$, $G^c = \tilde{G}^c - \bar{G}^c = (\tilde{G}^c_u - \bar{G}^c_u) \oplus 0 \in \partial_C \lambda_r(A^c_k)$.

If $\bar{G}^c_u = \tilde{G}^c_u$, then stop, and A^c_k is the approximated minimizer. Otherwise, go to Step 2.
Step 2. ($\mathcal{U}$-Step)
Solving

$$H_{\mathcal{U}}\lambda_r(\bar{A})\; U = -(\tilde{G}^c_u - \bar{G}^c_u), \tag{4.1}$$

and obtain the solution $U = U_k$.
Step 3. (Corrector)
Compute $A_{k+1} = A^c_k + U_k \oplus 0 = A_k + U_k \oplus V_k$, replace k by $k+1$, and go to Step 1.

Now we give the convergence of D.C. $\mathcal{VU}$-decomposition algorithm.

Theorem 4.1. *Suppose the following conditions holds:*

1. *$\bar{A}$ is the minimizer of (P_1);*

2. *$\mathcal{U}$-Hessian $H_{\mathcal{U}}(\lambda_r(\bar{A}))$ is positive-definite;*
3. $\tilde{G} = \nabla\phi_1(\bar{A}) \in \mathrm{ri}\partial(\phi_2(\bar{A}))$ *and* $\tilde{G}_{\mathcal{V}} = 0$.
Then the sequence $\{A_k^c\}_{k=1}^{\infty}$ produced by the Algorithm 1 converge to $\bar{A}$ superlinearly, i.e.,
$$\|A_{k+1}^c - \bar{A}\| = o(\|A_k^c - \bar{A}\|).$$

Proof. We first show that $\|(A_{k+1}^c - \bar{A})_{\mathcal{U}}\|_{\mathcal{U}} = o(\|(A_k^c - \bar{A})_{\mathcal{U}}\|_{\mathcal{U}})$. Since $\bar{A}$ is a minimizer of (P_1), one has $\tilde{G} = \nabla\phi_1(\bar{A}) \in \partial(\phi_2(\bar{A}))$. Taking $\bar{G} = \tilde{G}$ in (3.4), we have $\bar{G}_{\mathcal{U}} - \tilde{G}_{\mathcal{U}} = 0$. For $U \in \mathcal{U}$ small enough, it follows from (3.6) that

$$\begin{array}{rl} & \{\tilde{G}_u^c - \bar{G}_u^c | (\tilde{G}_u^c - \bar{G}_u^c) \oplus (\tilde{G}_{\mathcal{V}} - \bar{G}_{\mathcal{V}}) \in \partial_C \lambda_r(\bar{A} + U \oplus W), W \in W(U)\} \\ \subset & \tilde{G}_{\mathcal{U}} - \bar{G}_{\mathcal{U}} + H_{\mathcal{U}}\lambda_r(\bar{A})U + o(\|U\|_{\mathcal{U}})B_{\mathcal{U}} \\ = & H_{\mathcal{U}}\lambda_r(\bar{A})U + o(\|U\|_{\mathcal{U}})B_{\mathcal{U}}. \end{array} \tag{4.2}$$

Because U_k is a solution of (4.1) in Step 2, one has from $\bar{G}_{\mathcal{V}}^c = \tilde{G}_{\mathcal{V}}^c$ in Step 1 that

$$-H_{\mathcal{U}}\lambda_r(\bar{A})U_k = \tilde{G}_{\mathcal{U}} - \bar{G}_{\mathcal{U}} \in H_{\mathcal{U}}\lambda_r(\bar{A})(A_k^c - \bar{A})_{\mathcal{U}} + o(\|(A_k^c - \bar{A})_{\mathcal{U}}\|_{\mathcal{U}})B_{\mathcal{U}},$$

and hence
$$-H_{\mathcal{U}}\lambda_r(\bar{A})(U_k + (A_k^c - \bar{A})_{\mathcal{U}}) \in o(\|(A_k^c - \bar{A})_{\mathcal{U}}\|_{\mathcal{U}})B_{\mathcal{U}}.$$
The equality, $\|U_k + (A_k^c - \bar{A})_{\mathcal{U}}\| = o(\|(A_k^c - \bar{A})_{\mathcal{U}}\|_{\mathcal{U}})$, holds by the positive definiteness of $H_{\mathcal{U}}\lambda_r(\bar{A})$. In consequence, for $A_{k+1}^c = A_{k+1} + 0 \oplus V_{k+1} = A_k^c + U_k \oplus V_{k+1}$, the following equalities hold,

$$\begin{array}{rcl} \|(A_{k+1}^c - \bar{A})_{\mathcal{U}}\|_{\mathcal{U}} & = & \|(A_k^c - \bar{A} + U_k \oplus V_{k+1})_{\mathcal{U}}\|_{\mathcal{U}} \\ & = & \|U_k + (A_k^c - \bar{A})_{\mathcal{U}}\|_{\mathcal{U}} \\ & = & o(\|(A_k^c - \bar{A})_{\mathcal{U}}\|_{\mathcal{U}}). \end{array} \tag{4.3}$$

