SIAM J. OPTIM.
Vol. 21, No. 4, pp. 1392–1417

NONSINGULARITY CONDITIONS FOR THE FISCHER–BURMEISTER SYSTEM OF NONLINEAR SDPS*

SHUJUN BI†, SHAOHUA PAN†, AND JEIN-SHAN CHEN‡

Abstract. For a locally optimal solution to the nonlinear semidefinite programming problem, under Robinson's constraint qualification, we show that the nonsingularity of Clarke's Jacobian of the Fischer–Burmeister (FB) nonsmooth system is equivalent to the strong regularity of the Karush–Kuhn–Tucker point. Consequently, from Sun's paper [*Math. Oper. Res.*, 31 (2006), pp. 761–776] the semismooth Newton method applied to the FB system may attain the locally quadratic convergence under the strong second order sufficient condition and constraint nondegeneracy.

Key words. nonlinear semidefinite programming problem, the FB system, Clarke's Jacobian, nonsingularity, strong regularity

AMS subject classifications. 90C22, 90C25, 90C31, 65K05

DOI. 10.1137/110824577

1. Introduction.

Let $\mathbb{X}$ be a finite dimensional real vector space endowed with an inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\|\cdot\|$. Consider the nonlinear semidefinite programming problem (NLSDP)

$$\begin{aligned} &\min_{x\in\mathbb{X}} \quad f(x) \\ &\text{s.t.} \quad h(x) = 0, \\ &\qquad\; g(x) \in \mathbb{S}^n_+, \end{aligned} \tag{1}$$

where $f : \mathbb{X} \to \mathbb{R}, h : \mathbb{X} \to \mathbb{R}^m$ and $g : \mathbb{X} \to \mathbb{S}^n$ are twice continuously differentiable functions, $\mathbb{S}^n$ is the linear space of all $n\times n$ real symmetric matrices, and $\mathbb{S}^n_+$ is the cone of all $n \times n$ positive semidefinite matrices. By introducing a slack variable $X \in \mathbb{S}^n_+$ for the conic constraint $g(x) \in \mathbb{S}^n_+$, we can rewrite the NLSDP (1) as follows:

$$\begin{aligned} &\min_{(x,X)\in\mathbb{X}\times\mathbb{S}^n} f(x) \\ &\text{s.t.} \quad h(x) = 0, \\ &\qquad\; g(x) - X = 0, \\ &\qquad\; X \in \mathbb{S}^n_+. \end{aligned} \tag{2}$$

In this paper, we will concentrate on this equivalent formulation of the NLSDP (1).

The Karush–Kuhn–Tucker (KKT) condition for the NLSDP (2) takes the form

$$\mathcal{J}_{x,X} L(x, X, \mu, S, Y) = 0,\ h(x) = 0,\ g(x) - X = 0,\ -Y \in \mathcal{N}_{\mathbb{S}^n_+}(X), \tag{3}$$

*Received by the editors February 14, 2011; accepted for publication (in revised form) August 26, 2011; published electronically December 6, 2011. This work was supported by the National Young Natural Science Foundation (10901058) and the Guangdong Natural Science Foundation (9251802902000001).

http://www.siam.org/journals/siopt/21-4/82457.html

†Department of Mathematics, South China University of Technology, Guangzhou, China (beamilan@163.com, shhpan@scut.edu.cn).

‡Corresponding author. Department of Mathematics, National Taiwan Normal University, Taipei, Taiwan 11677 (jschen@math.ntnu.edu.tw). Member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. This author's work is partially supported by the National Science Council of Taiwan.

where the Lagrangian function $L\colon \mathbb{X}\times\mathbb{S}^n\times\mathbb{R}^m\times\mathbb{S}^n\times\mathbb{S}^n\to\mathbb{R}$ is defined by

$$L(x,X,\mu,S,Y) := f(x) + \langle \mu, h(x)\rangle + \langle S, g(x) - X\rangle - \langle X, Y\rangle,$$

$\mathcal{J}_{x,X}L(x,X,\mu,S,Y)$ is the derivative of L at (x,X,μ,S,Y) with respect to (x,X), and $\mathcal{N}_{\mathbb{S}^n_+}(X)$ denotes the normal cone of $\mathbb{S}^n_+$ at X in the sense of convex analysis [17]:

$$\mathcal{N}_{\mathbb{S}^n_+}(X) = \begin{cases} \{Z\in\mathbb{S}^n :\ \langle Z, W-X\rangle \le 0 \quad \forall W\in\mathbb{S}^n\} & \text{if } X\in\mathbb{S}^n_+,\\ \emptyset & \text{if } X\notin\mathbb{S}^n_+. \end{cases}$$

Recall that $\Phi\colon \mathbb{S}^n\times\mathbb{S}^n\to\mathbb{S}^n$ is a semidefinite cone (SDC) complementarity function if

$$\Phi(X,Y)=0 \iff X\in\mathbb{S}^n_+,\ Y\in\mathbb{S}^n_+,\ \langle X,Y\rangle = 0 \iff -Y\in\mathcal{N}_{\mathbb{S}^n_+}(X).$$

Then, with an SDC complementarity function Φ, the KKT optimality conditions in (3) can be reformulated as the following nonsmooth system:

$$E(x,X,\mu,S,Y) := \begin{bmatrix} \mathcal{J}_{x,X}L(x,X,\mu,S,Y)\\ h(x)\\ g(x)-X\\ \Phi(X,Y)\end{bmatrix} = 0. \tag{4}$$

The most popular SDC complementarity functions include the matrix-valued natural residual (NR) function and the Fischer–Burmeister (FB) function, which are defined as

$$\Phi_{\rm NR}(X,Y) \ :=\ X - \Pi_{\mathbb{S}^n_+}(X-Y) \qquad \forall X,Y\in\mathbb{S}^n$$

and

$$\Phi_{\rm FB}(X,Y) \ :=\ (X+Y) - \sqrt{X^2+Y^2} \qquad \forall X,Y\in\mathbb{S}^n, \tag{5}$$

respectively, where $\Pi_{\mathbb{S}^n_+}(\cdot)$ denotes the projection operator onto $\mathbb{S}^n_+$. It turns out that $\Phi_{\rm FB}$ has almost all favorable properties of $\Phi_{\rm NR}$ (see [21]). Also, the squared norm of $\Phi_{\rm FB}$ induces a continuously differentiable merit function whose derivative is globally Lipschitz continuous [**?**, 24]. This greatly facilitates the globalization of the semismooth Newton method [15, 16] for solving the FB system of (2). The FB system and the NR system mean $E_{\rm FB}(x,X,\mu,S,Y)=0$ and $E_{\rm NR}(x,X,\mu,S,Y)=0$, respectively, with the mappings $E_{\rm FB}$ and $E_{\rm NR}$ defined as in E except that Φ is specified as $\Phi_{\rm FB}$ and $\Phi_{\rm NR}$, respectively.

The strong regularity is one of the important concepts in sensitivity and perturbation analysis introduced by Robinson in his seminal paper [18]. For the NLSDP (1), Sun [22] offered a characterization for the strong regularity via the study of the nonsingularity of Clarke's Jacobian of the NR system under the strong second order sufficient condition and constraint nondegeneracy, and he established its equivalence to other characterizations discussed in a wide range of literature. Later, for the linear semidefinite programming problem (SDP), Chan and Sun [3] gained more insightful characterizations for the strong regularity via the study of the nonsingularity of Clarke's Jacobian of the NR system, too. Then, it is natural for us to ask the following question: is it possible to give a characterization for the strong regularity of NLSDPs by studying the nonsingularity of Clarke's Jacobian of the FB system? Note that up to now one does not even know whether the B-subdifferential of the FB system is nonsingular or not without strict complementarity of locally optimal solutions.

In this work, for a locally optimal solution to the NLSDP (2), we prove that under Robinson's constraint qualification, the nonsingularity of Clarke's Jacobian of the FB system is equivalent to the strong regularity of the KKT point, which by [22, Theorem 4.1] is further equivalent to the strong second order sufficient condition and constraint nondegeneracy. This result is of interest since, on one hand, it relates the nonsingularity of Clarke's Jacobian of the FB system to Robinson's strong regularity condition, and, on the other hand, it allows us to obtain the quadratic convergence of the semismooth Newton method [16, 15] for the FB system without strict complementarity assumption. In addition, it also extends the result of [9, Corollary 3.7] for the variational inequality with the polyhedral cone constraints to the setting of semidefinite cones. It is worthwhile to point out that [22, Theorem 4.1] plays a key role in achieving this objective.

