Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2014, Article ID 845017, 12 pages
http://dx.doi.org/10.1155/2014/845017

Research Article

The Space Decomposition Theory for a Class of Semi-Infinite Maximum Eigenvalue Optimizations

Ming Huang, Li-Ping Pang, Xi-Jun Liang, and Zun-Quan Xia

CORA, School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China

Correspondence should be addressed to Li-Ping Pang; lppang@dlut.edu.cn

Received 14 June 2013; Accepted 6 November 2013; Published 17 February 2014

Academic Editor: Jinde Cao

We study optimization problems involving eigenvalues of symmetric matrices. We present a nonsmooth optimization technique for a class of nonsmooth functions which are semi-infinite maxima of eigenvalue functions. Our strategy uses generalized gradients and $\mathcal{UV}$ space decomposition techniques suited for the norm and other nonsmooth performance criteria. For the class of max-functions, which possesses the so-called primal-dual gradient structure, we compute smooth trajectories along which certain second-order expansions can be obtained. We also give the first- and second-order derivatives of primal-dual function in the space of decision variables R^m under some assumptions.

1. Introduction

H_∞ output feedback control is an important example of a design problem, where the feedback controller has to respond favorably to several performance specifications. Typically in H_∞ synthesis, the H_∞ channel is used to enhance the robustness of the design. Due to its prominence in practice, H_∞ control has been addressed in various ways over the years.

In nominal H_∞ synthesis, feedback controllers are computed via semidefinite programming (SDP) [1] or algebraic Riccati equations [2]. When structural constraints on the controller are added, the H_∞ synthesis problem is no longer convex. Some of the problems above have even been recognized as NP-hard or as rationally undecidable. These mathematical concepts indicate the inherent difficulty of H_∞ synthesis under constraints on the controller. The H_∞ synthesis problem involves finding an output feedback control matrix K that minimizes the H_∞ norm of a certain transfer function, subject to the constraint that K is stabilizing. This is a challenging problem and even finding a stabilizing K can be difficult. Indeed, if the entries of K are restricted to lie in prescribed intervals, then finding a stabilizing K is an NP-hard problem [3].

H_∞ feedback controller synthesis was one of the motivating application for the development of our work. We consider a linear time invariant dynamical system in the standard LFT form

$$\begin{bmatrix} \dot{x} \\ z \\ y \end{bmatrix} = \begin{bmatrix} A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{bmatrix} \cdot \begin{bmatrix} x \\ w \\ u \end{bmatrix}, \tag{1}$$

where $x \in R^{n_x}$ is the state, $y \in R^{n_y}$ is the output, $u \in R^{n_u}$ is the command input, and $w \in R^{n_w}$, $z \in R^{n_z}$ are the performance channel. To cancel direct transmission from input u to output y, the assumption $D_{22} = 0$ is made. This is without loss of generality (see [4], chapter 17).

Let K be a static feedback controller; then the closed-loop state space data and transfer function $T(K, \cdot)$ read

$$\begin{aligned} \dot{x} &= A(K)x + B(K)w, \\ z &= C(K)x + D(K)w, \\ T(K, jw) &= C(K)\left(jwI - A(K)\right)^{-1}B(K) + D(K), \end{aligned} \tag{2}$$

where

$$A(K) = A + B_2KC_2, \qquad B(K) = B_1 + B_2KD_{21},$$
$$C(K) = C_1 + D_{12}KC_2, \qquad D(K) = D_{11} + D_{12}KD_{21}. \tag{3}$$

Dynamic controllers can be addressed in the same way by prior augmentation of the plant (3) (see, e.g., [5]).

In H_∞ synthesis, we compute K to minimize the H_∞ norm of the transfer function $T(K,\cdot)$; that is,

$$\|T(K,\cdot)\|_\infty := \sup_{w\in[0,\infty]} \sigma_1(T(K,jw)); \tag{4}$$

(see, e.g., [4]). The standard approach to H_∞ synthesis in the literature uses the Kalman-Yakubovich-Popov Lemma and leads to a bilinear matrix inequality (BMI) [6]. Here we use a different and much more direct approach based on our $\mathcal{UV}$-decomposition method. The advantage of this is that Lyapunov variables can be avoided, which is beneficial because they are a source of numerical trouble. Not only does their number grow quadratically with the system order, but they may also cause strong disparity between the optimization variables. The price to be paid for avoiding them is that a difficult semi-infinite and nonsmooth program has to be solved. To synthesize a dynamic controller K of order $n_k \in N$, $n_k \le n_x$, the objective $f : R^{(n_k+n_u)\times(n_k+n_y)} \to R^+$ is defined as follows:

$$f(K) := \max_{w\in[0,\infty]} \lambda_1\left(T(K,jw)^H T(K,jw)\right) = \|T(K,\cdot)\|_\infty^2, \tag{5}$$

which is nonsmooth and nonconvex with two sources of nonsmoothness, the infinite max-operator and the maximum eigenvalue function. In addition, Z^H stands for the conjugate transpose of the complex matrix Z.

The application we have in mind is optimizing the H_∞-norm, which is structurally of the form

$$f(x) = \sup_{\omega\in[0,\infty]} \lambda_1(A(x,\omega)), \tag{6}$$

where $A : R^m \times [0,\infty] \to S_n$ is an operator with values in the space S_n of $n\times n$ symmetric or Hermitian matrices, equipped with the scalar product $X \bullet Y = \mathrm{Tr}(XY)$, and λ_1 denotes the maximum eigenvalue function on S_n.

The above problem (5) can be recast as a case of (6). The program we wish to solve in this paper is

$$\min_{x\in R^m} f(x), \tag{7}$$

where the function f has the form (6).

f is nonsmooth with two possible sources of nonsmoothness: (a) the infinite max-operator and (b) the nonsmoothness of λ_1, which may lead to nonsmoothness of $\lambda_1(A(x,\omega))$ for fixed ω.

Optimization of the H_∞-norm is a prominent application in feedback synthesis, which has been pioneered by Polak and coworkers; see, for instance, [7, 8] and the references given there. Existing methods for the H_∞ synthesis problem are often based on first reformulating the problem into one involving linear matrix inequalities (LMIs) and an additional nonconvex rank constraint or nonconvex equality constraint. Solving methods for such reformulations of the problem include those based on linearization method [9], alternating projections method [10], augmented Lagrangian method [11], and sequential semidefinite programming method [12]. The H_∞ synthesis problem can also be reformulated into a problem involving bilinear matrix inequalities (BMIs). Dealing with such reformulations of the problem includes [12, 13] (see also the references therein). A disadvantage of these approaches is that they require the introduction of Lyapunov variables. As the number of Lyapunov variables grows quadratically with the number of state variables, the total number of variables can be quite large and even problems of moderate size can lead to numerical difficulties [14].

In this paper, the H_∞ synthesis problem is posed as an unconstrained, nonsmooth, nonconvex minimization problem and requires special optimization techniques. Our approach avoids the use of Lyapunov variables; hence, it is well suited for optimizing our reformulation of the H_∞ synthesis problem. We develop the local nonsmooth optimization strategy, a superlinear space decomposition algorithm, which is suited for optimizing the H_∞-norm. Problem (7) implies the smoothness information; we can adopt variable space decomposition form. Meanwhile, since the problem (7) has the special structure called primal-dual gradient structure (PDG), which has been introduced in [15], it is possible to identify smooth tracks. So we can design a method which has fast convergent rate. The approach which is taken to solve this problem is based on using the recently developed local optimization algorithm presented in [15, 16]. In light of the $\mathcal{UV}$-space decomposition, this method is introduced in [15] (see also [17, 18]). Moreover, it is applied to many applications such as nonlinear programming and second-order cone programming (see [19–22]). The idea is to decompose R^n into two orthogonal subspaces $\mathcal{V}$ and $\mathcal{U}$ at a point $\overline{x}$ that the nonsmoothness of f is concentrated essentially on $\mathcal{V}$, and the smoothness of f appears on $\mathcal{U}$-subspace. More precisely, for a given $\overline{g} \in \partial f(\overline{x})$, where $\partial f(\overline{x})$ denotes the Clarke subdifferential of f at $\overline{x}$. Then R^n can be decomposed as direct sum of two orthogonal subspaces, that is, $R^n = \mathcal{U} \oplus \mathcal{V}$, where $\mathcal{V} = \mathrm{lin}(\partial f(\overline{x}) - \overline{g})$, and $\mathcal{U} = \mathcal{V}^\perp$. Then we define the primal-dual Lagrangian, an approximation of the original function, and show along certain manifolds it can be used to create as second-order expansion for a nondifferentiable function. As a result, we can design an algorithm frame that makes a step in the $\mathcal{V}$-space, followed by a $\mathcal{U}$-Newton step in order to obtain superlinear convergence, and show that it improves the situation considerably.

The rest of the paper is organized as follows. In Section 2, we recall some basic concepts about $\mathcal{UV}$ decomposition theory. In Section 3, we reformulate these problems as unconstrained max-finite function optimization problems under the hypothesis of the multiplicity one of the largest eigenvalue. We also mention some of the issues involved in trying

to solve such problems. Using primal-dual gradient structure (PDG), we give an important conclusion about second-order expansion of the function. Likewise, Section 4 outlines the optimization approach as in Section 3 and presents a different way to deal with the supposition of multiplying largest eigenvalues. The paper ends with some concluding remarks.