We now show that $\|(A_{k+1}^c - \bar{A})_{\mathcal{V}}\|_{\mathcal{V}} = o(\|(A_k^c - \bar{A})_{\mathcal{U}}\|_{\mathcal{U}})$. Because

$$A_{k+1} + (0 \oplus V) = \bar{A} + ((A_k^c - \bar{A})_{\mathcal{U}} + U_k) \oplus ((A_k^c - \bar{A})_{\mathcal{V}} + V),$$

and
$$\begin{array}{rcl} V_{k+1} & \in & \arg\min_{V\in\mathcal{V}} \lambda_r(A_{k+1} + (0 \oplus V)) \\ & = & \arg\min_{V\in\mathcal{V}} \lambda_r(\bar{A} + ((A_k^c - \bar{A})_{\mathcal{U}} + U_k) \oplus ((A_k^c - \bar{A})_{\mathcal{V}} + V)), \end{array}$$

one has
$$\begin{array}{rcl} V_{k+1} + ((A_k^c - \bar{A})_{\mathcal{V}}) & \in & \arg\min_{V\in\mathcal{V}} \lambda_r(\bar{A} + ((A_k^c - \bar{A})_{\mathcal{U}} + U_k) \oplus V) \\ & = & \arg\min_{V\in\mathcal{V}} \{\lambda_r(\bar{A} + ((A_k^c - \bar{A})_{\mathcal{U}} + U_k) \oplus V) - \langle \bar{G}_{\mathcal{V}}, V\rangle_{\mathcal{V}}\}, \end{array}$$

where $\bar{G}_{\mathcal{V}} = \tilde{G}_{\mathcal{V}} = 0$. Thus,

$$V_{k+1} + (A_k^c - \bar{A})_{\mathcal{V}} \in W((A_k^c - \bar{A})_{\mathcal{U}} + U_k).$$

According to the assumptions and nonempty property of $W(U)$, one has

$$V_{k+1} + (A_k^c - \bar{A})_{\mathcal{V}} = o(\|(A_k^c - \bar{A})_{\mathcal{U}} + U_k\|_{\mathcal{U}}). \tag{4.4}$$

It follows from $(A_{k+1}^c - \bar{A})_{\mathcal{V}} = (A_k^c - \bar{A})_{\mathcal{V}} + V_{k+1}$, combining (4.3) and (4.4), that

$$\begin{array}{rcl} \|(A_{k+1}^c - \bar{A})_{\mathcal{V}}\|_{\mathcal{V}} & = & (A_k^c - \bar{A})_{\mathcal{V}} + V_{k+1} \\ & = & o(\|(A_k^c - \bar{A})_{\mathcal{U}} + U_k\|_{\mathcal{U}}) \\ & = & o(\|(A_{k+1}^c - \bar{A})_{\mathcal{U}}\|_{\mathcal{U}}) \\ & = & o(\|(A_k^c - \bar{A})_{\mathcal{U}}\|_{\mathcal{U}}). \end{array} \tag{4.5}$$

Futhermore, by (4.2) and (4.5) we have

$$\|A_{k+1}^c - \bar{A}\| = \|(A_{k+1}^c - \bar{A})_{\mathcal{U}}\|_{\mathcal{U}} + \|(A_{k+1}^c - \bar{A})_{\mathcal{V}}\|_{\mathcal{V}} = o(\|A_k^c - \bar{A}\|_{\mathcal{U}}) = o(\|A_k^c - \bar{A}\|).$$

The proof is complete.

□

Remark 4.2. Our technique above is highly conceptual, because we need to generate convergent estimates of $\mathcal{V}$, $\mathcal{U}$ and a positive definite $\mathcal{U}$-Hessian $H_{\mathcal{U}}\lambda_r(\bar{A})$ with relation to an optimal solution and a zero subgradient.