Throughout this paper, $\mathcal{J}_z f(z)$ and $\mathcal{J}_{zz}^2 f(z)$ denote the derivative and the second order derivative, respectively, of a twice differentiable function f with respect to z, and $\mathcal{I}$ denotes an identity operator. For any $n \times m$ real matrices A and B, $\langle A, B\rangle$ means their Frobenius inner product, and $\|A\|$ denotes the norm of A induced by the Frobenius inner product. For $X \in \mathbb{S}^n$, we write $X \succeq 0$ (respectively, $X \succ 0$) to mean $X \in \mathbb{S}^n_+$ (respectively, $X \in \mathbb{S}^n_{++}$). For a linear operator $\mathcal{A}$, we denote by $\mathcal{A}^*$ the adjoint of $\mathcal{A}$, and by $\|\mathcal{A}\|_2$ the operator norm of $\mathcal{A}$. For a linear operator $\mathcal{A} : \mathbb{S}^n \to \mathbb{S}^n$, we write $\mathcal{A} \succeq 0$ (respectively, $\mathcal{A} \succ 0$) if $\langle W, \mathcal{A}(W)\rangle \geq 0$ for any $W \in \mathbb{S}^n$ (respectively, $\langle W, \mathcal{A}(W)\rangle > 0$ for any nonzero $W \in \mathbb{S}^n$). For any given sets of indices α and β, we designate by $A_{\alpha\beta}$ the submatrix of A whose row indices belong to α and whose column indices belong to β, and we use $|\alpha|$ to denote the number of elements in the set α.

2. Preliminary results. Let $\mathbb{X}$ and $\mathbb{Y}$ be two arbitrary finite dimensional real vector spaces each equipped with a scalar product $\langle \cdot, \cdot\rangle$ and its induced norm $\|\cdot\|$. Let $\mathcal{O}$ be an open set in $\mathbb{X}$ and $\Xi : \mathcal{O} \to \mathbb{Y}$ be a locally Lipschitz continuous function on the set $\mathcal{O}$. By Rademacher's theorem, Ξ is almost everywhere Fréchet-differentiable (F-differentiable) in $\mathcal{O}$. We denote by $\mathcal{D}_\Xi$ the set of points in $\mathcal{O}$ where Ξ is F-differentiable. Then Clarke's Jacobian of Ξ at x is well defined [6]:

$$\partial\Xi(x) := \operatorname{conv}\{\partial_B \Xi(x)\},$$

where "conv" means the convex hull and $\partial_B\Xi(x)$ is the B-subdifferential of Ξ at x,

$$\partial_B\Xi(x) := \left\{V :\ V = \lim_{k\to\infty} \mathcal{J}_x\Xi(x^k),\ x^k \to x,\ x^k \in \mathcal{D}_\Xi\right\}.$$

For the concepts of (strong) semismoothness, please refer to the literature [16, 15, 20].

The following matrix inequalities are used in the proof of Lemma 3.3; see the appendix.

LEMMA 2.1. *For any $n \times m$ real matrices A, B and any $Z \in \mathbb{S}^n_+$, it holds that*

$$(A + B)^T Z(A + B) \preceq 2(A^T ZA + B^T ZB), \tag{6}$$

$$(A - B)^T Z(A - B) \preceq 2(A^T ZA + B^T ZB). \tag{7}$$

Proof. Fix any $Z \in \mathbb{S}^n_+$. Then, for any $n \times m$ real matrices A and B, we have that

$$0 \preceq (A - B)^T Z(A - B) = (A^T ZA + B^T ZB) - (A^T ZB + B^T ZA),$$
$$0 \preceq (A + B)^T Z(A + B) = (A^T ZA + B^T ZB) + (A^T ZB + B^T ZA).$$

The first equation means that $(A^TZB+B^TZA) \preceq (A^TZA+B^TZB)$, which along with the second equality yields (6). The second equation implies that $-(A^TZB+B^TZA) \preceq (A^TZA+B^TZB)$, which along with the first equality yields (7). □

LEMMA 2.2. *Let $X, Y \in \mathbb{S}^n$ with $X^2+Y^2 \succ 0$. Then for any $n \times m$ real matrices A, B,*

$$A^TA + B^TB - (A^TX + B^TY)(X^2 + Y^2)^{-1}(XA + YB) \succeq 0.$$

Proof. Note that $A^TA + B^TB - (A^TX + B^TY)(X^2 + Y^2)^{-1}(XA + YB)$ is the Schur complement of $X^2 + Y^2$ in the following block symmetric matrix:

$$\Sigma = \begin{bmatrix} X^2 + Y^2 & XA + YB \\ (XA + YB)^T & A^TA + B^TB \end{bmatrix}.$$

We need only prove $\Sigma \succeq 0$ (see [10, Theorem 7.7.6]). For any $\zeta = (\zeta_1, \zeta_2) \in \mathbb{R}^n \times \mathbb{R}^m$,

$$\begin{aligned} \zeta^T\Sigma\zeta &= \zeta_1^T(X^2 + Y^2)\zeta_1 + 2\zeta_1^T(XA + YB)\zeta_2 + \zeta_2^T(A^TA + B^TB)\zeta_2 \\ &= \|X\zeta_1 + A\zeta_2\|^2 + \|Y\zeta_1 + B\zeta_2\|^2 \geq 0, \end{aligned}$$

which shows that $\Sigma \succeq 0$. The proof is then complete. □

For any given $X \in \mathbb{S}^n$, let $\mathcal{L}_X : \mathbb{S}^n \to \mathbb{S}^n$ be the Lyapunov operator associated with X:

$$\mathcal{L}_X(Y) := XY + YX \qquad \forall Y \in \mathbb{S}^n.$$

We next study several properties of the Lyapunov operators associated with $X, Y \in \mathbb{S}^n$ and $Z \in \mathbb{S}^n_+$ with $Z^2 \succeq X^2 + Y^2$. To this end, we need to establish two trace inequalities.

LEMMA 2.3. *Let $X, Y \in \mathbb{S}^n$ with $X \succeq |Y|$. Then, for any $W \in \mathbb{S}^n$, it holds that*

$$\text{Trace}(WXWX) \geq \text{Trace}(WYWY).$$

Proof. Fix any $W \in \mathbb{S}^n$. By the trace property of symmetric matrices, we have that

$$\begin{aligned} &\text{Trace}(WXWX) - \text{Trace}(WYWY) \\ &= \text{Trace}\,[WXW(X - Y)] + \text{Trace}\,[W(X - Y)WY] \\ &= \text{Trace}\,[W(X - Y)WX] + \text{Trace}\,[W(X - Y)WY] \\ &= \text{Trace}\,[W(X - Y)W(X + Y)]\,. \end{aligned}$$

Since $X \succeq |Y|$, we have $W(X - Y)W \succeq 0$ and $X + Y \succeq 0$. From [10, Theorem 7.6.3], it then follows that $\text{Trace}\,[W(X - Y)W(X + Y)] \geq 0$. The result is thus proved. □

LEMMA 2.4. *For any given $X, Y \in \mathbb{S}^n$ and $Z \in \mathbb{S}^n_+$ satisfying $Z \succeq \sqrt{X^2 + Y^2}$, we have*

$$\text{Trace}(WZWZ) \geq \text{Trace}(W|X|W|X|) + \text{Trace}(W|Y|W|Y|) \qquad \forall W \in \mathbb{S}^n.$$

Proof. Fix any $W \in \mathbb{S}^n$. Applying Lemma 2.3, we readily obtain that

$$\text{Trace}(WZWZ) \geq \text{Trace}\left(W\sqrt{X^2 + Y^2}W\sqrt{X^2 + Y^2}\right). \tag{8}$$

In addition, from [1, Theorem IX.6.1], we know that $\varphi(A,B) := \mathrm{Trace}(W\sqrt{A}W\sqrt{B})$ is a jointly concave function on $\mathbb{S}^n_+ \times \mathbb{S}^n_+$, which means that for any $A_1, A_2, B_1, B_2 \in \mathbb{S}^n_+$,

$$\varphi\left(\frac{A_1+A_2}{2}, \frac{B_1+B_2}{2}\right) \geq \frac{1}{2}\left[\varphi(A_1,B_1)+\varphi(A_2,B_2)\right].$$

Using this inequality with $A_1 = B_1 = X^2$ and $A_2 = B_2 = Y^2$, we obtain that

$$2\varphi\left(\frac{X^2+Y^2}{2}, \frac{X^2+Y^2}{2}\right) \geq \mathrm{Trace}(W|X|W|X|) + \mathrm{Trace}(W|Y|W|Y|).$$

This, together with the definition of φ and inequality (8), implies the result. ∎

The following proposition, extending the result of [8, Proposition 3.4] associated with second order cones to SDCs, is used to prove Proposition 2.2. Among other methods, Proposition 2.2 is the key to characterizing the properties of Clarke's Jacobian of $\Phi_{\rm FB}$; see section 4.