2. Preparation and Preliminary Results

We recall the $\mathcal{UV}$-theory developed in [15]. Let $f : R^m \to R$ be a finite-valued convex function. For a given $x \in R^m$, we start by defining a decomposition of the space $R^m = \mathcal{U}(x) \oplus \mathcal{V}(x)$. The subspaces $\mathcal{U}(x)$ and $\mathcal{V}(x)$ are equivalently defined as follows:

$$\mathcal{U}(x) := \left\{d \in R^m : f'(x;d) = -f'(x;-d)\right\}, \qquad \mathcal{V}(x) := \mathcal{U}(x)^{\perp}. \tag{8}$$

In other words, $\mathcal{U}$ is the subspace where $f(x+\cdot)$ appears to be differentiable at 0. Likewise, we can obtain the following result, which is stated in [15].

Proposition 1. *Let $f : R^m \to R$ be a proper convex function for a given point x; one has the following.*

(1) *$\mathcal{V}(x)$ is the subspace parallel to aff $\partial f(x)$ and $\mathcal{U}(x) = \mathcal{V}(x)^{\perp}$.*

(2) *For any $g \in ri\ \partial f(x)$, $\mathcal{U}(x)$ and $\mathcal{V}(x)$ are, respectively, the normal and tangent cones to $\partial f(x)$ at g, where ri C stands for the relative interior respect to a given set C.*

We give the Clarke generalized gradient for local Lipschitz function.

Definition 2 (see [23, 24]). Let f be local Lipschitz on R^n; the *generalized gradient* of f at x, denoted by $\partial f(x)$, is defined by

$$\partial f(x) := \{\xi \in R^n : f^{\circ}(x;d) \geq \langle \xi, d\rangle, \forall d \in R^n\}, \tag{9}$$

where $f^{\circ}(x;d) = \limsup_{y \to x, t \to 0^+} (f(y+td) - f(y))/t$ is the *generalized directional derivative* of f at x in the directive d.

The following results come from [23]; we will use these properties in later sections; as for their proofs, we will omit them.

Proposition 3. *Suppose $\{f_i\}$ is a finite collection of functions ($i = 1, \ldots, n$), each of which is Lipschitz near x. The function f is defined by*

$$f(x) = \max_{i=1,\ldots,n} f_i(x). \tag{10}$$

Then one has

$$\partial f(x) \subset \operatorname{conv}\{\partial f_i(x) : i \in I(x)\}, \tag{11}$$

where $I(x) := \{i : f_i(x) = f(x)\}$, and if f_i is regular at x for each $i \in I(x)$, then equality holds and f is regular at x.

Notation. We introduce the basic notation in the remainder parts. S_n is the space of $n\times n$ symmetric matrices and S_n^+ stands for the cone of $n\times n$ positive semidefinite symmetric matrices. $A \cdot B := \operatorname{tr} AB$ denotes Fröbenius scalar product of $A, B \in S_n$. Let $p \geq 1$ be the multiplicity of the largest eigenvalue $\lambda_1(A)$ of A; that is, A lies on the submanifold $\mathcal{M}_p := \{A \in S_n : \lambda_1(A) = \cdots = \lambda_p(A) > \lambda_{p+1}(A)\}$, where $\mathcal{M}_p$ is a C^{∞}-submanifold of S_n. Let $E_1(A)$ be the eigenspace associated with λ_1, let $Q_1(A)$ be an orthonormal basis of $E_1(A)$, let $P_1(A)$ be an orthonormal basis associated with $\lambda_1, \ldots, \lambda_p$, $T_{\mathcal{M}}(A)$, and let $N_{\mathcal{M}}(A)$ be, respectively, the tangent and normal spaces to the submanifold $\mathcal{M}$ at $A \in \mathcal{M}$. $\mathcal{A}^* : S_n \to R^m$ is the adjoint operator of the linear operator $\mathcal{A} : R^m \to S_n$. Much of the additional notation comes from [25, 26].

3. $\mathcal{UV}$-Space Decomposition for Single Eigenvalue

3.1. $\mathcal{UV}$-Theory of the Single Eigenvalue Function. In this section we will analyse the case where the multiplicity of $\lambda_1(A(x,\omega))$ is one at all active frequencies ω. This is motivated by practical considerations because nonsmoothness (b) never occurred in our tests. The necessary changes required for the general case will be discussed in next section.

Lemma 4. *For a closed-loop stabilizing controller x, the set of active frequencies $\Omega(x) := \{\omega \in [0,\infty] : f(x) = \lambda_1(A(x,\omega))\}$ is either finite or $\Omega(x) = [0,\infty]$; that is, $f(x) = \lambda_1(A(x,\omega))$ for all ω.*

A system where $\Omega(x) = [0,\infty]$ is called *all-pass*. This is rarely encountered in practice. For the technical formulas we will concentrate on those x's, where the set of active frequencies or peaks $\Omega(x) = \{\omega \in [0,\infty] : f(x) = \lambda_1(A(x,\omega))\}$ is finite.

From what follows we will analyse the case where the multiplicity of $f(x,\omega) := \lambda_1(A(x,\omega))$ is one at all active frequencies ω. This is motivated by practical considerations because nonsmoothness about λ_1 never happened in our tests. The necessary changes required for the general case will be discussed in Section 4.

In [27], three approaches to semi-infinite programming are discussed: exchange of constraints, discretization, and local reduction. We will use a local reduction method here. The main ideas are recalled below.

Let $\overline{x}$ be a local solution of (7). Indexing the active frequencies $\Omega(x) := \{\overline{\omega}_1, \ldots, \overline{\omega}_p\}$ at $\overline{x}$, we suppose that the following conditions are satisfied.

Assumption 5. Consider

(i) $f'_{\omega}(\overline{x}, \overline{\omega}_i) = 0$, $i = 1, \ldots, p$,

(ii) $f''_{\omega\omega}(\overline{x}, \overline{\omega}_i) < 0$, $i = 1, \ldots, p$,

(iii) $f(\overline{x}, \omega) < f(\overline{x})$, for every $\omega \notin \Omega(\overline{x}) = \{\overline{\omega}_1, \ldots, \overline{\omega}_p\}$.

These assumptions define the setting denoted as the standard case in semi-infinite programming [27]. The three conditions express the fact that the frequencies $\overline{\omega}_i \in \Omega(\overline{x})$ are the strict global maximizers of $f(\overline{x}, \cdot)$. Notice that condition (iii) is the finiteness hypothesis already mentioned, justified by Lemma 4.

Lemma 6. *Under conditions (i)–(iii), the neighborhood U of $\overline{x}$ may be chosen such that* $\max_{\omega\in[0,\infty]} f(x,\omega) = \max_{i=1,\dots,p} f(x,\omega_i(x))$ *for every $x \in U$. In particular, $\Omega(x) \subset \{\omega_1(x),\dots,\omega_p(x)\}$ for every $x \in U$.*

So we have that program (6) is locally equivalent to the standard following nonlinear program:

$$\min_{x\in R^n} \sup_{i=1,\dots,p} f(x,\omega_i(x)), \tag{12}$$

where $f(x,\omega_i(x)) = \lambda_1(A(x,\omega_i(x)))$; then we may solve (12) via the so-called $\mathcal{UV}$-decomposition method.

Assumption 7. $f'_x(\overline{x},\overline{\omega}_1),\dots,f'_x(\overline{x},\overline{\omega}_p)$ are linearly independent.

Under the hypothesis of Assumption 7, local convergence of this approach will be assured because this guarantees that (12) satisfies the linear independence constraint qualification hypothesis.

We denote $F(x) := \sup_{i=1,\dots,p}\lambda_1(A(x,\omega_i(x)))$, and $q_1^i(x)$, $i = 1,\dots,p$, stands for the eigenvector associated with the largest eigenvalue of $A(x,\omega_i(x))$.

Next a special kind of structure of $F(x)$, called primal-dual structure (PDG), will be seen.

Proposition 8. *There exists a ball about $\overline{x}$, denoted by $B(\overline{x})$, p functions*

$$f_i(x) := f(x,\omega_i(x)) = \lambda_1(A(x,\omega_i(x))), \quad \text{for } i = 1,\dots,p; \tag{13}$$

the multiplicity of $\lambda_1(A(x,\omega_i(x)))$ is single, so $f_i(x)$ are C^∞ on $B(\overline{x})$; in addition,

(1) *$\overline{x} \in B(\overline{x})$ and $f_i(\overline{x}) = F(\overline{x})$ for $i = 1,\dots,p$;*

(2) *for each $x \in B(\overline{x})$, $F(x) = \max_{i=1,\dots,p} f_i(x)$;*

(3) *$\triangle_1$ is the unit simplex in R^p given by*

$$\triangle_1 := \left\{(\alpha_1,\alpha_2,\dots,\alpha_p) : \sum_{i=1}^p \alpha_i = 1, \alpha_i \ge 0, i = 1,\dots,p\right\}; \tag{14}$$

(4) *on the basis of the property about the subdifferential of the maximum functions, for each $x \in B(\overline{x})$, $g \in \partial F(x)$ if and only if*

$$g = \sum_{i=1}^p \alpha_i \nabla f_i(x), \tag{15}$$

where $\alpha_i = 0$ if $f_i(x) < F(x)$ and $\alpha := (\alpha_1,\dots,\alpha_p) \in \triangle_1$.

