Appendix A. Proof of Theorem 1

In this appendix, we first derive a canonical form of the pencil $L_G - \lambda L_H$, and then prove the variational principle in Theorem 1. For the simplicity of notation, in this appendix, we denote $A = L_G$ and $B = L_H$. We begin with the following lemma.

Lemma 1. *If $A - \lambda B$ is a symmetric matrix pencil of order n with $A \succeq 0$ and $B \succeq 0$, then there exists an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ such that*

$$Q^T A Q = \begin{array}{c} r \\ n_1 \\ m \end{array} \overset{\begin{array}{ccc} r & n_1 & m \end{array}}{\begin{bmatrix} \widehat{A}_{11} & \widehat{A}_{12} & \\ \widehat{A}_{12}^T & \widehat{A}_{22} & \\ & & 0 \end{bmatrix}} \equiv \begin{array}{c} r+n_1 \\ m \end{array} \overset{\begin{array}{cc} r+n_1 & m \end{array}}{\begin{bmatrix} \widehat{A} & \\ & 0 \end{bmatrix}}, \tag{1}$$

$$Q^T B Q = \begin{array}{c} r \\ n_1 \\ m \end{array} \overset{\begin{array}{ccc} r & n_1 & m \end{array}}{\begin{bmatrix} \widehat{B}_{11} & & \\ & 0 & \\ & & 0 \end{bmatrix}} \equiv \begin{array}{c} r+n_1 \\ m \end{array} \overset{\begin{array}{cc} r+n_1 & m \end{array}}{\begin{bmatrix} \widehat{B} & \\ & 0 \end{bmatrix}}, \tag{2}$$

where $\widehat{A}_{22} \succ 0$ and $\widehat{B}_{11} \succ 0$. Furthermore, the sub-pencil $\widehat{A} - \lambda \widehat{B}$ is regular and $\widehat{A} \succeq 0$ and $\widehat{B} \succeq 0$.

Proof. Since $B \succeq 0$, there exists an orthogonal matrix $Q_1 \in \mathbb{R}^{n\times n}$ such that

$$B^{(0)} \equiv Q_1^T B Q_1 = \begin{array}{c} r \\ d \end{array} \overset{\begin{array}{cc} r & d \end{array}}{\begin{bmatrix} \widehat{B}_{11} & \\ & 0 \end{bmatrix}}, \tag{3}$$

where $\widehat{B}_{11} \succ 0$. Applying transformation Q_1 to matrix A, we have

$$A^{(0)} \equiv Q_1^T A Q_1 = \begin{array}{c} r \\ d \end{array} \overset{\begin{array}{cc} r & d \end{array}}{\begin{bmatrix} \widehat{A}_{11} & A_{12} \\ A_{12}^T & A_{22} \end{bmatrix}}.$$

Note that $A_{22} \succeq 0$ due to the fact that $A \succeq 0$.

For the $d \times d$ block matrix A_{22}, there exists an orthogonal matrix $Q_{22} \in \mathbb{R}^{d\times d}$ such that

$$Q_{22}^T A_{22} Q_{22} = \begin{array}{c} n_1 \\ m \end{array} \overset{\begin{array}{cc} n_1 & m \end{array}}{\begin{bmatrix} \widehat{A}_{22} & \\ & 0 \end{bmatrix}},$$

where $\widehat{A}_{22} \succ 0$.

Let $Q_2 = \mathrm{diag}(I_r, Q_{22})$. Then we have

$$A^{(1)} \equiv Q_2^T A^{(0)} Q_2 = \begin{array}{c} r \\ n_1 \\ m \end{array} \overset{\begin{array}{ccc} r & n_1 & m \end{array}}{\begin{bmatrix} \widehat{A}_{11} & \widehat{A}_{12} & \widehat{A}_{13} \\ \widehat{A}_{12}^T & \widehat{A}_{22} & \\ \widehat{A}_{13}^T & & 0 \end{bmatrix}},$$

$$B^{(1)} \equiv Q_2^T B^{(0)} Q_2 = \begin{array}{c} r \\ n_1 \\ m \end{array} \overset{\begin{array}{ccc} r & n_1 & m \end{array}}{\begin{bmatrix} \widehat{B}_{11} & & \\ & 0 & \\ & & 0 \end{bmatrix}},$$