PROPOSITION 2.1. *For any given $X, Y \in \mathbb{S}^n$ and $Z \in \mathbb{S}^n_+$, the following implication holds:*

$$Z^2 \succeq X^2 + Y^2 \implies \mathcal{L}_Z^2 \succeq \mathcal{L}_X^2 + \mathcal{L}_Y^2.$$

Proof. Since $Z^2 \succeq X^2 + Y^2$ and $Z \in \mathbb{S}^n_+$, from [1, Proposition V.1.8] it follows that

$$Z \succeq \sqrt{X^2+Y^2}.$$

Now choose a matrix $W \in \mathbb{S}^n$ arbitrarily. Then, a simple computation yields that

$$\begin{aligned}
\langle W, (\mathcal{L}_Z^2 - \mathcal{L}_X^2 - \mathcal{L}_Y^2)W\rangle &= 2\big[\mathrm{Trace}(WZWZ) + \mathrm{Trace}(W^2Z^2) - \mathrm{Trace}(WXWX) \\
&\qquad - \mathrm{Trace}(W^2X^2) - \mathrm{Trace}(W^2Y^2) - \mathrm{Trace}(WYWY)\big] \\
&= 2\big[\mathrm{Trace}\big(W^2(Z^2-X^2-Y^2)\big) + \mathrm{Trace}(WZWZ) \\
&\qquad - \mathrm{Trace}(WXWX) - \mathrm{Trace}(WYWY)\big] \\
&\geq 2\left[\mathrm{Trace}(WZWZ) - \mathrm{Trace}(WXWX) - \mathrm{Trace}(WYWY)\right] \\
&\geq 0,
\end{aligned}$$

where the first inequality is due to $Z^2 \succeq X^2+Y^2$ and the second uses $Z \succeq \sqrt{X^2+Y^2}$ and Lemmas 2.4 and 2.3. Since W is arbitrary in $\mathbb{S}^n$, the result follows. ∎

PROPOSITION 2.2. *For any given $X, Y \in \mathbb{S}^n$ and $Z \in \mathbb{S}^n_{++}$, define $\mathcal{A}\colon \mathbb{S}^n \times \mathbb{S}^n \to \mathbb{S}^n$ by*

$$\mathcal{A}(\triangle U, \triangle V) := \mathcal{L}_Z^{-1}\mathcal{L}_X(\triangle U) + \mathcal{L}_Z^{-1}\mathcal{L}_Y(\triangle V) \quad \forall \triangle U, \triangle V \in \mathbb{S}^n.$$

If $Z^2 \succeq X^2 + Y^2$, then the linear operator $\mathcal{A}$ satisfies $\|\mathcal{A}\|_2 \leq 1$, and consequently

$$\left\|\mathcal{L}_Z^{-1}\mathcal{L}_X(\triangle U) + \mathcal{L}_Z^{-1}\mathcal{L}_Y(\triangle V)\right\| \leq \sqrt{\|\triangle U\|^2 + \|\triangle V\|^2} \quad \forall \triangle U, \triangle V \in \mathbb{S}^n. \tag{9}$$

Proof. Assume that $Z^2 \succeq X^2 + Y^2$. By the definition of $\mathcal{A}$ and Proposition 2.1, we have

$$\mathcal{A}\mathcal{A}^* = \mathcal{L}_Z^{-1}(\mathcal{L}_X^2 + \mathcal{L}_Y^2)\mathcal{L}_Z^{-1} \preceq \mathcal{L}_Z^{-1}\mathcal{L}_Z^2\mathcal{L}_Z^{-1} = \mathcal{I}.$$

This means that the largest eigenvalue of $\mathcal{A}\mathcal{A}^*$ is less than 1, and consequently,

$$\|\mathcal{A}\|_2 = \sqrt{\|\mathcal{A}^*\mathcal{A}\|_2} = \sqrt{\lambda_{\max}(\mathcal{A}^*\mathcal{A})} = \sqrt{\lambda_{\max}(\mathcal{A}\mathcal{A}^*)} \le 1.$$

This completes the proof of the first part. By the definition of operator norm, we have

$$\|\mathcal{L}_Z^{-1}\mathcal{L}_X(\triangle U) + \mathcal{L}_Z^{-1}\mathcal{L}_Y(\triangle V)\| = \|\mathcal{A}(\triangle U, \triangle V)\| \le \|\mathcal{A}\|_2\, \|(\triangle U, \triangle V)\|.$$

Together with the first part, we prove that the inequality (9) holds. ∎

Let α, β, and γ be disjoint index sets with $\alpha \cup \beta \cup \gamma = \{1, 2, \ldots, n\}$. Define

$$\Gamma(X, Y) := (X_{\beta\beta}^2 + Y_{\beta\beta}^2 + X_{\beta\gamma}X_{\gamma\beta} + Y_{\beta\alpha}Y_{\alpha\beta})^{1/2} \quad \forall X, Y \in \mathbb{S}^n. \tag{10}$$

The following property of the function Γ will be used in the subsequent sections.

Proposition 2.3. *Let $X, Y \in \mathbb{S}^n$ be such that $\Gamma(X, Y) \succ 0$. Then for any $G, H \in \mathbb{S}^n$,*

$$\|\mathcal{L}^{-1}_{\Gamma(X,Y)}(X_{\beta\gamma}G_{\gamma\beta} + G_{\beta\gamma}X_{\gamma\beta})\| \le 2\sqrt{|\beta||\gamma|}\, \|G_{\gamma\beta}\|,$$
$$\|\mathcal{L}^{-1}_{\Gamma(X,Y)}(Y_{\beta\alpha}H_{\alpha\beta} + H_{\beta\alpha}Y_{\alpha\beta})\| \le 2\sqrt{|\beta||\alpha|}\, \|H_{\alpha\beta}\|.$$

Proof. Let $\Gamma(X, Y) = Q_\beta \operatorname{diag}(\lambda_1, \ldots, \lambda_{|\beta|})Q_\beta^T$ be the spectral decomposition of $\Gamma(X, Y)$, where $\lambda_i > 0$ for each i. Let Q_γ and Q_α be arbitrary but fixed $|\gamma| \times |\gamma|$ and $|\alpha| \times |\alpha|$ orthogonal matrices, respectively. Define $\widetilde{X}_{\beta\gamma} := Q_\beta^T X_{\beta\gamma} Q_\gamma$ and $\widetilde{Y}_{\beta\alpha} := Q_\beta^T Y_{\beta\alpha} Q_\alpha$. Then, from the expression of $\Gamma(X, Y)$ and its spectral decomposition, it is easy to get that

$$\lambda_i^2 \ge \sum_{k=1}^{|\gamma|} \widetilde{X}_{ik}^2 + \sum_{l=1}^{|\alpha|} \widetilde{Y}_{il}^2 \quad \text{for all} \quad i = 1, \ldots, |\beta|.$$

This means that for $1 \le k \le |\gamma|$, $1 \le l \le |\alpha|$, $1 \le i \le |\beta|$, and $1 \le j \le |\beta|$,

$$\frac{|\widetilde{X}_{ik}|}{\lambda_i + \lambda_j} \le 1, \ \frac{|\widetilde{X}_{kj}|}{\lambda_i + \lambda_j} \le 1, \ \frac{|\widetilde{Y}_{il}|}{\lambda_i + \lambda_j} \le 1, \ \frac{|\widetilde{Y}_{lj}|}{\lambda_i + \lambda_j} \le 1. \tag{11}$$