We have the following result.

Theorem 9. *Suppose the set $\Omega(x)$ is finite. Then the Clarke subdifferential of F at x is the set as follows:*

$$\begin{aligned}
&\partial F(x) \\
&\quad = \Big\{G \mid G \\
&\qquad = \textstyle\sum_{i\in I(x)} \alpha_i \nabla f(x,\omega_i(x)) \\
&\qquad = \textstyle\sum_{i\in I(x)} \alpha_i \Big(A'_x(x,\omega_i(x)) \\
&\qquad\qquad + A'_\omega(x,\omega_i(x))\,\omega'_i(x)\Big)^* \\
&\qquad\qquad \Big(q_1^i(x)\left(q_1^i(x)\right)^T\Big), \alpha \in \Delta_{|I(x)|}\Big\},
\end{aligned} \tag{16}$$

where

$$I(x) = \{i \in \{1,\dots,p\} \mid F(x) = f(x,\omega_i(x))\} \tag{17}$$

is the set of active indices at x and

$$\Delta_s = \left\{\alpha \in R^s \mid \alpha_i \ge 0, \sum_{i=1}^s \alpha_i = 1\right\}. \tag{18}$$

Proof. Because $F(x)$ is the finite maximum functions, we can directly make use of the Clarke subdifferential of it and derivative of the eigenvalue function with multiplicity one to get

$$\begin{aligned}
&\partial F(x) \\
&\quad = \sum_{i\in I(x)} \alpha_i \nabla_x (\lambda_1(A(x,\omega_i(x)))) \\
&\quad = \sum_{i\in I(x)} \alpha_i \left(A'\right)^* \left(q_1^i(x)\left(q_1^i(x)\right)^T\right) \\
&\quad = \sum_{i\in I(x)} \alpha_i \Big(A'_x(x,\omega_i(x)) \\
&\qquad\qquad + A'_\omega(x,\omega_i(x))\,\omega'_i(x)\Big)^* \left(q_1^i(x)\left(q_1^i(x)\right)^T\right),
\end{aligned} \tag{19}$$

and the proof is done. □

Theorem 10. *Suppose Assumptions 5 and 7 hold. Then one has the following results at $\overline{x}$.*

(1) *The Clarke subdifferential of $F(\overline{x})$ has the following expression:*

$$\begin{aligned}
&\partial F(\overline{x}) \\
&\quad = \left\{g \mid g = \sum_{i=1}^p \alpha_i \left(A'_x(\overline{x},\overline{\omega}_i)\right)^* \left(q_1^i(\overline{x})\left(q_1^i(\overline{x})\right)^T\right), \alpha \in \Delta_1\right\}.
\end{aligned} \tag{20}$$

(2) *Let $\mathscr{V}$ denote the subspace generated by the subdifferential $\partial F(\overline{x})$. Then*

$$\mathscr{V} = \operatorname{lin}\{\nabla f_i(\overline{x}) - \nabla f_1(\overline{x})\}, \quad i = 2, \ldots, p,$$

$$\mathscr{U} = \{d \in R^n \mid \langle d, \nabla f_i(\overline{x}) - \nabla f_1(\overline{x})\rangle = 0,\ i = 2, \ldots, p\}, \tag{21}$$

where lin C *stands for linear hull of a set* C.

Proof. With Theorem 9 and Assumption 5, we can get the conclusion (1).

Let $\alpha_1 = 1, \alpha_i = 0,\ i \neq 1$, and we have $\nabla f_1(\overline{x}) \in \partial F(\overline{x})$. Then it follows from the definition of space $\mathscr{V}$ that

$$\begin{aligned}\mathscr{V} &= \operatorname{lin}(\partial F(\overline{x}) - \nabla f_1(\overline{x}))\\ &= \operatorname{lin}\{\nabla f_i(\overline{x}) - \nabla f_1(\overline{x})\},\end{aligned} \tag{22}$$

and $\mathscr{U} = \mathscr{V}^{\perp}$ means that the second formula holds. The proof is completed. □

Remark 11. (i) Since the subspaces $\mathscr{U}$ and $\mathscr{V}$ generate the whole space R^m, every vector can be decomposed along its $\mathscr{V}\mathscr{U}$-components at $\overline{x}$. In particular, any $x \in R^m$ can be expressed as follows:

$$R^n \ni x = \overline{x} + u \oplus v = \overline{x} + \overline{U}u + \overline{V}v, \tag{23}$$

where $\overline{V} = [\nabla f_i(\overline{x}) - \nabla f_1(\overline{x}), i = 2, \ldots, p]$ and $\overline{U} = \overline{V}^{\perp}$.

(ii) For any $\overline{s} \in \partial F(\overline{x})$, we have

$$\overline{s} = \overline{s}_{\mathscr{U}} \oplus \overline{s}_{\mathscr{V}} = \overline{U}^T\overline{s} + \overline{V}^T\overline{s}. \tag{24}$$

From the above Theorem 10, the $\mathscr{U}$-component of a subgradient $s \in \partial F(\overline{x})$ is the same as that of any other subgradient at $\overline{x}$; that is, $\overline{s}_{\mathscr{U}} = \overline{U}^T\overline{s}$.

3.2. Smooth Trajectory and Second-Order Properties. Given that $\overline{g} = \overline{g}_{\mathscr{U}} \oplus \overline{g}_{\mathscr{V}} \in \partial F(\overline{x})$, the Lagrangian-like function of F can be formulated in

$$L(u; \overline{g}_{\mathscr{V}}) = \inf_{v \in \mathscr{V}} \{F(\overline{x} + u \oplus v) - \langle \overline{g}_{\mathscr{V}}, v\rangle_{\mathscr{V}}\}. \tag{25}$$

Theorem 12. *Suppose Assumption 7 holds. Then, for all u small enough, there exists the following.*

(i) *The solution of the nonlinear system with variables $(u, v) \in \mathscr{U} \times \mathscr{V}$*

$$f_i(\overline{x} + u \oplus v) - f_1(\overline{x} + u \oplus v) = 0, \quad 1 \neq i, \tag{26}$$

is unique and $v = v(u)$, where $v(u) : \mathscr{U} \to \mathscr{V}$ is a C^2 function.

(ii) *For the C^2 solution function $v(u)$ in (i) one has*

$$Jv(u) = -\left(V(u)^T\overline{V}\right)^{-1}V(u)^T\overline{U}, \tag{27}$$

where $V(u) = [\{\nabla f_i(\mathscr{X}(u)) - \nabla f_1(\mathscr{X}(u))\}, 1 \neq i]$.

The trajectory $\mathscr{X}(u) = \overline{x} + u \oplus v(u)$ is C^2 and

$$J\mathscr{X}(u) = \overline{U} + \overline{V}Jv(u). \tag{28}$$

In particular, $\mathscr{X}(0) = \overline{x}$, $Jv(0) = 0$, and $J\mathscr{X}(0) = \overline{U}$.

(iii) $F(\mathscr{X}(u)) = f_i(\mathscr{X}(u)) = f(\mathscr{X}(u), \omega_i(\mathscr{X}(u))), \omega_i \in \Omega(\overline{x})$.

Proof. (i) Differentiating the left hand side of (26) with respect to v gives

$$[\nabla f_i(\overline{x} + u \oplus v) - \nabla f_1(\overline{x} + u \oplus v)]^T\overline{V}, \quad 1 \neq i \in \{2, \ldots, p\}. \tag{29}$$

This Jacobian at $(u, v) = (0, 0)$ is $\overline{V}^T\overline{V}$, which is nonsingular because of Assumption 7. There is also a Jacobian with respect to u, so by the implicit function theorem, there is a C^1 function $v(u)$ defined on a neighborhood of $u = 0$ such that $v(0) = 0$.

(ii) From (i), we have that $v(u)$ is C^1. Thus, the Jacobian $Jv(u)$ and $J\mathscr{X}(u)$ exist and are continuous. Differentiating the system $f_i(\mathscr{X}(u)) - f_1(\mathscr{X}(u))$ with respect to u, we obtain that

$$\begin{aligned}[\nabla f_i(\mathscr{X}(u)) - \nabla f_1(\mathscr{X}(u))]^T J\mathscr{X}(u) = 0,\\ \omega_i(\mathscr{X}(u)) \in \Omega(\overline{x}),\end{aligned} \tag{30}$$

or, in matrix form, $V(u)J\mathscr{X}(u) = 0$. Using the expression $J\mathscr{X}(u) = \overline{U} + \overline{V}Jv(u)$, we have that

$$V(u)^T\left(\overline{U} + \overline{V}Jv(u)\right) = 0. \tag{31}$$

By virtue of the continuity of $V(u)$, $V(u)^T\overline{V}$ is nonsingular. Hence,

$$Jv(u) = -\left(V(u)^T\overline{V}\right)^{-1}V(u)^T\overline{U}. \tag{32}$$

Furthermore, $V(u)$ is C^1 because $f_i(x)$, $i \in \{1, \ldots, p\}$, is C^2; then $Jv(u)$ is C^1. Thus, $\mathscr{X}(u)$ and $v(u)$ are C^2. From the definition of the $\mathscr{U}\mathscr{V}$ spaces, we have $\mathscr{V} \perp \mathscr{U}$. Hence, $\overline{V}^T\overline{U} = 0$. So $Jv(0) = 0$ and $J\mathscr{X}(0) = \overline{U}$.