where $\begin{bmatrix} \widehat{A}_{12}, & \widehat{A}_{13} \end{bmatrix} = A_{12} Q_{22}$. Note that since $A^{(1)} \succeq 0$, we must have $\widehat{A}_{13} = 0$. Otherwise, if there exists an element $a_{ij} \neq 0$ in $\widehat{A}_{13}$, then the 2 by 2 sub-matrix $\begin{bmatrix} \widehat{a}_{ii} & a_{ij} \\ a_{ij} & 0 \end{bmatrix}$ of $A^{(1)}$ is indefinite, where $\widehat{a}_{ii}$ is the i-th diagonal element of $\widehat{A}_{11}$. This contradicts to the positive semi-definiteness of $A^{(1)} \succeq 0$.

Denote $Q = Q_1 Q_2$. Then Q is orthogonal, and $Q^T AQ,\ Q^T BQ$ have the form (1).

Finally, we show the pencil $\widehat{A} - \lambda\widehat{B}$ is regular. For any $\lambda \in \mathbb{C}$, straightforward calculation gives that

$$\begin{aligned}\det(\widehat{A} - \lambda\widehat{B}) &= \det\begin{pmatrix} \widehat{A}_{11} - \lambda\widehat{B}_{11} & \widehat{A}_{12} \\ \widehat{A}_{12}^T & \widehat{A}_{22} \end{pmatrix} \\ &= \det\begin{pmatrix} \widehat{A}_{11} - \widehat{A}_{12}\widehat{A}_{22}^{-1}\widehat{A}_{12}^T - \lambda\widehat{B}_{11} & \\ \widehat{A}_{12}^T & \widehat{A}_{22} \end{pmatrix} \\ &= \det(\widehat{A}_{22})\det(\widehat{A}_{11} - \widehat{A}_{12}\widehat{A}_{22}^{-1}\widehat{A}_{12}^T - \lambda\widehat{B}_{11}).\end{aligned}$$

Recall that $\widehat{A}_{22} \succ 0$. Furthermore, since $\widehat{B}_{11} \succ 0$, $\det(\widehat{A}_{11} - \widehat{A}_{12}\widehat{A}_{22}^{-1}\widehat{A}_{12}^T - \lambda\widehat{B}_{11}) \not\equiv 0$. Hence, $\det(\widehat{A} - \lambda\widehat{B}) \not\equiv 0$. This means the pencil $\widehat{A} - \lambda\widehat{B}$ is regular. □

By Lemma 1, we have the following canonical form of the matrix pair $\{A, B\}$ to show that the matrices A and B are simultaneously diagonalizable with a congruence transformation.

Lemma 2. *If $A - \lambda B$ is a symmetric matrix pencil of order n with $A \succeq 0$ and $B \succeq 0$, then there exists a nonsingular matrix $X \in \mathbb{R}^{n\times n}$ such that*

$$X^T AX = \begin{matrix} r \\ n_1 \\ m \end{matrix}\overset{\begin{matrix} r & n_1 & m \end{matrix}}{\begin{bmatrix} \Lambda_r & & \\ & I & \\ & & 0 \end{bmatrix}}, \quad X^T BX = \begin{matrix} r \\ n_1 \\ m \end{matrix}\overset{\begin{matrix} r & n_1 & m \end{matrix}}{\begin{bmatrix} I & & \\ & 0 & \\ & & 0 \end{bmatrix}}, \tag{4}$$

where Λ_r is a diagonal matrix of non-negative diagonal elements $\lambda_1, \ldots, \lambda_r$, $r = \operatorname{rank}(B)$, $m = \dim(\mathcal{N}(A)\cap\mathcal{N}(B))$ and $n_1 = \dim(\mathcal{N}(B)) - m$.