For any $G, H \in \mathbb{S}^n$, with $\widetilde{G}_{\beta\gamma} = Q_\beta^T G_{\beta\gamma} Q_\gamma$ and $\widetilde{H}_{\beta\alpha} = Q_\beta^T H_{\beta\alpha} Q_\alpha$, we calculate that

$$Q_\beta^T \mathcal{L}^{-1}_{\Gamma(X,Y)}(X_{\beta\gamma}G_{\gamma\beta} + G_{\beta\gamma}X_{\gamma\beta})Q_\beta = \left[\frac{\sum_{k=1}^{|\gamma|}(\widetilde{X}_{ik}\widetilde{G}_{kj} + \widetilde{G}_{ik}\widetilde{X}_{kj})}{\lambda_i + \lambda_j}\right]_{1 \le i,j \le |\beta|},$$
$$Q_\beta^T \mathcal{L}^{-1}_{\Gamma(X,Y)}(Y_{\beta\alpha}H_{\alpha\beta} + H_{\beta\alpha}Y_{\alpha\beta})Q_\beta = \left[\frac{\sum_{l=1}^{|\alpha|}(\widetilde{Y}_{il}\widetilde{H}_{lj} + \widetilde{H}_{il}\widetilde{Y}_{lj})}{\lambda_i + \lambda_j}\right]_{1 \le i,j \le |\beta|}.$$

Using the inequalities in (11) and noting that the Frobenius norm is orthogonally invariant, from the last two equalities we obtain the desired result. ∎

In the subsequent sections, we always use $C\colon \mathbb{S}^n \times \mathbb{S}^n \to \mathbb{S}^n$ to denote the function

$$C(X, Y) := \sqrt{X^2 + Y^2} \quad \forall X, Y \in \mathbb{S}^n, \tag{12}$$

and for any given $X, Y \in \mathbb{S}^n$ assume that $C(X, Y)$ has the spectral decomposition

$$C(X, Y) = P \operatorname{diag}(\lambda_1, \ldots, \lambda_n)P^T = PDP^T, \tag{13}$$

where P is an $n\times n$ orthogonal matrix, and $D=\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$ with $\lambda_i\ge 0$ for all i. Define the index sets κ and β associated with the eigenvalues of $C(X,Y)$ by

$$\kappa \;:=\; \{i:\lambda_i>0\} \quad\text{and}\quad \beta \;:=\; \{i:\lambda_i=0\}\,.$$

Then, by permuting the rows and columns of $C(X,Y)$ if necessary, we may assume that

$$D=\begin{bmatrix} D_\kappa & 0\\ 0 & D_\beta\end{bmatrix}=\begin{bmatrix} D_\kappa & 0\\ 0 & 0\end{bmatrix}.$$

3. Directional derivative and B-subdifferential. The function $\Phi_{\rm FB}$ is directionally differentiable everywhere in $\mathbb{S}^n\times\mathbb{S}^n$; see [21, Corollary 2.3]. But, to our best knowledge, the expression of its directional derivative is not given in the literature. Next we derive it and use it to show that the B-subdifferential of $\Phi_{\rm FB}$ at a general point coincides with that of its directional derivative function at the origin.[1]

PROPOSITION 3.1. *For any given $X,Y\in\mathbb{S}^n$, let $C(X,Y)$ have the spectral decomposition as in* (13). *Then, the directional derivative $\Phi'_{\rm FB}((X,Y);(G,H))$ of $\Phi_{\rm FB}$ at (X,Y) with the direction $(G,H)\in\mathbb{S}^n\times\mathbb{S}^n$ has the expression*

$$(G+H)-P\begin{bmatrix}\mathcal{L}_{D_\kappa}^{-1}\big(\mathcal{L}_{\widetilde{X}_{\kappa\kappa}}(\widetilde{G}_{\kappa\kappa})+\mathcal{L}_{\widetilde{Y}_{\kappa\kappa}}(\widetilde{H}_{\kappa\kappa})\big) & D_\kappa^{-1}(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta}+\widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta})\\ (\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa}+\widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D_\kappa^{-1} & \Theta(\widetilde{G},\widetilde{H})\end{bmatrix}P^T, \tag{14}$$

where $\widetilde{X}:=P^TXP$, $\widetilde{Y}:=P^TYP$, $\widetilde{G}:=P^TGP$, $\widetilde{H}:=P^THP$, and Θ is defined by

$$\begin{aligned}\Theta(U,V):=\Big[&U_{\beta\beta}^2+V_{\beta\beta}^2+U_{\beta\kappa}U_{\kappa\beta}+V_{\beta\kappa}V_{\kappa\beta}\\ &-(U_{\beta\kappa}\widetilde{X}_{\kappa\kappa}+V_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D_\kappa^{-2}(\widetilde{X}_{\kappa\kappa}U_{\kappa\beta}+\widetilde{Y}_{\kappa\kappa}V_{\kappa\beta})\Big]^{1/2}\quad\forall U,V\in\mathbb{S}^n.\end{aligned}\tag{15}$$

Proof. Fix any $G,H\in\mathbb{S}^n$. Assume that $(X,Y)\neq(0,0)$. Then, for any $t>0$, we have

$$\Phi_{\rm FB}(X+tG,Y+tH)-\Phi_{\rm FB}(X,Y)=t(G+H)-\triangle(t) \tag{16}$$

with

$$\triangle(t)\equiv\big[C^2(X,Y)+t(\mathcal{L}_X(G)+\mathcal{L}_Y(H))+t^2(G^2+H^2)\big]^{1/2}-C(X,Y).$$

Let $\widetilde{X},\widetilde{Y},\widetilde{G}$, and $\widetilde{H}$ be defined as in the proposition. It is easy to see that

$$\widetilde{\triangle}(t):=P^T\triangle(t)P=(D^2+\widetilde{W})^{1/2}-D, \tag{17}$$

where

$$\widetilde{W}=t\left(\widetilde{X}\widetilde{G}+\widetilde{G}\widetilde{X}+\widetilde{Y}\widetilde{H}+\widetilde{H}\widetilde{Y}\right)+t^2\left(\widetilde{G}^2+\widetilde{H}^2\right).$$

[1] When we were preparing this manuscript, we learned that these results were obtained by Zhang, Zhang, and Pang (see [26]) via the singular value decomposition. We achieved them independently by eigenvalue decomposition in order to obtain Proposition 3.2.

Since $\widetilde{X}^2 + \widetilde{Y}^2 = D^2$ and $D_\beta = 0$, we have $\widetilde{X} = \mathrm{diag}(\widetilde{X}_{\kappa\kappa}, 0)$ and $\widetilde{Y} = \mathrm{diag}(\widetilde{Y}_{\kappa\kappa}, 0)$. So,

$$\widetilde{W} = t \begin{bmatrix} \mathcal{L}_{\widetilde{X}_{\kappa\kappa}}(\widetilde{G}_{\kappa\kappa}) + \mathcal{L}_{\widetilde{Y}_{\kappa\kappa}}(\widetilde{H}_{\kappa\kappa}) & \widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta} \\ \widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa} & 0 \end{bmatrix} + \begin{bmatrix} o(t) & o(t) \\ o(t) & t^2\left(\widetilde{G}^2_{\beta\beta} + \widetilde{H}^2_{\beta\beta} + \widetilde{G}_{\beta\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{H}_{\beta\kappa}\widetilde{H}_{\kappa\beta}\right) \end{bmatrix}.$$

By (17) and [24, Lemma 6.2], we know that

$$\begin{cases} \widetilde{\triangle}(t)_{\kappa\kappa} = \mathcal{L}^{-1}_{D_\kappa}(\widetilde{W}_{\kappa\kappa}) + o(\|\widetilde{W}\|), \\ \widetilde{\triangle}(t)_{\kappa\beta} = D_\kappa^{-1}\widetilde{W}_{\kappa\beta} + o(\|\widetilde{W}\|), \\ \widetilde{W}_{\beta\beta} = \widetilde{\triangle}(t)^T_{\kappa\beta}\widetilde{\triangle}(t)_{\kappa\beta} + \widetilde{\triangle}(t)^2_{\beta\beta}. \end{cases} \tag{18}$$