(iii) The conclusion can be directly obtained in terms of (i) and the definition of $\mathscr{X}(u)$. □

So far we have developed a primal track $\mathscr{X}(u)$. Now we take our attention to an associated dual object, which is also a smooth function of $u \in \mathscr{U}$; we study a multiplier vector function $\alpha(u)$, which depends on structure function gradients, $\mathscr{X}(u)$, and an arbitrary subgradient at $\overline{x}$.

Lemma 13. *Given that $\overline{g} \in \partial F(\overline{x})$, the system with $\{\alpha_i(u)\}$, $i \in \{1, \ldots, p\}$,*

$$\begin{aligned}\overline{V}^T\left[\sum_{i=1}^{p}\alpha_i(u)\nabla f_i(\mathscr{X}(u)) - \overline{g}\right] = 0,\\ \sum_{i=1}^{p}\alpha_i(u) = 1,\end{aligned} \tag{33}$$

has a unique solution $\alpha = \alpha(u)$, which is given by

$$\{\alpha_i(u)\}_{i\neq 1} = \left(\overline{V}^T V(u)\right)^{-1}\overline{V}\left(\overline{g} - \nabla f_1(\mathcal{X}(u))\right),$$
$$\alpha_1(u) = 1 - \sum_{i\neq 1}\alpha_i(u), \tag{34}$$

in particular, for all $i \in \{1,\ldots,p\}$, $\alpha_i(0) = \overline{\alpha}_i$.

Theorem 14. *Given that $\overline{g} \in \partial F(\overline{x})$, at the trajectory $\mathcal{X}(u) = \overline{x} + \overline{U}u + \overline{V}v(u)$, one has*

$$L(u;\overline{g}_{\mathcal{V}}) = f_i(\mathcal{X}(u)) - \langle \overline{g}_{\mathcal{V}}, v(u)\rangle_{\mathcal{V}}, \quad i \in \{1,\ldots,p\}. \tag{35}$$

Proof. According to the definition of L in (25) and the item (iii) from Theorem 12, we get

$$\begin{aligned} L(u;\overline{g}_{\mathcal{V}}) &= F(\mathcal{X}(u)) - \langle \overline{g}_{\mathcal{V}}, v(u)\rangle_{\mathcal{V}} \\ &= f_i(\mathcal{X}(u)) - \langle \overline{g}_{\mathcal{V}}, v(u)\rangle_{\mathcal{V}}. \end{aligned} \tag{36}$$

□

Theorem 15. *Given that $\overline{g} \in \partial F(\overline{x})$ and supposing Assumption 7 holds; then for u small enough, the following assertions are true.*

(i) *L is a C^2 function of u and satisfies the Lagrangian-like result $L(u;0) = f_i(\mathcal{X}(u))$, for $i = 1,\ldots,p$.*

(ii) *The gradient of L is given by $\nabla L(u;\overline{g}_{\mathcal{V}}) = \overline{U}^T(\sum_{i=1}^p \alpha_i(u)\nabla f_i(\mathcal{X}(u)))$, and, in particular, when $u = 0$, one has*

$$\nabla L(0;\overline{g}_{\mathcal{V}}) = \overline{U}^T\left(\sum_{i=1}^p \overline{\alpha}_i \nabla f_i(\overline{x})\right) = \overline{U}^T\overline{g}. \tag{37}$$

(iii) *The Hessian of L is given by*

$$\nabla^2 L(u;\overline{g}_{\mathcal{V}}) = J(\mathcal{X}(u))^T\sum_{i=1}^p \alpha_i(u)\nabla^2 f_i((\mathcal{X}(u)))J(\mathcal{X}(u)). \tag{38}$$

In particular, when $u = 0$, one has

$$\nabla^2 L(0;\overline{g}_{\mathcal{V}}) = \overline{U}^T\sum_{i=1}^p \overline{\alpha}_i \nabla^2 f_i(\overline{x})\overline{U}. \tag{39}$$

Proof. (i) This conclusion follows from C^2 of f_i.

(ii) Using the chain rule, the differential of the Lagrangian-like with respect to u can be written as follows:

$$\nabla L(u;\overline{g}_{\mathcal{V}}) = J(\mathcal{X}(u))^T\nabla f_i(\mathcal{X}(u)) - Jv(u)^T\overline{V}^T\overline{g}. \tag{40}$$

Multiplying each equation by the appropriate $\alpha_i(u)$, summing the results, and using the fact that $\sum_i \alpha_i(u) = 1$ yield

$$\nabla L(u;\overline{g}_{\mathcal{V}}) = J(\mathcal{X}(u))^T\sum_i \alpha_i(u)\left(\nabla f_i(\mathcal{X}(u))\right) - Jv(u)^T\overline{V}^T\overline{g}. \tag{41}$$

Using the transpose of the expression of $J(\mathcal{X}(u))$, we get

$$\begin{aligned} \nabla L(u;\overline{g}_{\mathcal{V}}) &= \overline{U}^T\sum_i \alpha_i(u)\left(\nabla f_i(\mathcal{X}(u))\right) \\ &\quad + Jv(u)^T\overline{V}^T\left(\sum_i \alpha_i(u)\left(\nabla f_i(\mathcal{X}(u))\right) - \overline{g}\right), \end{aligned} \tag{42}$$

which together with (6.11) in [28] yields the desired result.

If $u = 0$, then $v(0) = 0$, and $\mathcal{X}(0) = \overline{x}$. By Remark 11(ii), we have

$$\nabla L_{\mathcal{U}}(0;\overline{g}_{\mathcal{V}}) = \overline{U}^T\sum_{i=1}^p \overline{\alpha}_i \nabla f_i(\overline{x}) = \overline{U}^T\overline{g}. \tag{43}$$

(iii) Differentiating the following equation with respect to u;

$$\nabla L(u;\overline{g}_{\mathcal{V}}) = \overline{U}^T\left(\sum_{i=1}^p \alpha_i(u)\nabla f_i(\mathcal{X}(u))\right), \tag{44}$$

we obtain

$$\begin{aligned} \nabla^2 L(u;\overline{g}_{\mathcal{V}}) &= \overline{U}^T M(u) J(\mathcal{X}(u)) \\ &\quad + \overline{U}^T\sum_{i=1}^p \alpha_i(u)\nabla f_i(\mathcal{X}(u)) J\alpha_i(u), \end{aligned} \tag{45}$$

where $M(u) = \sum_{i=1}^p \alpha_i(u)\nabla^2 f_i(\mathcal{X}(u), \omega_i(\mathcal{X}(u)))$. It follows from the proof of Theorem 6.3 in [28] that

$$\begin{aligned} &\sum_{i=1}^p \alpha_i(u)\nabla f_i(\mathcal{X}(u)) J\alpha_i(u) \\ &\quad = -V(u)\left(\overline{V}V(u)\right)^{-1}\overline{V}^T M(u) J(\mathcal{X}(u)). \end{aligned} \tag{46}$$

Then

$$\begin{aligned} \nabla^2 L(u;\overline{g}_{\mathcal{V}}) &= \overline{U}^T M(u) J(\mathcal{X}(u)) \\ &\quad - \overline{U}^T V(u)\left(\overline{V}V(u)\right)^{-1}\overline{V}^T M(u) J(\mathcal{X}(u)) \\ &= \overline{U}^T M(u) J(\mathcal{X}(u)) \\ &\quad + Jv(u)^T\overline{V}^T M(u) J(\mathcal{X}(u)) \\ &= \left[\overline{U}^T + Jv(u)^T\overline{V}^T\right] M(u) J(\mathcal{X}(u)) \\ &= J\mathcal{X}(u)^T M(u) J(\mathcal{X}(u)), \end{aligned} \tag{47}$$

when $u = 0$,

$$\nabla^2 L(0;\overline{g}_{\mathcal{V}}) = \overline{U}^T M(0)\overline{U} = \overline{U}^T\sum_{i=1}^p \overline{\alpha}_i\left(\nabla^2 f_i(\overline{x})\right)\overline{U}. \tag{48}$$

We finish the proof. □

Theorem 16. *Suppose Assumption 7 holds and $\overline{g} \in \partial F(\overline{x})$. Then for u small enough, there holds the second-order expansion of f along with the trajectory $\mathcal{X}(u) = \overline{x} + u \oplus v(u)$,*

$$F(\mathcal{X}(u)) = F(\overline{x}) + \langle \overline{g}, u \oplus v(u) \rangle + \frac{1}{2} u^T \nabla L^2 (0; \overline{g}_{\mathcal{V}}) u + o\left(\|u\|_{\mathcal{U}}^2\right). \tag{49}$$