5 Concluding Remarks

In this paper, we have discussed the $\mathcal{VU}$-decomposition technique for a class of eigenvalues functions. For the nonconvex eigenvalue with ranking last in the group of equal eigenvalues, the $\mathcal{VU}$-space decomposition approaches are presented. Nonconvex eigenvalue function mentioned in this paper is actually a class of D.C. functions (difference of convex functions). By introducing a new $\mathcal{U}$-Lagrangian function, the first- and second-order derivatives of the $\mathcal{U}$-Lagrangian function can be obtained by $\mathcal{U}$-Lagrangian theory. In addition, a conceptual $\mathcal{VU}$-decomposition algorithm is proposed.

Only the conceptual algorithm for solving such special class of eigenvalue optimization is presented. Is it can be operated in practical application? This is our subsequent work: we will consider to investigate the performance of its rapidly convergent executable algorithm for solving combinatorial optimization problems and optimal control problem, and study how to use bundle techniques to approximate proximal points and other $\mathcal{VU}$-related quantities. In addition, the eigenvalue optimization problem we studied only ranks the last of the equal ones. When it ranks other positions, this case appears much trickier in view of its subdifferential structure. Then, how is the $\mathcal{VU}$-decomposition and $\mathcal{U}$-Lagrangian? It is also a meaningful topic and worth studying deeply in later work.

References

[1] P. Apkarian, D. Noll, J.-B. Thevenet and H.D. Tuan, A spectral quadratic-SDP method with applications to fixed-order H_2 and H_∞ synthesis, *Eur. J. Control* 10 (2004) 527–538.

[2] J.V. Burke, A.S. Lewis and M.L. Overton, Optimal stability and eigenvalue multiplicity, *Found. Comput. Math.* 1 (2001) 205–225.

[3] F. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, 1983; reprinted, SIAM, Philadelphia, 1990.

[4] F.H. Clarke, Y.S. Ledyaev, R.J. Stern and P.R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer-Verlag, New York, 1998.

[5] J. Cullum, W. Donath and P. Wolfe, The minimization of certain nondifferentiable sums of eigenvalues of symmetric matrices, *Math. Program. Stud.* 3 (1975) 35–55.

[6] S. Cox and R. Lipton, Extremal eigenvalue problems for two-phase conductors, *Arch. Ration. Mech. Anal.* 136 (1996) 101–117.

[7] A.R. Diaaz and N. Kikuchi, Solution to shape and topology eigenvalue optimization problems using a homogenization method, *J. Numer. Methods Engineer.* 35 (2005) 1487–1502.

[8] R. Fletcher, Semi-definite matrix constraints in optimization, *SIAM J. Control Optim.* 23 (1985) 493–513.

[9] Z. Gong, C. Liu, J. Sun and K.L. Teo, Distributionally robust L_1-estimation in multiple linear regression, *Optim. Lett.* 3 (2019) 935–947.

[10] C. Helmberg, F. Rendl and R. Weismantel, A semidefinite programming approach to the quadratic knapsack problem, *J. Comb. Optim.* 4 (2000) 197–215.

[11] J.-B. Hiriart-Urruty and D. Ye, Sensivity analysis of all eigenvalues of a symmetric matrix, *Numer. Math.* 70 (1995) 45–72.

[12] C. Lemaréchal, F. Oustry and C. Sagastizábal, The $\mathcal{U}$-Lagrangian of a convex function, *Trans. AMS* 352 (2000) 711–729.

[13] A.S. Lewis and M.L. Overton, Eigenvalue optimization, *Acta Numer.* 5 (1996) 149–190.

[14] B. Li, Y. Rong, J. Sun and K.L. Teo, A distributionally robust linear receiver design for multi-access space-time block coded MIMO systems, *IEEE T. Wirel. Commun.* 16 (2017) 464–474.

[15] B. Li, Y. Rong, J. Sun and K.L. Teo, A distributionally robust minimum variance beamformer design, *IEEE Signal Proc. Let.* 25 (2018) 105–109.

[16] B. Li, X. Qian, J. Sun, K.L. Teo and C.J. Yu, A model of distributionally robust two-stage stochastic convex programming with linear recourse, *Appl. Math. Model.* 58 (2018) 86–97.

[17] C. Liu, Z. Gong, H.W.J. Lee and K.L. Teo, Robust bi-objective optimal control of 1, 3-propanediol microbial batch production process, *J. Process Contr.* 78 (2019) 170–182.