From the second equality of (18) and the expression of $\widetilde{W}_{\kappa\beta}$, it follows that

$$\widetilde{\triangle}(t)_{\kappa\beta} = tD_\kappa^{-1}\left(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta}\right) + o(t), \tag{19}$$

and consequently,

$$\widetilde{\triangle}(t)^T_{\kappa\beta}\widetilde{\triangle}(t)_{\kappa\beta} = t^2\left(\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa}\right)D_\kappa^{-2}\left(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta}\right) + o(t^2).$$

This, together with the third equation of (18) and the expression of $\widetilde{W}_{\beta\beta}$, implies that

$$\begin{aligned} \widetilde{\triangle}(t)^2_{\beta\beta} = {} & t^2\left(\widetilde{G}_{\beta\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{H}_{\beta\kappa}\widetilde{H}_{\kappa\beta} + \widetilde{G}^2_{\beta\beta} + \widetilde{H}^2_{\beta\beta}\right) \\ & - t^2\left(\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa}\right)D_\kappa^{-2}\left(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta}\right) + o(t^2). \end{aligned}$$

Since $D_\beta = 0$, the expression of $\widetilde{\triangle}(t)$ in (17) implies that $\widetilde{\triangle}(t)_{\beta\beta} \succeq 0$. Therefore,

$$\lim_{t\downarrow 0}\frac{\widetilde{\triangle}(t)_{\beta\beta}}{t} = \lim_{t\downarrow 0}\frac{[\widetilde{\triangle}(t)^2_{\beta\beta}]^{1/2}}{t} = \Theta(\widetilde{G}, \widetilde{H}).$$

In addition, from the first equation in (18) and the expression of $\widetilde{W}_{\kappa\kappa}$, we have

$$\widetilde{\triangle}(t)_{\kappa\kappa} = t\mathcal{L}^{-1}_{D_\kappa}\left(\mathcal{L}_{\widetilde{X}_{\kappa\kappa}}(\widetilde{G}_{\kappa\kappa}) + \mathcal{L}_{\widetilde{Y}_{\kappa\kappa}}(\widetilde{H}_{\kappa\kappa})\right) + o(t).$$

Combining the last two equations with (19), we immediately obtain that

$$\lim_{t\downarrow 0}\frac{\widetilde{\triangle}(t)}{t} = \begin{bmatrix} \mathcal{L}^{-1}_{D_\kappa}\left(\mathcal{L}_{\widetilde{X}_{\kappa\kappa}}(\widetilde{G}_{\kappa\kappa}) + \mathcal{L}_{\widetilde{Y}_{\kappa\kappa}}(\widetilde{H}_{\kappa\kappa})\right) & D_\kappa^{-1}(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta}) \\ (\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D_\kappa^{-1} & \Theta(\widetilde{G}, \widetilde{H}) \end{bmatrix}.$$

This, along with (16), shows that $\Phi'_{\rm FB}((X,Y);(G,H))$ has the expression given by (14).

When $(X,Y) = (0,0)$, by the positive homogeneity of $\Phi_{\rm FB}$, we immediately have

$$\Phi'_{\rm FB}((X,Y);(G,H)) = (G+H) - \sqrt{G^2 + H^2} = \Phi_{\rm FB}(G,H).$$

Note that this is a special case of (14) with $\kappa = \emptyset$. The result then follows. $\square$

Note that the function Θ in (15) is always well defined since, by Lemma 2.2,

$$U_{\beta\kappa}U_{\kappa\beta} + V_{\beta\kappa}V_{\kappa\beta} - (U_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + V_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D_\kappa^{-2}(\widetilde{X}_{\kappa\kappa}U_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}V_{\kappa\beta}) \succeq 0$$

for all $U, V \in \mathbb{S}^n$. As a consequence of Proposition 3.1, we readily obtain the following necessary and sufficient characterization for the differentiable points of the function $\Phi_{\rm FB}$.

COROLLARY 3.1. *The function $\Phi_{\rm FB}$ is F-differentiable at (X, Y) if and only if $C(X, Y) \succ 0$. Furthermore, when $C(X, Y) \succ 0$, we have for any $(G, H) \in \mathbb{S}^n \times \mathbb{S}^n$*

$$\mathcal{J}\Phi_{\rm FB}(X, Y)(G, H) = (G + H) - \mathcal{L}_{C(X,Y)}^{-1}\left(\mathcal{L}_X(G) + \mathcal{L}_Y(H)\right). \tag{20}$$

Proof. The "if" part is direct by [1, Theorem V.3.3] or [5, Proposition 4.3]. We next prove the "only if" part by contradiction. Suppose that $\Phi_{\rm FB}$ is F-differentiable at (X, Y), but $C(X, Y) \succ 0$ does not hold. Then $|\beta| \neq \emptyset$. Since $\Phi_{\rm FB}$ is F-differentiable at (X, Y), $\Phi'_{\rm FB}((X, Y); (\cdot, \cdot))$ is a linear operator. But, letting $(G_1, H_1), (G_2, H_2) \in \mathbb{S}^n \times \mathbb{S}^n$ be such that $G_1 = G_2 = 0$, $H_1 = \mathrm{diag}(0, I_{|\beta|})$, and $H_2 = -H_1$, we obtain that

$$\begin{aligned} 0 &= \Phi'_{\rm FB}((X, Y); (G_1, H_1) + (G_2, H_2)) \\ &= \Phi'_{\rm FB}((X, Y); (G_1, H_1)) + \Phi'_{\rm FB}((X, Y); (G_2, H_2)) \\ &= -2P\begin{pmatrix} 0 & 0 \\ 0 & I_{|\beta|} \end{pmatrix} P^T, \end{aligned}$$

which is a contradiction. This contradiction shows that the "only if" part holds. The formula in (20) follows by [4, Lemma 2] or [11, 12, Theorem 3.4]. $\square$

Next we derive the expression of the directional derivative of Θ at (U, V) with the direction $(G, H) \in \mathbb{S}^n \times \mathbb{S}^n$, which is used to characterize the F-differentiable points of Θ in Lemma 3.2 below. Define $\Omega_1 : \mathbb{S}^n \times \mathbb{S}^n \to \mathbb{R}^{|\beta|\times|\kappa|}$ and $\Omega_2 : \mathbb{S}^n \times \mathbb{S}^n \to \mathbb{R}^{|\beta|\times|\kappa|}$ by

$$\Omega_1(U, V) := U_{\beta\kappa} - (U_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + V_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D_\kappa^{-2}\widetilde{X}_{\kappa\kappa} \quad \forall U, V \in \mathbb{S}^n$$

and

$$\Omega_2(U, V) := V_{\beta\kappa} - (U_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + V_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D_\kappa^{-2}\widetilde{Y}_{\kappa\kappa} \quad \forall U, V \in \mathbb{S}^n,$$

respectively. Noting that $\widetilde{X}_{\kappa\kappa}^2 + \widetilde{Y}_{\kappa\kappa}^2 = D_\kappa^2$, we can rewrite the function Θ in (15) as

$$\Theta(U, V) = \left[U_{\beta\beta}^2 + V_{\beta\beta}^2 + \Omega_1(U, V)\Omega_1(U, V)^T + \Omega_2(U, V)\Omega_2(U, V)^T\right]^{1/2} \ \forall U, V \in \mathbb{S}^n. \tag{21}$$

For any given $U, V \in \mathbb{S}^n$, assume that $\Theta(U, V)$ has the spectral decomposition

$$\Theta(U, V) = R\Lambda R^T = R\,\mathrm{diag}(\vartheta_1, \ldots, \vartheta_{|\beta|})R^T,$$

where $\Lambda = \mathrm{diag}(\vartheta_1, \ldots, \vartheta_{|\beta|})$ is the diagonal matrix of eigenvalues of $\Theta(U, V)$ and R is a corresponding matrix of orthonormal eigenvectors. Define the index sets I and J associated with the eigenvalues of $\Theta(U, V)$ by

$$I := \{i : \vartheta_i > 0\} \quad \text{and} \quad J := \{i : \vartheta_i = 0\}.$$