Proof. From the definition of L, we have

$$L(u; \overline{g}_{\mathcal{V}}) = F(\mathcal{X}(u)) - \langle \overline{g}_{\mathcal{V}}, v(u) \rangle_{\mathcal{V}}. \tag{50}$$

Since $L \in C^2$, we get

$$\begin{aligned} L(u; \overline{g}_{\mathcal{V}}) &= L(0; \overline{g}_{\mathcal{V}}) + \langle \nabla L(0; \overline{g}_{\mathcal{V}}), u \rangle_{\mathcal{U}} + \frac{1}{2} u^T \nabla^2 L(0; \overline{g}_{\mathcal{V}}) u + o\left(\|u\|_{\mathcal{U}}^2\right) \\ &= F(\overline{x}) + \langle \overline{g}_{\mathcal{U}}, u \rangle_{\mathcal{U}} + \frac{1}{2} u^T \nabla^2 L(0; \overline{g}_{\mathcal{V}}) u + o\left(\|u\|_{\mathcal{U}}^2\right). \end{aligned} \tag{51}$$

Therefore,

$$\begin{aligned} F(\mathcal{X}(u)) &= F(\overline{x}) + \langle \overline{g}_{\mathcal{U}}, u \rangle_{\mathcal{U}} + \langle \overline{g}_{\mathcal{V}}, v(u) \rangle_{\mathcal{V}} + \frac{1}{2} u^T \nabla^2 L(0; \overline{g}_{\mathcal{V}}) u + o\left(\|u\|_{\mathcal{U}}^2\right) \\ &= F(\overline{x}) + \langle \overline{g}, u \oplus v(u) \rangle + \frac{1}{2} u^T \nabla^2 L(0; \overline{g}_{\mathcal{V}}) u + o\left(\|u\|_{\mathcal{U}}^2\right). \end{aligned} \tag{52}$$

□

4. $\mathcal{UV}$-Decomposition for Multiple Eigenvalues

4.1. $\mathcal{UV}$-Theory of the Multiple Eigenvalue Function. The working hypothesis of the previous section was that leading eigenvalues $\lambda_1(A(x, \omega_i(x)))$ had multiplicity 1 for all frequencies in the set $\{\omega_1(x), \ldots, \omega_p(x)\}$ and for all x in a neighborhood of $\overline{x}$. This hypothesis is motivated by our numerical experience, where we have never encountered multiple eigenvalues. This is clearly in contrast with experience in pure eigenvalue optimization problems. However, our approach is still functional if the hypothesis of single eigenvalues at active frequencies is abandoned. Based on the weaker assumption that the eigenvalue multiplicities $\overline{r}_i$ at the limit point $\overline{x}$ are known for all active frequencies $\overline{\omega}_i$, $i = 1, \ldots, p$, and on the information at the current iterate point, we have good technique to dependably guess $\overline{r}_i$.

This situation has been discussed by several authors (see, e.g., [27, 29–31]). Consider $X \in S_n$ where $\lambda_1(X)$ has multiplicity r. We replace the maximum eigenvalue function λ_1 by the average of the first r eigenvalues

$$\widehat{\lambda}_r(X) = \frac{1}{r} \sum_{j=1}^{r} \lambda_j(X). \tag{53}$$

This function is smooth and convex in a neighborhood of the smooth manifold

$$\mathcal{M}_r = \{X \in S_n : \lambda_1(X) = \cdots = \lambda_r(X) > \lambda_{r+1}(X)\} \tag{54}$$

of the matrices $X \in S_n$ with the largest eigenvalue multiplicity r, and $\lambda_1 = \widehat{\lambda}_r$ on $\mathcal{M}_r$. Then we may replace the nonsmooth information contained in λ_1 by the smooth information contained in the function $\widehat{\lambda}_r$ by adding the constraint $A(x, \omega_i(x)) \in \mathcal{M}_{\overline{r}_i}$. The manifold has codimension $d := ((r(r+1))/2) - 1$ in S_n and in a neighborhood of $\overline{X}$ may be described by d equations $h_1(X) = 0, \ldots, h_d(X) = 0$, which has been presented independently in [16, 32]. The extension to the semi-infinite eigenvalue optimization is clear under the finiteness assumption (iii). We may then approach minimization of the H_∞-norm along the same lines and obtain the finite program

$$\begin{aligned} \min_x \max_{i=1,\ldots,p} \quad & \widehat{\lambda}_{\overline{r}_i}(A(x, \omega_i(x))) \\ \text{s.t.} \quad & A(x, \omega_i(x)) \in \mathcal{M}_{\overline{r}_i}, \end{aligned} \tag{55}$$

where $\overline{r}_i$ stands for the multiplicity of the largest eigenvalue $\lambda_1(A(\overline{x}, \overline{\omega}_i))$; we denote $Q_{\overline{r}_i} Q_{\overline{r}_i}{}^T$ by an orthonor-mal basis associated with the eigenvector of $\lambda_1(A(\overline{x}, \overline{\omega}_i))$. According to the foregoing analysis, we can transform the above constrained optimization problem into the following form:

$$\begin{aligned} \min_x \max_{i=1,\ldots,p} \quad & \widehat{\lambda}_{\overline{r}_i}(A(x, \omega_i(x))) \\ \text{s.t.} \quad & h_1(A(x, \omega_1(x))) = 0 \\ & \vdots \\ & h_{\overline{d}_1}(A(x, \omega_1(x))) = 0 \\ & h_{\overline{d}_1+1}(A(x, \omega_2(x))) = 0 \\ & \vdots \\ & h_{\overline{d}_1+\overline{d}_2}(A(x, \omega_2(x))) = 0 \\ & \vdots \\ & h_{\overline{d}_1+\cdots+\overline{d}_{p-1}+1}\left(A\left(x, \omega_p(x)\right)\right) = 0 \\ & \vdots \\ & h_{\overline{d}_1+\cdots+\overline{d}_p}\left(A\left(x, \omega_p(x)\right)\right) = 0. \end{aligned} \tag{56}$$

For convenience, we denote

$$\begin{aligned} f_1(x) &= \widehat{\lambda}_{\overline{r}_1}(A(x, \omega_1(x))), \\ &\vdots \\ f_p(x) &= \widehat{\lambda}_{\overline{r}_p}\left(A\left(x, \omega_p(x)\right)\right), \end{aligned}$$

$$\begin{aligned}
\phi_1(x) &= h_1(A(x,\omega_1(x))),\\
&\vdots\\
\phi_{\overline{d}_1}(x) &= h_{\overline{d}_1}(A(x,\omega_1(x))),\\
&\vdots\\
\phi_{\overline{d}_1+\cdots+\overline{d}_{p-1}+1}(x) &= h_{\overline{d}_1+\cdots+\overline{d}_{p-1}+1}\left(A\left(x,\omega_p(x)\right)\right),\\
&\vdots\\
\phi_{\overline{d}_1+\cdots+\overline{d}_p}(x) &= h_{\overline{d}_1+\cdots+\overline{d}_p}\left(A\left(x,\omega_p(x)\right)\right).
\end{aligned} \tag{57}$$

So $\widehat{F}(x) = \max_{i=1,\ldots,p}\widehat{\lambda}_{\overline{r}_i}(A(x,\omega_i(x))) = \max_{i=1,\ldots,p} f_i(x)$ because $\lambda_1(A(x,\omega_i(x))) = \widehat{\lambda}_{\overline{r}_i}(A(x,\omega_i(x)))$; we still label $f(x,\omega_i(x))$ as $\widehat{\lambda}_{\overline{r}_i}(A(x,\omega_i(x)))$. In view of the property about the indicated functions, we can transform (56) into an unconstrained optimization. Thus, for the problem (56), we just need to deal with the following form equivalently:

$$\min_x \widehat{F}(x) := \max_{i=1,\ldots,p} f_i(x) + I_{\{0\}}(\phi_1(x)) + \cdots + I_{\{0\}}\left(\phi_{N_p}(x)\right), \tag{58}$$

where $N_p := \overline{d}_1 + \cdots + \overline{d}_p$, and $I_C(y)$ means the indicated function at y for the set C.

Similarly as in Proposition 8, we can obtain that the problem (58) possesses PDG.