Proof. By Lemma 1, there exists an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ such that

$$A^{(1)} \equiv Q^T AQ = \begin{matrix} r \\ n_1 \\ m \end{matrix}\overset{\begin{matrix} r & n_1 & m \end{matrix}}{\begin{bmatrix} \widehat{A}_{11} & \widehat{A}_{12} & \\ \widehat{A}_{12}^T & \widehat{A}_{22} & \\ & & 0 \end{bmatrix}} \quad \text{and} \quad B^{(1)} \equiv Q^T BQ = \begin{matrix} r \\ n_1 \\ m \end{matrix}\overset{\begin{matrix} r & n_1 & m \end{matrix}}{\begin{bmatrix} \widehat{B}_{11} & & \\ & 0 & \\ & & 0 \end{bmatrix}}.$$

Let

$$X_1 = \begin{bmatrix} I_r & & \\ -\widehat{A}_{22}^{-1}\widehat{A}_{12}^T & \widehat{A}_{22}^{-1/2} & \\ & & I_s \end{bmatrix}.$$

Then

$$A^{(2)} \equiv X_1^T A^{(1)} X_1 = \begin{bmatrix} \widehat{A}_{11} - \widehat{A}_{12}^{(1)}\widehat{A}_{22}^{-1}\widehat{A}_{12}^T & & \\ & I_{n_1} & \\ & & 0_s \end{bmatrix} \quad \text{and} \quad B^{(2)} \equiv X_1^T B^{(1)} X_1 = \begin{bmatrix} \widehat{B}_{11} & & \\ & 0_{n_1} & \\ & & 0_s \end{bmatrix}.$$

Since $\widehat{B}_{11} \succ 0$, there exists a nonsingular matrix $\widehat{X}_2$ such that

$$\widehat{X}_2^T[\widehat{A}_{11} - \widehat{A}_{12}^{(1)}\widehat{A}_{22}^{-1}\widehat{A}_{12}^T]\widehat{X}_2 = \Lambda, \ \widehat{X}_2^T\widehat{B}_{11}\widehat{X}_2 = I_r.$$

Let $X_2 = \operatorname{diag}(\widehat{X}_2, I_{n_1}, I_s)$. Then we have

$$X_2^T A^{(2)} X_2 = \operatorname{diag}(\Lambda, I_{n_1}, 0_s), \ X_2^T B^{(2)} X_2 = \operatorname{diag}(I_r,\ 0_{n_1},\ 0_s).$$

Denote $X = QX_1X_2$. Then we obtain (4). The remaining results are eaily obtained from the canonical form (2). □

The following remarks are in order:

1. By Lemma 2, we know (i) there are $r = \text{rank}(B)$ finite eigenvalues of the pencil $A - \lambda B$ and all finite eigenvalues are real, nonnegative and non-defective. and (ii) there are $n_1 = \dim(\mathcal{N}(B)) - \dim(\mathcal{N}(A) \cap \mathcal{N}(B))$ non-defective infinite eigenvalues.

2. The canonical form (4) has been derived in [Newcomb, 1961]. Here we give the values of indices r, n_1, m in (4) and our proof seems more compact.

3. Lemma 3.8 in [Liang et al., 2013] deals with the canonical form of a general positive semi-definite pencil. Obviously, the pencil $A - \lambda B$ considered here is a special case of positive semi-definite pencil. So Lemma 3.8 is applicable here. Our proof is constructive based on Fix-Heiberger's reduction [Fix and Heiberger, 1972].

We now provide a proof of the variational principle in Theorem 1. Without loss of generality, we assume that pencil $A - \lambda B$ is in the canonical form (4), i.e.,

$$A = \begin{array}{c} \scriptstyle r \\ \scriptstyle n_1 \\ \scriptstyle m \end{array} \overset{\begin{array}{ccc} \scriptstyle r & \scriptstyle n_1 & \scriptstyle m \end{array}}{\begin{bmatrix} \Lambda_r & & \\ & I & \\ & & 0 \end{bmatrix}}, \quad B = \begin{array}{c} \scriptstyle r \\ \scriptstyle n_1 \\ \scriptstyle m \end{array} \overset{\begin{array}{ccc} \scriptstyle r & \scriptstyle n_1 & \scriptstyle m \end{array}}{\begin{bmatrix} I & & \\ & 0 & \\ & & 0 \end{bmatrix}}. \tag{5}$$

Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a subspace of dimension $n+1-i$, where $1 \leqslant i \leqslant r$ and $x \in \mathcal{X}$ be partitioned into $x = [x_1^T, x_2^T, x_3^T]^T$ conformally with the form (5), then

$$\inf_{\substack{x \in \mathcal{X} \\ x^T B x > 0}} \frac{x^T A x}{x^T B x} = \inf_{\substack{x \in \mathcal{X} \\ x_1^T x_1 > 0}} \frac{x_1^T \Lambda_r x_1 + x_2^T x_2}{x_1^T x_1} = \inf_{\substack{x \in \mathcal{X} \\ x_1^T x_1 > 0}} \frac{x_1^T \Lambda_r x_1}{x_1^T x_1}. \tag{6}$$

Let $\mathcal{X}^{(1)} = \{[I_r, 0_{n-r}]x \mid x \in \mathcal{X}\}$. Evidently, $\mathcal{X}^{(1)}$ is a subspace of $\mathbb{R}^r$. Moreover,

$$n + 1 - i \geqslant \dim(\mathcal{X}^{(1)}) \geqslant n + 1 - i - n_1 - s = r + 1 - i.$$

Then there exists a subspace $\widetilde{\mathcal{X}} \subseteq \mathbb{R}^r$ of dimension $r + 1 - i$ such that $\widetilde{\mathcal{X}} \subseteq \mathcal{X}^{(1)}$. For the matrix Λ_r, by Courant-Fischer min-max principle, we have

$$\inf_{\substack{x \in \mathcal{X} \\ x_1^T x_1 > 0}} \frac{x_1^T \Lambda_r x_1}{x_1^T x_1} = \min_{\substack{x_1 \in \mathcal{X}^{(1)} \\ x_1^T x_1 > 0}} \frac{x_1^T \Lambda_r x_1}{x_1^T x_1} \leqslant \min_{\substack{x_1 \in \widetilde{\mathcal{X}} \\ x_1^T x_1 > 0}} \frac{x_1^T \Lambda_r x_1}{x_1^T x_1} \leqslant \max_{\substack{\dim(\mathcal{S}) = r+1-i \\ \mathcal{S} \subseteq \mathbb{R}^r}} \min_{\substack{x_1 \in \mathcal{S} \\ x_1^T x_1 > 0}} \frac{x_1^T \Lambda_r x_1}{x_1^T x_1} = \lambda_i.$$

Combining above equation with (6), we know that for any subspace $\mathcal{X} \subseteq \mathbb{R}^n$ with dimension $n + 1 - i$,

$$\min_{\substack{x \in \mathcal{X} \\ x^T B x > 0}} \frac{x^T A x}{x^T B x} \leqslant \lambda_i. \tag{7}$$

On the other hand, let us consider a special choice of the subspace $\mathcal{X}$:

$$\mathcal{S}_i = \mathcal{R}(S_i),$$

where

$$S_i = \begin{array}{c} \scriptstyle i-1 \\ \scriptstyle r+1-i \\ \scriptstyle n-r \end{array} \overset{\begin{array}{cc} \scriptstyle r+1-i & \scriptstyle n-r \end{array}}{\begin{bmatrix} 0 & \\ I & 0 \\ & I \end{bmatrix}}.$$

Then $\dim(\mathcal{S}_i) = n + 1 - i$, and

$$S_i^T A S_i = \text{diag}(\widetilde{\Lambda}_i,\ I_{n_1},\ 0_s), \quad S_i^T B S_i = \text{diag}(I_{r+1-i},\ 0_{n_1},\ 0_s),$$

where $\widetilde{\Lambda}_i = \text{diag}(\lambda_i, \cdots, \lambda_r)$. Let $x_* = S_i e_1 \in \mathcal{S}_i$, where e_1 is a unit vector of dimension $n + r - i$, then

$$\frac{x_*^T A x_*}{x_*^T B x_*} = \lambda_i.$$

Consequently, Eq.17 (Sec.4) follows from above equation and (7). Taking $i = 1$ in (7), we get Eq.18 (Sec.4).