Then, by permuting the rows and columns of $\Theta(U,V)$ if necessary, we may assume that

$$\Lambda = \begin{bmatrix} \Lambda_I & 0 \\ 0 & \Lambda_J \end{bmatrix} = \begin{bmatrix} \Lambda_I & 0 \\ 0 & 0 \end{bmatrix}.$$

From (21) and the spectral decomposition of $\Theta(U,V)$, it is easy to obtain that

$$[R^T U_{\beta\beta}]_{J\beta} = 0,\ [R^T V_{\beta\beta}]_{J\beta} = 0,\ [R^T \Omega_1(U,V)]_{J\kappa} = 0,\ [R^T \Omega_2(U,V)]_{J\kappa} = 0. \tag{22}$$

LEMMA 3.1. *For any given $(U,V) \in \mathbb{S}^n \times \mathbb{S}^n$, assume that $\Theta(U,V)$ has the spectral decomposition as above. Then, the directional derivative $\Theta'((U,V);(G,H))$ of Θ at (U,V) with the direction $(G,H) \in \mathbb{S}^n \times \mathbb{S}^n$ has the expression*

$$R \begin{bmatrix} \mathcal{L}_{\Lambda_I}^{-1}[\widetilde{W}_{II}] & \Lambda_I^{-1}\widetilde{W}_{IJ} \\ \widetilde{W}_{IJ}^T \Lambda_I^{-1} & (\widetilde{\Theta}_{JJ} - \widetilde{W}_{IJ}^T \Lambda_I^{-2} \widetilde{W}_{IJ})^{1/2} \end{bmatrix} R^T, \tag{23}$$

where $\widetilde{\Theta} := R^T \Theta^2(G,H) R$ and $\widetilde{W} := R^T W(G,H) R$ with $W(G,H)$ given by

$$\begin{aligned} W(G,H) := {} & \Omega_1(U,V)\Omega_1(G,H)^T + \Omega_1(G,H)\Omega_1(U,V)^T + \mathcal{L}_{U_{\beta\beta}}(G_{\beta\beta}) \\ & + \mathcal{L}_{V_{\beta\beta}}(H_{\beta\beta}) + \Omega_2(U,V)\Omega_2(G,H)^T + \Omega_2(G,H)\Omega_2(U,V)^T. \end{aligned}$$

Proof. Assume that $\Theta(U,V) \neq 0$. For any $t > 0$, we calculate that

$$\begin{aligned} \Delta(t) &:= \Theta(U + tG, V + tH) - \Theta(U,V) \\ &= \left[\Theta^2(U,V) + tW(G,H) + t^2\Theta^2(G,H)\right]^{1/2} - \Theta(U,V). \end{aligned}$$

From the spectral decomposition of $\Theta(U,V)$, it then follows that

$$\widetilde{\Delta}(t) := R^T \Delta(t) R = \left(\Lambda^2 + t\widetilde{W} + t^2\widetilde{\Theta}\right)^{1/2} - \Lambda, \tag{24}$$

where $\widetilde{\Theta}$ and $\widetilde{W}$ are defined as in the lemma. From (24) and [24, Lemma 6.2], we have

$$\begin{cases} \widetilde{\Delta}(t)_{II} = t\mathcal{L}_{\Lambda_I}^{-1}[\widetilde{W}_{II}] + o(t), \\ \widetilde{\Delta}(t)_{IJ} = t\Lambda_I^{-1}\widetilde{W}_{IJ} + o(t), \\ t\widetilde{W}_{JJ} + t^2\widetilde{\Theta}_{JJ} = \widetilde{\Delta}(t)_{IJ}^T\widetilde{\Delta}(t)_{IJ} + \widetilde{\Delta}(t)_{JJ}^2. \end{cases} \tag{25}$$

By (22) and the definition of $\widetilde{W}$, we have $\widetilde{W}_{JJ} = 0$. Then, from the last two equalities of (25), it follows that

$$\widetilde{\Delta}(t)_{JJ}^2 = t^2\widetilde{\Theta}_{JJ} - \widetilde{\Delta}(t)_{IJ}^T\widetilde{\Delta}(t)_{IJ} = t^2\left(\widetilde{\Theta}_{JJ} - \widetilde{W}_{IJ}^T\Lambda_I^{-2}\widetilde{W}_{IJ}\right) + o(t^2).$$

Since $\Lambda_J = 0$, the expression of $\widetilde{\Delta}(t)$ in (24) implies that $\widetilde{\Delta}(t)_{JJ} \succeq 0$. Therefore,

$$\lim_{t\downarrow 0} \frac{\widetilde{\Delta}(t)_{JJ}}{t} = \lim_{t\downarrow 0} \frac{\sqrt{\widetilde{\Delta}(t)_{JJ}^2}}{t} = \left(\widetilde{\Theta}_{JJ} - \widetilde{W}_{IJ}^T\Lambda_I^{-2}\widetilde{W}_{IJ}\right)^{1/2}.$$

This, together with the first two equalities of (25), yields that

$$\Theta'((U,V);(G,H)) = \lim_{t\downarrow 0} \frac{R\widetilde{\Delta}(t)R^T}{t} = R \begin{bmatrix} \mathcal{L}_{\Lambda_I}^{-1}[\widetilde{W}_{II}] & \Lambda_I^{-1}\widetilde{W}_{IJ} \\ \widetilde{W}_{IJ}^T\Lambda_I^{-1} & (\widetilde{\Theta}_{JJ} - \widetilde{W}_{IJ}^T\Lambda_I^{-2}\widetilde{W}_{IJ})^{1/2} \end{bmatrix} R^T.$$

If $\Theta(U,V)=0$, then $U_{\beta\beta}=0, V_{\beta\beta}=0$, $\Omega_1(U,V)=0$, and $\Omega_2(U,V)=0$. By this, it is easy to compute that $\Theta'((U,V);(G,H))=\Theta(G,H)$. Note that $\Theta(G,H)$ is a special case of (23) with $I=\emptyset$. The result then follows. ◻

Remark 3.1. Lemma 3.1 shows that the function Θ defined by (15) is directionally differentiable everywhere in $\mathbb{S}^n\times\mathbb{S}^n$. In fact, Θ is also globally Lipschitz continuous and strongly semismooth in $\mathbb{S}^n\times\mathbb{S}^n$. Let $\Psi(U,V):=[U_{\beta\beta}\ \ V_{\beta\beta}\ \ \Omega_1(U,V)\ \ \Omega_2(U,V)]$ for $U,V\in\mathbb{S}^n$, and $G^{\rm mat}(A):=\sqrt{AA^T}$ for $A\in\mathbb{R}^{|\beta|\times 2n}$. Comparing with (21), we have that $\Theta(U,V)\equiv G^{\rm mat}(\Psi(U,V))$. By [21, Theorem 2.2], $G^{\rm mat}$ is globally Lipschitz continuous and strongly semismooth everywhere in $\mathbb{R}^{|\beta|\times 2n}$. Since Ψ is a linear function, the composition of $G^{\rm mat}$ and Ψ, i.e., the function Θ, is globally Lipschitz continuous and strongly semismooth everywhere in $\mathbb{S}^n\times\mathbb{S}^n$ by [7, Theorem 19].

By the expression of the directional derivative of Θ, we may present the necessary and sufficient characterization for the differentiable points of Θ.