[18] C. Liu, Z. Gong, K.L. Teo, J. Sun and L. Caccetta Robust multi-objective optimal switching control arising in 1, 3-propanediol microbial fed-batch process, *Nonlinear Anal.-Hybri.* 25 (2017) 1–20.

[19] C. Liu, R. Loxton, Q. Lin and K.L. Teo, Dynamic optimization for switched time-delay systems with state-dependent switching conditions, *SIAM J. Control Optim.* 56 (2018) 3499–3523.

[20] R. Mifflin and C. Sagastizábal, $\mathcal{VU}$-smoothness and proximal point results for some nonconvex functions, *Optim. Methods Softw.* 19 (2004) 463–478.

[21] D. Noll, M. Torki and P. Apkarian, Partially augmented Lagrangian method for matrix inequality constraints, *SIAM J. Optim.* 15 (2004) 161–184.

[22] F. Oustry, The $\mathcal{U}$-Lagrangian of the maximum eigenvalue functions, *SIAM J. Optim.* 9 (1999) 526–549.

[23] M.L. Overton, Large-scale optimization of eigenvalues, *SIAM J. Optim.* 2 (1992) 88–120.

[24] M.L. Overton and R.S. Womersley, Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrices, *Math. Program.* 62 (1993) 321–357.

[25] M.L. Overton and R.S. Womersley, Second derivatives for optimizing eigenvalues of symmetric matrices, *SIAM J. Matrix Anal. Appl.* 16 (1995) 697–718.

[26] E. Polak and Y. Wardi, A nondifferential optimization algorithm for the design of control systems subject to singular value inequalities over the frequency range, *Automatica* 18 (1982) 267–283.

[27] A. Shapiro and M.K.H. Fan, On eigenvalue optimization, *SIAM J. Optim.* 5 (1995) 552-556.

[28] M. Torki, First- and second-order epi-differentiability in eigenvalue optimization, *J. Math. Anal. Appl.* 234 (1999) 391–416.

[29] L. Vandenberghe and S. Boyd, Semidefinite programming, *SIAM Rev.* 38 (1996) 49–95.

[30] L. Wang, J. Yuan, C. Wu and X. Wang, Practical algorithm for stochastic optimal control problem about microbial fermentation in batch culture, *Optim. Lett.* 13 (2019) 527–541.

[31] P. Wolfe, A method of conjugate subgradients for minimizing nondifferentiable functions, in: *Nondifferentiable Optimization*, 3, North-Holland, Amsterdam, 1975, pp. 145–173.

[32] J. Yuan, C.Z. Wu, J.X. Ye and J. Xie, Robust identification of nonlinear state-dependent impulsive switched system with switching duration constraints, *Nonlinear Anal. Hybr.* 36 (2020) 100879.

[33] J. Yuan, J. Xie, M. Huang, H. Fan, E. Feng and Z. Xiu, Robust optimal control problem with multiple characteristic time points in the objective for a batch nonlinear time-varying process using parallel global optimization, *Optim. Eng.* Accepted, (2019), https://doi.org/10.1007/s11081-019-09456-z

[34] J. Yuan, J. Xie, C. Liu, K.L. Teo, M. Huang, H. Fan, E. Feng and Z. Xiu, Robust optimization for a nonlinear switched time-delay system with noisy output measurements using hybrid optimization algorithm, *J. Franklin I.* 356 (2019) 9730–9762.

[35] Z. Zhao, B.Z. Braams, M. Fukuda, M.L. Overton and J.K. Percus, The reduced density matrix method for electronic structure calculations and the role of three-index representability conditions, *J. Chem. Phys.* 120 (2004) 2095–2104.

Manuscript received 24 October 2018
revised 13 May 2019
accepted for publication 14 May 2019

Ming Huang
School of Science, Dalian Maritime University
Dalian 116026, China
School of Control Science and Engineering, Dalian University of Technology
Dalian 116024, China
E-mail address: huangming0224@163.com

Sida Lin
School of Control Science and Engineering, Dalian University of Technology
Dalian 116024, China
E-mail address: lsdsl@163.com

Jinlong Yuan
School of Science, Dalian Maritime University
Dalian 116026, China
College of Transportation Engineering, Dalian Maritime University
Dalian, Liaoning 116026, PR China
E-mail address: yuanjinlong0613@163.com

Zhaohua Gong
School of Mathematics and Information Science
Shandong Institute of Business and Technology
Yantai, Shandong 264005, China
E-mail address: zhaohuagong@163.com

Xijun Liang
College of Science, China University of Petroleum
Qingdao 266580, China
E-mail address: liangxijunsd@163.com