Proposition 17. *$\widehat{F}(x)$ in (58) is a primal-dual gradient structured (PDG) function.*

Proof. First, $f_i(x) = \widehat{\lambda}_{\overline{r}_i}(A(x,\omega_i(x))) = (1/\overline{r}_i)\sum_{j=1}^{\overline{r}_i}\lambda_j(A(x,\omega_i(x)))$; in [32] Shapiro and Fan had given the fact that the function $\sum_{j=1}^{\overline{r}_i}\lambda_j(A(x,\omega_i(x)))$, $h_l(A(x,\omega_i(x))) \in C^\infty$, for $i = 1,\ldots,p$ and $l = 1,\ldots,N_p$, so $f_i(x) \in C^\infty$ too. In this way, there exists a ball about $\overline{x}$, $B(\overline{x})$ and a dual multiplier set

(1) $\overline{x} \in \mathscr{P} := \{x \in B(\overline{x}) : \phi_l(x) = 0 \text{ for } l = 1,\ldots,N_p\}$ and $f_i(\overline{x}) = \widehat{F}(\overline{x})$ for $i = 1,\ldots,p$;

(2) for each $x \in \mathscr{P}$, $\widehat{F}(x) = \max_{i=1,\ldots,p} f_i(x)$;

(3) $\triangle$ is a closed convex set such that

(a) if $\gamma := (\alpha_1,\ldots,\alpha_p,\beta_1,\ldots,\beta_{N_p}) \in \triangle$, then we know that $\widehat{F}(x) = \sum_{i=1}^p \alpha_i f_i(x)$; that is, $(\alpha_1,\ldots,\alpha_p) \in \triangle_1$, which is defined in (14);

(b) for each $i = 1,\ldots,p$, $\mathbf{1}_i \in \triangle$, where $\mathbf{1}_j$ is the jth unit vector in R^{p+N_p};

(c) for each $l = 1,2,\ldots,N_p$, there exists $\gamma^l \in \triangle$ such that $\beta_l^l \neq 0$ and $\beta_i^l = 0$ for $i \in \{1,\ldots,N_p\} \setminus \{l\}$;

(4) for each $x \in B(\overline{x})$, $g \in \partial\widehat{F}(x)$ if and only if

$$g = \sum_{i=1}^p \alpha_i \nabla f_i(x) + \sum_{j=1}^{N_p} \beta_j \nabla\phi_j(x), \tag{59}$$

where $\alpha_i = 0$ if $f_i(x) < \widehat{F}(x)$ and $\alpha := (\alpha_1,\ldots,\alpha_p) \in \triangle_1$. Moreover, γ satisfies dual feasibility: $\gamma = (\alpha_1,\ldots,\alpha_p,\beta_1,\ldots,\beta_{N_p}) \in \triangle$. □

Assumption 18. $\{\nabla f_i(\overline{x}) - \nabla f_1(\overline{x})\}_{i=2}^p \bigcup \{\nabla\phi_j(\overline{x})\}_{j=1}^{N_p}$ are linearly independent.

Theorem 19. *Suppose Assumption 18 holds. Then one has the following results at $\overline{x}$.*

(1) *The Clarke subdifferential of $\widehat{F}(\overline{x})$ has the following expression:*

$$\partial\widehat{F}(\overline{x}) = \left\{g \mid g = \sum_{i=1}^p \alpha_i \nabla f_i(\overline{x}) + \sum_{j=1}^{N_p} \beta_j \nabla\phi_j(\overline{x}), \alpha \in \Delta_1\right\}, \tag{60}$$

where $N_p := \overline{d}_1 + \cdots + \overline{d}_p$.

(2) *Let $\mathscr{V}$ denote the subspace generated by the subdifferential $\partial F(\overline{x})$. Then*

$$\begin{aligned}
\mathscr{V} &= \operatorname{lin}\left(\{\nabla f_i(\overline{x}) - \nabla f_1(\overline{x})\}_{i=2}^p \bigcup \{\nabla\phi_j(x)\}_{j=1}^{N_p}\right)\\
\mathscr{U} &= \{d \in R^n \mid \langle d, \nabla f_i(\overline{x}) - \nabla f_1(\overline{x})\rangle = 0,\ i = 2,\ldots,p\},
\end{aligned} \tag{61}$$

where $\nabla f_i(\overline{x}) = (1/\overline{r}_i)(A'_x(\overline{x},\overline{\omega}_i))^*(Q_{\overline{r}_i}Q_{\overline{r}_i}{}^T)$,

$$\begin{aligned}
&\nabla f_i(\overline{x}) \in \partial\widehat{F}(\overline{x}) \quad \text{for } i = 1,\ldots,p;\\
&\nabla\phi_j(\overline{x}) \in \mathscr{V} \quad \text{for } j = 1,\ldots,N_p.
\end{aligned} \tag{62}$$

4.2. Smooth Trajectory and Second-Order Properties. We give the smooth trajectory information about the function $\widehat{F}(x)$ in the following theorem; with respect to its proofs, it is similar to Theorem 12.

Theorem 20. *Suppose Assumption 18 holds. Then for all u small enough, there exists the following.*

(i) *The nonlinear system with variables $(u,v) \in \mathscr{U} \times \mathscr{V}$*

$$\begin{aligned}
f_i(\overline{x} + u \oplus v) - f_1(\overline{x} + u \oplus v) &= 0, \quad i = 2,\ldots,p,\\
\phi_j(\overline{x} + u \oplus v) &= 0, \quad j = 1,\ldots,N_p,
\end{aligned} \tag{63}$$

has a unique solution $v = v(u)$, where $v : \mathscr{U} \to \mathscr{V}$ is a C^2-function.

(ii) *For C^2 function $v(u)$, one has*

$$Jv(u) = -\left(V(u)^T\overline{V}\right)^{-1} V(u)^T\overline{U}, \tag{64}$$

where $V(u) = [\{\nabla f_i(\mathcal{X}(u)) - \nabla f_1(\mathcal{X}(u))\}_{i=2}^{p} \bigcup \{\nabla \phi_j(\mathcal{X}(u))\}_{j=1}^{N_p}]$.

The trajectory $\mathcal{X}(u) = \overline{x} + u \oplus v(u)$ *is* C^2, *and*

$$J\mathcal{X}(u) = \overline{U} + \overline{V} J v(u). \tag{65}$$

In particular, $\mathcal{X}(0) = \overline{x}$, $Jv(0) = 0$, *and* $J\mathcal{X}(0) = \overline{U}$.

(iii) $v(u) = O(\|u\|^2)$, *and the smooth trajectory* $\mathcal{X}(u)$ *is tangent to* $\mathcal{U}$ *at* $\mathcal{X}(0) = \overline{x}$.

(iv) $\widehat{F}(\mathcal{X}(u)) = f_i(\mathcal{X}(u))$, $i = 1, \ldots, p$, *and* $\phi_j(\mathcal{X}(u)) = 0$, $j = 1, \ldots, N_p$.

Now we pay our attention to an associated dual object, that is, also a smooth function of $u \in \mathcal{U}$.

Theorem 21. *We suppose Assumption 18 is holding, with trajectory* $\mathcal{X}(u) = \overline{x} + \overline{U}u + \overline{V}v(u)$ *and to a subgradient* $\overline{g} \in \partial\widehat{F}(\overline{x})$ *at* $\overline{x}$, *for all u small enough; the linear system with variables* α_i, $i = 1, \ldots, p$; β_j, $j = 1, \ldots, N_p$,

$$\overline{V}^T \left[\sum_{i=1}^{p} \alpha_i \nabla f_i(\mathcal{X}(u)) + \sum_{j=1}^{N_p} \nabla \phi_j(\mathcal{X}(u)) \right] = \overline{V}^T \overline{g}, \qquad \sum_{i=1}^{p} \alpha_i = 1, \tag{66}$$

has a unique solution $\gamma(u) = (\alpha(u), \beta(u))$, *given by*

$$\{\alpha_i(u)\}_{i=2}^{p}, \{\beta_j(u)\}_{j=1}^{N_p} = \left(\overline{V}^T V(u)\right)^{-1} \overline{V}(\overline{g} - \nabla f_0(\mathcal{X}(u))), \qquad \alpha_1(u) = 1 - \sum_{i=2}^{p} \alpha_i(u), \tag{67}$$

where $V(u)$ *is defined in Theorem 20.*

Next we consider the following primal-dual function

$$\mathcal{L}(u; z) := \widehat{F}(\mathcal{X}(u)) - z^T v(u). \tag{68}$$

Theorem 22. *Given* $\overline{g} \in \partial\widehat{F}(\overline{x})$, *at the trajectory* $\mathcal{X}(u) = \overline{x} + \overline{U}u + \overline{V}v(u)$, *one has*

$$\mathcal{L}(u; \overline{g}_{\mathcal{V}}) = f_i(\mathcal{X}(u)) - \langle \overline{g}_{\mathcal{V}}, v(u) \rangle_{\mathcal{V}}, \quad i \in \{1, \ldots, p\}. \tag{69}$$

Theorem 23. *Given* $\overline{g} \in \partial\widehat{F}(\overline{x})$ *and supposing that Assumption 18 holds, then for u small enough, the following assertions are true.*

(i) $\mathcal{L}$ *is a* C^2 *function of u and*

$$\mathcal{L}(u; 0) = \left[\sum_{i=1}^{p} \alpha_i(u) f_i(\mathcal{X}(u)) + \sum_{j=1}^{N_p} \beta_j(u) \phi_j(\mathcal{X}(u)) \right]. \tag{70}$$

(ii) *The gradient of* $\mathcal{L}$ *is given by* $\nabla\mathcal{L}(u; \overline{g}_{\mathcal{V}}) = \overline{U}^T g(u; \overline{g}_{\mathcal{V}})$, *where*

$$g(u; \overline{g}_{\mathcal{V}}) = \sum_{i=1}^{p} \alpha_i(u) \nabla f_i(\mathcal{X}(u)) + \sum_{j=1}^{N_p} \beta_j(u) \nabla \phi_j(\mathcal{X}(u)), \tag{71}$$

and, in particular, when $u = 0$, *one has*

$$\nabla\mathcal{L}(0; \overline{g}_{\mathcal{V}}) = \overline{U}^T \left[\sum_{i=1}^{p} \overline{\alpha}_i \nabla f_i(\overline{x}) + \sum_{j=1}^{N_p} \overline{\beta}_j \nabla \phi_j(\overline{x}) \right] = \overline{U}^T \overline{g} = \overline{g}_{\mathcal{U}}. \tag{72}$$