LEMMA 3.2. *The function Θ is F-differentiable at (U,V) if and only if $\Theta(U,V)\succ 0$. Furthermore, when $\Theta(U,V)\succ 0$, we have for any $(G,H)\in\mathbb{S}^n\times\mathbb{S}^n$,*

$$\begin{aligned}\mathcal{J}\Theta(U,V)(G,H)=\mathcal{L}^{-1}_{\Theta(U,V)}\Big[&(U_{\beta\kappa}G_{\kappa\beta}+G_{\beta\kappa}U_{\kappa\beta})+(V_{\beta\kappa}H_{\kappa\beta}+H_{\beta\kappa}V_{\kappa\beta})\\&-\Big(G_{\beta\kappa}\widetilde{X}_{\kappa\kappa}+H_{\beta\kappa}\widetilde{Y}_{\kappa\kappa}\Big)D_\kappa^{-2}\Big(\widetilde{X}_{\kappa\kappa}U_{\kappa\beta}+\widetilde{Y}_{\kappa\kappa}V_{\kappa\beta}\Big)\\&-\Big(U_{\beta\kappa}\widetilde{X}_{\kappa\kappa}+V_{\beta\kappa}\widetilde{Y}_{\kappa\kappa}\Big)D_\kappa^{-2}\Big(\widetilde{X}_{\kappa\kappa}G_{\kappa\beta}+\widetilde{Y}_{\kappa\kappa}H_{\kappa\beta}\Big)\\&+\mathcal{L}_{U_{\beta\beta}}(G_{\beta\beta})+\mathcal{L}_{V_{\beta\beta}}(H_{\beta\beta})\Big].\end{aligned}$$

Proof. We need only prove the "only if" part. If Θ is F-differentiable at (U,V), then $\Theta'((U,V);(G,H))$ is a linear function of (G,H) which, by (23) implies that $(\widetilde{\Theta}_{JJ}-\widetilde{W}_{IJ}^T\Lambda_I^{-2}\widetilde{W}_{IJ})^{1/2}$ is a linear function of (G,H). We next argue that this holds true only if $J=\emptyset$. Indeed, if $J\neq\emptyset$, by taking $G=\left[\begin{smallmatrix}0&0\\0&G_{\beta\beta}\end{smallmatrix}\right]$ and $H=\left[\begin{smallmatrix}0&0\\0&H_{\beta\beta}\end{smallmatrix}\right]$ with $G_{\beta\beta}\succ 0$ and $H_{\beta\beta}\succ 0$, we have $\Omega_1(G,H)=0$ and $\Omega_2(G,H)=0$, which, together with $[R^TU_{\beta\beta}]_{J\beta}=0$ and $[R^TV_{\beta\beta}]_{J\beta}=0$, implies that $\widetilde{W}_{JI}=[R^TW(G,H)R]_{JI}=0$. Note that $\Theta^2(G,H)=G_{\beta\beta}^2+H_{\beta\beta}^2$. Then, $(\widetilde{\Theta}_{JJ}-\widetilde{W}_{IJ}^T\Lambda_I^{-2}\widetilde{W}_{IJ})^{1/2}=\sqrt{[R^T(G_{\beta\beta}^2+H_{\beta\beta}^2)R]_{JJ}}$, which is clearly nonlinear. The Jacobian formula of Θ is direct by a simple computation. ◻

Remark 3.2. Combining Proposition 3.1 with Lemma 3.2, we immediately obtain that $\Phi'_{\rm FB}((X,Y);(\cdot,\cdot))$ is F-differentiable at (G,H) if and only if $\Theta(P^TGP,P^THP)\succ 0$.

By the definition of Θ and Lemma 3.2, we can prove the following result (see the appendix for the proof), which corresponds to the property of $\Phi_{\rm NR}$ in [14, Lemma 11].

LEMMA 3.3. *For any given $X,Y\in\mathbb{S}^n$, let $\Psi_{\rm FB}(\cdot,\cdot)\equiv\Phi'_{\rm FB}((X,Y);(\cdot,\cdot))$. Then,*

$$\partial_B\Phi_{\rm FB}(X,Y)=\partial_B\Psi_{\rm FB}(0,0).$$

Now Lemma 3.3 and Proposition 3.1 allow us to obtain the main result of this section.

PROPOSITION 3.2. *For any given $X,Y\in\mathbb{S}^n$, let $C(X,Y)$ have the spectral decomposition as in* (13). *Then, a $(\mathcal{U},\mathcal{V})\in\partial_B\Phi_{\rm FB}(X,Y)$ (respectively, $\partial\Phi_{\rm FB}(X,Y)$) if and only if there exists a $(\mathcal{G},\mathcal{H})\in\partial_B\Theta(0,0)$ (respectively, $\partial\Theta(0,0)$) such that for* *any* $G, H \in \mathbb{S}^n$,

$$(\mathcal{I}-\mathcal{U})(G)+(\mathcal{I}-\mathcal{V})(H)$$
$$= P\begin{bmatrix} \mathcal{L}_{D_\kappa}^{-1}\left(\mathcal{L}_{\widetilde{X}_{\kappa\kappa}}(\widetilde{G}_{\kappa\kappa})+\mathcal{L}_{\widetilde{Y}_{\kappa\kappa}}(\widetilde{H}_{\kappa\kappa})\right) & D_\kappa^{-1}\left(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta}+\widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta}\right) \\ \left(\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa}+\widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa}\right)D_\kappa^{-1} & \mathcal{G}(\widetilde{G})+\mathcal{H}(\widetilde{H}) \end{bmatrix}P^T, \tag{26}$$

where $\widetilde{X} := P^TXP$, $\widetilde{Y} := P^TYP$, $\widetilde{G} := P^TGP$, *and* $\widetilde{H} := P^THP$.

Proof. For any $G, H \in \mathbb{S}^n$, let $\Psi(G,H) := (P^TGP, P^THP)$. Define $\Xi\colon \mathbb{S}^n\times\mathbb{S}^n \to \mathbb{S}^n$ by

$$\Xi(S,T) := P\begin{bmatrix} \mathcal{L}_{D_\kappa}^{-1}\left(\mathcal{L}_{\widetilde{X}_{\kappa\kappa}}(S_{\kappa\kappa})+\mathcal{L}_{\widetilde{Y}_{\kappa\kappa}}(T_{\kappa\kappa})\right) & D_\kappa^{-1}(\widetilde{X}_{\kappa\kappa}S_{\kappa\beta}+\widetilde{Y}_{\kappa\kappa}T_{\kappa\beta}) \\ (S_{\beta\kappa}\widetilde{X}_{\kappa\kappa}+T_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D_\kappa^{-1} & \Theta(S,T) \end{bmatrix}P^T.$$

By Proposition 3.1, clearly, $\Psi_{\rm FB}(G,H) = (G+H)-\Xi(\Psi(G,H))$ for any $G,H\in\mathbb{S}^n$. Note that Ξ is globally Lipschitz continuous in $\mathbb{S}^n\times\mathbb{S}^n$ by the remarks after (21), and $\mathcal{J}\Psi(G,H)$ for any $G,H\in\mathbb{S}^n$ is onto. Applying [3, Lemma 2.1] to the composite mapping $\Xi\circ\Psi$ at $(0,0)$, we have that $\partial_B(\Xi\circ\Psi)(0,0) = \partial_B\Xi(\Psi(0,0))\mathcal{J}\Psi(0,0) = \partial_B\Xi(0,0)\Psi$. So,

$$\partial_B\Psi_{\rm FB}(0,0) = (\mathcal{I},\mathcal{I}) - \partial_B\Xi(0,0)\Psi.$$

This, together with Lemma 3.3 and the expression of Ξ, completes the proof. ∎

4. Nonsingularity conditions.

This section will show that the Clarke's Jacobian of $E_{\rm FB}$ at a KKT point is nonsingular if and only if the KKT point is a strongly regular solution to the generalized equation

$$0 \in \begin{bmatrix} \mathcal{J}_{x,X}L(x,X,\mu,S,Y) \\ h(x) \\ g(x)-X \\ X \end{bmatrix} + \begin{bmatrix} \mathcal{N}_{\mathbb{X}\times\mathbb{S}^n}(x,X) \\ \mathcal{N}_{\mathbb{R}^m}(\mu) \\ \mathcal{N}_{\mathbb{S}^n}(S) \\ \mathcal{N}_{\mathbb{S}^n_+}(Y) \end{bmatrix}. \tag{27}$$

Let $(\overline{x},\overline{X},\overline{\mu},\overline{S},\overline{Y}) \in \mathbb{X}\times\mathbb{S}^n_+\times\mathbb{R}^m\times\mathbb{S}^n\times\mathbb{S}^n_+$ be a KKT point of the NLSDP (2), i.e., a point satisfying the KKT condition (3). Let $\overline{C}\equiv C(\overline{X},\overline{Y})$. Noting that

$$\overline{X}\in\mathbb{S}^n_+,\quad \overline{Y}\in\mathbb{S}^n_+,\ \text{ and }\ \langle\overline{X},\overline{Y}\rangle = 0, \tag{28}$$

we may assume that $\overline{C}$ has the spectral decomposition as in (13) with $\kappa=\alpha\cup\gamma$,

$$\overline{X} = P\begin{bmatrix} D_\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}P^T,\ \text{ and }\ \overline{Y} = P\begin{bmatrix} 0 & 0 & 0 \\ 0 & D_\gamma & 0 \\ 0 & 0 & 0 \end{bmatrix}P^T. \tag{29}$$

By this, we write $P = [P_\alpha\ \ P_\gamma\ \ P_\beta]$ with $P_\alpha\in\mathbb{R}^{n\times|\alpha|}$, $P_\gamma\in\mathbb{R}^{n\times|\gamma|}$, and $P_\beta\in\mathbb{R}^{n\times|\beta|}$.