Appendix B. Proof of Theorem 2

Similar to Appendix A, for the simplicity of notation, we denote $A = L_G$ and $B = L_H$. By the definitions of K and M in Theorem 2, we have

$$K = -B, \quad M = A + \mu B + ZSZ^T.$$

By Lemma 1, there exists an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ such that

$$Q^T A Q = \begin{array}{c} \\ {\scriptstyle n-m} \\ {\scriptstyle m} \end{array} \begin{array}{c} \begin{array}{cc} {\scriptstyle n-m} & {\scriptstyle m} \end{array} \\ \begin{bmatrix} \widehat{A} & \\ & 0 \end{bmatrix} \end{array}, \quad Q^T B Q = \begin{array}{c} \\ {\scriptstyle n-m} \\ {\scriptstyle m} \end{array} \begin{array}{c} \begin{array}{cc} {\scriptstyle n-m} & {\scriptstyle m} \end{array} \\ \begin{bmatrix} \widehat{B} & \\ & 0 \end{bmatrix} \end{array}, \tag{8}$$

where the $(n-m)\times(n-m)$ sub-pencil $\widehat{A} - \lambda\widehat{B}$ is regular and $\widehat{A} \succeq 0$ and $\widehat{B} \succeq 0$.

Let Q in (8) be conformally partitioned in the form $Q = [Q_1, Q_2]$, where $Q_2 \in \mathbb{R}^{n\times m}$. Then Q_2 is also an orthonormal basis of $\mathcal{N}(A) \cap \mathcal{N}(B)$, i.e.,

$$Z = Q_2 G \tag{9}$$

for some orthogonal matrix G.

For the regular pair $\{\widehat{A}, \widehat{B}\}$, by Lemma 2, there exists a nonsingular matrix $\widetilde{X} \in \mathbb{R}^{(n-m)\times(n-m)}$ such that

$$\widetilde{X}^T \widehat{A} \widetilde{X} = \mathrm{diag}(\Lambda_r, I_{n_1}), \;\; \widetilde{X}^T \widehat{B} \widetilde{X} = \mathrm{diag}(I_r, 0_{n_1}), \tag{10}$$

where $\Lambda_r = \mathrm{diag}(\lambda_1, \cdots, \lambda_r) \succeq 0$.

Let $X = Q\,\mathrm{diag}(\widetilde{X}, I_m)$. Then

$$\begin{aligned} X^T K X &= \mathrm{diag}(\widetilde{X}^T, I_m) Q^T (-B) Q\,\mathrm{diag}(\widetilde{X}, I_m) \\ &= \mathrm{diag}(\widetilde{X}^T, I_m)\,\mathrm{diag}(-\widehat{B}, 0_m)\,\mathrm{diag}(\widetilde{X}^T, I_m) \qquad \text{by (8)} \\ &= \mathrm{diag}(-I_r,\ 0_{n_1}, 0_m) \qquad \text{by (10)}, \end{aligned}$$

and

$$\begin{aligned} X^T M X &= \mathrm{diag}(\widetilde{X}^T,\ I_m) Q^T (A + \mu B + ZSZ^T) Q\,\mathrm{diag}(\widetilde{X},\ I_m) \\ &= \mathrm{diag}(\widetilde{X}^T,\ I_m)\,\mathrm{diag}(\widehat{A} + \mu\widehat{B}, GSG^T)\,\mathrm{diag}(\widetilde{X},\ I_m) \qquad \text{by (8) and (9)} \\ &= \mathrm{diag}(\Lambda_r + \mu I_r,\ I_{n_1}, GSG^T) \qquad \text{by (10)}. \end{aligned}$$

Since $\Lambda_r \succeq 0$, $S > 0$ and $\mu > 0$, $M \succ 0$. The nonzero eigenvalues of the pencil $K - \sigma M$ are $\sigma_i = -1/(\lambda_i + \mu)$ for $i = 1, \ldots, r$.

References

[Fix and Heiberger, 1972] Fix, G. and Heiberger, R. (1972). An algorithm for the ill-conditioned generalized eigenvalue problem. *SIAM Journal on Numerical Analysis*, 9:78–88.

[Liang et al., 2013] Liang, X., Li, R.-C., and Bai, Z. (2013). Trace minimization principles for positive semi-definite pencils. *Linear Algebra and its Applications*, 438:3085–3106.

[Newcomb, 1961] Newcomb, R. W. (1961). On the simultaneous diagonalization of two semi-definite matrices. *Quarterly of Applied Mathematics*, 19:144–146.