(iii) *The Hessian of* $\mathcal{L}$ *is given by*

$$\nabla^2\mathcal{L}(u; \overline{g}_{\mathcal{V}}) = J(\mathcal{X}(u))^T M(u, \overline{g}_{\mathcal{V}}) J(\mathcal{X}(u)), \tag{73}$$

where $M(u, \overline{g}_{\mathcal{V}})$ *is the* $n \times n$ *matrix function defined by*

$$M(u, \overline{g}_{\mathcal{V}}) := \sum_{i=1}^{p} \alpha_i(u) \nabla^2 f_i(\mathcal{X}(u)) + \sum_{j=1}^{N_p} \beta_j(u) \nabla^2 \phi_j(\mathcal{X}(u)). \tag{74}$$

In particular, when $u = 0$, *one has*

$$\nabla^2\mathcal{L}(0; \overline{g}_{\mathcal{V}}) = \overline{U}^T M(0; \overline{g}_{\mathcal{V}}) \overline{U}. \tag{75}$$

Proof. (i) Because $f_i(\mathcal{X}(u))$ is C^2, it follows from C^2 of $\mathcal{L}$ in the above theorem. At the same time, Assumption 18 holds; using (68) with $z = 0$ gives

$$\mathcal{L}(u; 0) = f_i(\mathcal{X}(u)), \quad i = 1, \ldots, p. \tag{76}$$

In addition,

$$\phi_j(\mathcal{X}(u)) = 0, \quad j = 1, \ldots, N_p. \tag{77}$$

Multiplying α_i and β_j, respectively, for the above equations and summing, we get the Lagrangian-like expression in item (i).

(ii) Using the chain rule, the differential of the Lagrangian-like functions (69) and (77) with respect to u can be written as follows:

$$\nabla\mathcal{L}(u; \overline{g}_{\mathcal{V}}) = J(\mathcal{X}(u))^T (\nabla f_i(\mathcal{X}(u))) - Jv(u)^T \overline{V}^T \overline{g}, \qquad J(\mathcal{X}(u))^T (\nabla \phi_j(\mathcal{X}(u))) = 0. \tag{78}$$

Multiplying each equation by the appropriate $\alpha_i(u)$ and $\beta_j(u)$, summing the results, and using the fact that $\sum_i \alpha_i(u) = 1$ yield

$$\begin{aligned}\nabla\mathscr{L}(u;\overline{g}_{\mathscr{V}}) &= J(\mathscr{X}(u))^T\left[\sum_{i=1}^{p}\alpha_i(u)\left(\nabla f_i(\mathscr{X}(u))\right)\right.\\ &\qquad\left.+\sum_{j=1}^{N_p}\beta_j(u)\left(\nabla\phi_j(\mathscr{X}(u))\right)\right]\\ &\quad - Jv(u)^T\overline{V}^T\overline{g}\\ &= J(\mathscr{X}(u))^T g(u;\overline{g}_{\mathscr{V}}) - Jv(u)^T\overline{V}^T\overline{g}.\end{aligned} \tag{79}$$

Using the transpose of the expression of $J(\mathscr{X}(u))$, we get

$$\nabla\mathscr{L}(u;\overline{g}_{\mathscr{V}}) = \overline{U}^T g(u;\overline{g}_{\mathscr{V}}) + Jv(u)^T\overline{V}^T\left(g(u;\overline{g}_{\mathscr{V}}) - \overline{g}\right), \tag{80}$$

which together with (6.11) in [28] yields the desired result.

If $u = 0$, then $v(0) = 0$, and $\mathscr{X}(0) = \overline{x}$, $J\mathscr{X}(0) = \overline{U}$, $\sum_{i=1}^{p}\alpha_i(0) = 1$, and by Theorem 19

$$\begin{aligned}&\text{either } \nabla f_i(\overline{x}) \in \partial\widehat{F}(\overline{x}) \quad \text{so } \overline{U}^T\nabla f_i(\overline{x}) = \overline{g}_{\mathscr{U}},\\ &\text{or } \nabla\phi_j(\overline{x}) \in \mathscr{V} \quad \text{so } \overline{U}^T\nabla\phi_j(\overline{x}) = 0;\end{aligned} \tag{81}$$

so we attain

$$\nabla\mathscr{L}(0;\overline{g}_{\mathscr{V}}) = \overline{U}^T\left[\sum_{i=1}^{p}\overline{\alpha}_i\nabla f_i(\overline{x}) + \sum_{j=1}^{N_p}\overline{\beta}_j\nabla\phi_j(\overline{x})\right] = \overline{U}^T\overline{g}. \tag{82}$$

(iii) Differentiating the following equation with respect to u,

$$\begin{aligned}\nabla\mathscr{L}(u;\overline{g}_{\mathscr{V}}) &= \overline{U}^T\left(\sum_{i=1}^{p}\alpha_i(u)\nabla f_i(\mathscr{X}(u))\right.\\ &\qquad\left.+\sum_{j=1}^{N_p}\beta_j(u)\nabla\phi_j(\mathscr{X}(u))\right),\end{aligned} \tag{83}$$

we obtain

$$\begin{aligned}\nabla^2\mathscr{L}(u;\overline{g}_{\mathscr{V}}) &= \overline{U}^T M(u) J(\mathscr{X}(u))\\ &\quad+\overline{U}^T\left[\sum_{i=1}^{p}\alpha_i(u)\nabla f_i(\mathscr{X}(u)) J\alpha_i(u)\right.\\ &\qquad\left.+\sum_{j=1}^{N_p}\beta_j(u)\nabla\phi_j(\mathscr{X}(u)) J\beta_j(u)\right],\end{aligned} \tag{84}$$

where $M(u) = \sum_{i=1}^{p}\alpha_i(u)\nabla^2 f_i(\mathscr{X}(u)) + \sum_{j=1}^{N_p}\beta_j(u)\nabla^2\phi_j(\mathscr{X}(u))$. It follows from the proof of Theorem 6.3 in [28] that

$$\begin{aligned}&\sum_{i=1}^{p}\alpha_i(u)\nabla f_i(\mathscr{X}(u)) J\alpha_i(u) + \sum_{j=1}^{N_p}\beta_j(u)\nabla\phi_j(\mathscr{X}(u)) J\beta_j(u)\\ &\quad= -V(u)\left(\overline{V}V(u)\right)^{-1}\overline{V}^T M(u) J(\mathscr{X}(u)).\end{aligned} \tag{85}$$

Then

$$\begin{aligned}&\nabla^2\mathscr{L}(u;\overline{g}_{\mathscr{V}})\\ &\quad= \overline{U}^T M(u) J(\mathscr{X}(u))\\ &\qquad- \overline{U}^T V(u)\left(\overline{V}V(u)\right)^{-1}\overline{V}^T M(u) J(\mathscr{X}(u))\\ &\quad= \overline{U}^T M(u) J(\mathscr{X}(u)) + Jv(u)^T\overline{V}^T M(u) J(\mathscr{X}(u))\\ &\quad= \left[\overline{U}^T + Jv(u)^T\overline{V}^T\right] M(u) J(\mathscr{X}(u))\\ &\quad= J(\mathscr{X}(u))^T M(u) J(\mathscr{X}(u)),\end{aligned} \tag{86}$$

when $u = 0$,

$$\nabla^2\mathscr{L}(0;\overline{g}_{\mathscr{V}}) = \overline{U}^T M(0)\overline{U}, \tag{87}$$

where $M(0;\overline{g}_{\mathscr{V}}) = \sum_{i=1}^{p}\overline{\alpha}_i\nabla^2 f_i(\overline{x}) + \sum_{j=1}^{N_p}\overline{\beta}_j\nabla^2\phi_j(\overline{x})$. □

We call the corresponding Hessian matrix of $\mathscr{L}$ at $u = 0$ a basic $\mathscr{U}$-Hessian for $\widehat{F}$ at $\overline{x}$ and denote it by $\overline{H} := \nabla^2\mathscr{L}(0;0)$. Using second-order $\mathscr{U}$-derivatives we can specify second-order expansions for $\widehat{F}$ and give related necessary conditions for optimization problem.

Theorem 24. *Suppose Assumption 18 holds and $\overline{g} \in \partial\widehat{F}(\overline{x})$. Then for u small enough, there holds the second-order expansion of f along the trajectory $\mathscr{X}(u) = \overline{x} + u \oplus v(u)$,*

$$\begin{aligned}\widehat{F}(\mathscr{X}(u)) &= \widehat{F}(\overline{x}) + \langle\overline{g}, u\oplus v(u)\rangle\\ &\quad+\frac{1}{2}u^T\nabla^2\mathscr{L}(0;\overline{g}_{\mathscr{V}})u + o\left(\|u\|_{\mathscr{U}}^2\right).\end{aligned} \tag{88}$$

Proof. From the definition of $\mathscr{L}$, we have

$$\mathscr{L}(u;\overline{g}_{\mathscr{V}}) = \widehat{F}(\mathscr{X}(u)) - \langle\overline{g}_{\mathscr{V}}, v(u)\rangle_{\mathscr{V}}. \tag{89}$$

Since $\mathscr{L} \in C^2$, we get

$$\begin{aligned}\mathscr{L}(u;\overline{g}_{\mathscr{V}}) &= \mathscr{L}(0;\overline{g}_{\mathscr{V}}) + \langle\nabla\mathscr{L}(0;\overline{g}_{\mathscr{V}}), u\rangle_{\mathscr{U}}\\ &\quad+\frac{1}{2}u^T\nabla^2\mathscr{L}(0;\overline{g}_{\mathscr{V}})u + o\left(\|u\|_{\mathscr{U}}^2\right)\\ &= \widehat{F}(\overline{x}) + \langle\overline{g}_{\mathscr{U}}, u\rangle_{\mathscr{U}}\\ &\quad+\frac{1}{2}u^T\nabla^2\mathscr{L}(0;\overline{g}_{\mathscr{V}})u + o\left(\|u\|_{\mathscr{U}}^2\right).\end{aligned} \tag{90}$$

Therefore,

$$\begin{aligned}\widehat{F}(\mathcal{X}(u)) &= \widehat{F}(\overline{x}) + \langle \overline{g}_{\mathcal{U}}, u\rangle_{\mathcal{U}} + \langle \overline{g}_{\mathcal{V}}, v(u)\rangle_{\mathcal{V}} \\ &\quad + \frac{1}{2}u^T\nabla^2\mathcal{L}(0;\overline{g}_{\mathcal{V}})\,u + o\left(\|u\|_{\mathcal{U}}^2\right) \\ &= \widehat{F}(\overline{x}) + \langle \overline{g}, u\oplus v(u)\rangle \\ &\quad + \frac{1}{2}u^T\nabla^2\mathcal{L}(0;\overline{g}_{\mathcal{V}})\,u + o\left(\|u\|_{\mathcal{U}}^2\right).\end{aligned} \tag{91}$$

□

Corollary 25. *Suppose Assumption 18 holds and $\overline{x}$ is a local minimizer of* (56). *Then $0 \in \partial\widehat{F}(\overline{x})$ and the associated basic $\mathcal{U}$-Hessian $\overline{H}$ is positive semidefinite.*

5. Conclusions

In this paper, we mainly study the $\mathcal{UV}$-theory to optimize the H_∞-norm or other nonsmooth criteria which are semi-infinite maxima of maximum eigenvalue functions. We use a methodology from semi-infinite programming to obtain a local nonlinear programming model and apply the $\mathcal{UV}$ decomposition method. With the so-called PDG that this problem possesses, Lagrangian-like theory is applied to the class of the functions. Under some hypothesis conditions, we can obtain the first- and second-order derivatives of the primal-dual Lagrangian function. This method can operate well in practice.

For further work, the need can be anticipated: in this paper we only give the theory analysis to solve the special class of eigenvalue optimization, we will continue to study its executable algorithm, and we will extend the $\mathcal{UV}$ algorithm of convex eigenvalues to nonconvex cases, where its related theory will be researched in later papers.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank Yuan Lu from Shenyang University for numerous fruitful discussions and an anonymous referee for their constructive criticism that helped to improve the presentation. This paper is supported by the National Natural Science Foundation of China, Grant nos. 11171049, 11226230, and 11301347, and General Project of the Education Department of Liaoning Province L2012427.

References

[1] P. Gahinet and P. Apkarian, "A linear matrix inequality approach to H_∞ control," *International Journal of Robust and Nonlinear Control*, vol. 4, no. 4, pp. 421–448, 1994.

[2] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, "State-space solutions to standard H_2 and H_∞ control problems," *IEEE Transactions on Automatic Control*, vol. 34, no. 8, pp. 831–847, 1989.

[3] V. Blondel and J. N. Tsitsiklis, "NP-hardness of some linear control design problems," *SIAM Journal on Control and Optimization*, vol. 35, no. 6, pp. 2118–2127, 1997.

[4] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice Hall, Upper Saddle River, NJ, USA, 1996.

[5] P. Apkarian and D. Noll, "Nonsmooth H_∞ synthesis," *IEEE Transactions on Automatic Control*, vol. 51, no. 1, pp. 71–86, 2006.

[6] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, vol. 15 of *SIAM Studies in Applied Mathematics*, SIAM, Philadelphia, Pa, USA, 1994.

[7] D. Q. Mayne and E. Polak, "Algorithms for the design of control systems subject to singular value inequalities," *Mathematical Programming Study*, vol. 18, pp. 112–134, 1982.

[8] E. Polak, "On the mathematical foundations of nondifferentiable optimization in engineering design," *SIAM Review*, vol. 29, no. 1, pp. 21–89, 1987.

[9] L. El Ghaoui, F. Oustry, and M. Ait Rami, "A cone complementarity linearization algorithm for static output-feedback and related problems," *IEEE Transactions on Automatic Control*, vol. 42, no. 8, pp. 1171–1176, 1997.

[10] K. M. Grigoriadis and R. E. Skelton, "Low-order control design for LMI problems using alternating projection methods," *Automatica*, vol. 32, no. 8, pp. 1117–1125, 1996.

[11] 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," *European Journal of Control*, vol. 10, no. 6, pp. 527–538, 2004.

[12] B. Fares, D. Noll, and P. Apkarian, "Robust control via sequential semidefinite programming," *SIAM Journal on Control and Optimization*, vol. 40, no. 6, pp. 1791–1820, 2002.

[13] J. B. Thevenet, D. Noll, and P. Apkarian, "Nonlinear spectral SDP method for BMI-constrained problems: applications to control design," in *Proceedings of the 1st International Conference on Informatics in Control, Automation and Robotics (ICINCO '04)*, Setubal, Portugal, 2004.

[14] V. Bompart, D. Noll, and P. Apkarian, "Second-order nonsmooth optimization for H_∞ synthesis," *Numerische Mathematik*, vol. 107, no. 3, pp. 433–454, 2007.

[15] C. Lemaréchal, F. Oustry, and C. Sagastizábal, "The U-Lagrangian of a convex function," *Transactions of the American Mathematical Society*, vol. 352, no. 2, pp. 711–729, 2000.

[16] F. Oustry, "The U-Lagrangian of the maximum eigenvalue function," *SIAM Journal on Optimization*, vol. 9, no. 2, pp. 526–549, 1999.

[17] R. Mifflin and C. Sagastizábal, "VU-smoothness and proximal point results for some nonconvex functions," *Optimization Methods & Software*, vol. 19, no. 5, pp. 463–478, 2004.

[18] C. Lemaréchal and C. Sagastizábal, "More than first-order developments of convex functions: primal-dual relations," *Journal of Convex Analysis*, vol. 3, no. 2, pp. 255–268, 1996.

[19] Y. Lu, L. P. Pang, and Z. Q. Xia, "A VU-decomposition method for a programming problem," *Applied Mathematics and Mechanics*, vol. 31, no. 2, pp. 263–270, 2010.

[20] Y. Lu, L. P. Pang, F. F. Guo, and Z. Q. Xia, "A superlinear space decomposition algorithm for constrained nonsmooth convex program," *Journal of Computational and Applied Mathematics*, vol. 234, no. 1, pp. 224–232, 2010.

[21] Y. Lu, L. P. Pang, X. J. Liang, and Z. Q. Xia, "An approximate decomposition algorithm for convex minimization," *Journal of Computational and Applied Mathematics*, vol. 234, no. 3, pp. 658–666, 2010.

[22] Y. Lu, L. P. Pang, J. Shen, and X. J. Liang, "A decomposition algorithm for convex nondierentiable minimization with errors," *Journal of Applied Mathematics*, vol. 2012, Article ID 215160, 15 pages, 2012.

[23] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley & Sons, New York, NY, USA, 1983.

[24] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, New York, NY, USA, 1998.

[25] J. B. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms*, Springer, Berlin, Germany, 1993.

[26] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, Germany, 1998.

[27] R. Hettich and K. O. Kortanek, "Semi-infinite programming: theory, methods, and applications," *SIAM Review*, vol. 35, no. 3, pp. 380–429, 1993.

[28] R. Mifflin and C. Sagastizábal, "On VU-theory for functions with primal-dual gradient structure," *SIAM Journal on Optimization*, vol. 11, no. 2, pp. 547–571, 2000.

[29] D. Noll and P. Apkarian, "Spectral bundle methods for nonconvex maximum eigenvalue functions: first-order methods," *Mathematical Programming*, vol. 104, no. 2-3, pp. 701–727, 2005.

[30] D. Noll and P. Apkarian, "Spectral bundle methods for nonconvex maximum eigenvalue functions: second-order methods," *Mathematical Programming*, vol. 104, no. 2-3, pp. 729–747, 2005.

[31] D. Noll, M. Torki, and P. Apkarian, "Partially augmented Lagrangian method for matrix inequality constraints," *SIAM Journal on Optimization*, vol. 15, no. 1, pp. 161–184, 2004.

[32] A. Shapiro and M. K. H. Fan, "On eigenvalue optimization," *SIAM Journal on Optimization*, vol. 5, no. 3, pp. 552–569, 1995.