By the spectral decomposition of $\overline{X}$ and $\overline{Y}$, we can simplify the function Θ involved in $\Phi'_{\rm FB}((\overline{X},\overline{Y});(\cdot,\cdot))$ as the function $\Gamma\colon\mathbb{S}^n\times\mathbb{S}^n\to\mathbb{S}^{|\beta|}$ defined by (10) with the above α,γ, and β. In view of this, we first characterize a property of Clarke's Jacobian of Γ at a general point, which will be used to prove Proposition 4.1 below. Particularly, it also implies the property of Clarke's Jacobian of $\Phi_{\rm FB}$ at a general point; see Remark 4.1.

LEMMA 4.1. *For any given $(U,V)\in\mathbb{S}^n\times\mathbb{S}^n$, let $(\mathcal{G},\mathcal{H})\in\partial\Gamma(U,V)$. Then, we have*

$$\|\mathcal{G}(G)+\mathcal{H}(H)\|\le\|(G_{\beta\beta},H_{\beta\beta})\|+2\sqrt{|\beta||\gamma|}\,\|G_{\beta\gamma}\|+2\sqrt{|\beta||\alpha|}\,\|H_{\alpha\beta}\|\quad\forall G,H\in\mathbb{S}^n.$$

Proof. Let $(\mathcal{G},\mathcal{H})\in\partial\Gamma(U,V)$. By Carathéodory's theorem, there exist a positive integer l and $(\mathcal{G}^i,\mathcal{H}^i)\in\partial_B\Gamma(U,V)$ for $i=1,\ldots,l$ such that $(\mathcal{G},\mathcal{H})=\sum_{i=1}^{l}\nu_i(\mathcal{G}^i,\mathcal{H}^i)$, where $\sum_{i=1}^{l}\nu_i=1$ and $\nu_i\ge 0,\ i=1,\ldots,l$. From Lemma 3.2, we know that Γ is F-differentiable at (U,V) if and only if $\Gamma(U,V)\succ 0$. Also, when $\Gamma(U,V)\succ 0$, we have

$$\begin{aligned}\mathcal{J}\Gamma(U,V)(G,H)=\mathcal{L}^{-1}_{\Gamma(U,V)}\big[&\mathcal{L}_{U_{\beta\beta}}(G_{\beta\beta})+\mathcal{L}_{V_{\beta\beta}}(H_{\beta\beta})+U_{\beta\gamma}G_{\gamma\beta}\\&+G_{\beta\gamma}U_{\gamma\beta}+V_{\beta\alpha}H_{\alpha\beta}+H_{\beta\alpha}V_{\alpha\beta}\big]\end{aligned}$$

for any $G,H\in\mathbb{S}^n$. Hence, for each $i\in\{1,\ldots,l\}$, by the definition of the elements in $\partial_B\Gamma(U,V)$, there exists a sequence $\{(U^{i_k},V^{i_k})\}$ in $\mathbb{S}^n\times\mathbb{S}^n$ converging to (U,V) with $\Gamma(U^{i_k},V^{i_k})\succ 0$ such that $(\mathcal{G}^i,\mathcal{H}^i)=\lim_{k\to\infty}\mathcal{J}\Gamma(U^{i_k},V^{i_k})$. Thus, for any $G,H\in\mathbb{S}^n$,

$$\begin{aligned}\mathcal{G}(G)+\mathcal{H}(H)=\lim_{k\to\infty}\sum_{i=1}^{l}\nu_i\mathcal{L}^{-1}_{\Gamma(U^{i_k},V^{i_k})}\Big[&\mathcal{L}_{U^{i_k}_{\beta\beta}}(G_{\beta\beta})+\mathcal{L}_{V^{i_k}_{\beta\beta}}(H_{\beta\beta})+U^{i_k}_{\beta\gamma}G_{\gamma\beta}\\&+G_{\beta\gamma}U^{i_k}_{\gamma\beta}+V^{i_k}_{\beta\alpha}H_{\alpha\beta}+H_{\beta\alpha}V^{i_k}_{\alpha\beta}\Big].\end{aligned}$$

Together with the continuity and convexity of $\|\cdot\|$, it follows that

$$\begin{aligned}&\|\mathcal{G}(G)+\mathcal{H}(H)\|\\&\quad\le\lim_{k\to\infty}\sum_{i=1}^{l}\nu_i\Big\|\mathcal{L}^{-1}_{\Gamma(U^{i_k},V^{i_k})}\Big[\mathcal{L}_{U^{i_k}_{\beta\beta}}(G_{\beta\beta})+\mathcal{L}_{V^{i_k}_{\beta\beta}}(H_{\beta\beta})\\&\qquad\qquad+U^{i_k}_{\beta\gamma}G_{\gamma\beta}+G_{\beta\gamma}U^{i_k}_{\gamma\beta}+V^{i_k}_{\beta\alpha}H_{\alpha\beta}+H_{\beta\alpha}V^{i_k}_{\alpha\beta}\Big]\Big\|\\&\quad\le\lim_{k\to\infty}\sum_{i=1}^{l}\nu_i\Big\{\Big\|\mathcal{L}^{-1}_{\Gamma(U^{i_k},V^{i_k})}\Big[\mathcal{L}_{U^{i_k}_{\beta\beta}}(G_{\beta\beta})+\mathcal{L}_{V^{i_k}_{\beta\beta}}(H_{\beta\beta})\Big]\Big\|\\&\qquad\qquad+\Big\|\mathcal{L}^{-1}_{\Gamma(U^{i_k},V^{i_k})}\Big(U^{i_k}_{\beta\gamma}G_{\gamma\beta}+G_{\beta\gamma}U^{i_k}_{\gamma\beta}+V^{i_k}_{\beta\alpha}H_{\alpha\beta}+H_{\beta\alpha}V^{i_k}_{\alpha\beta}\Big)\Big\|\Big\}.\end{aligned}$$

For each i and k, from Propositions 2.2 and 2.3, we have that

$$\Big\|\mathcal{L}^{-1}_{\Gamma(U^{i_k},V^{i_k})}\Big[\mathcal{L}_{U^{i_k}_{\beta\beta}}(G_{\beta\beta})+\mathcal{L}_{V^{i_k}_{\beta\beta}}(H_{\beta\beta})\Big]\Big\|\le\|(G_{\beta\beta},H_{\beta\beta})\|$$

and

$$\begin{aligned}&\Big\|\mathcal{L}^{-1}_{\Gamma(U^{i_k},V^{i_k})}\Big(U^{i_k}_{\beta\gamma}G_{\gamma\beta}+G_{\beta\gamma}U^{i_k}_{\gamma\beta}+V^{i_k}_{\beta\alpha}H_{\alpha\beta}+H_{\beta\alpha}V^{i_k}_{\alpha\beta}\Big)\Big\|\\&\quad\le 2\sqrt{|\beta||\gamma|}\,\|G_{\gamma\beta}\|+2\sqrt{|\beta||\alpha|}\,\|H_{\alpha\beta}\|.\end{aligned}$$

From the last three inequalities, we immediately obtain that for any $G,H\in\mathbb{S}^n$,

$$\|\mathcal{G}(G)+\mathcal{H}(H)\|\le\|(G_{\beta\beta},H_{\beta\beta})\|+2\sqrt{|\beta||\gamma|}\,\|G_{\gamma\beta}\|+2\sqrt{|\beta||\alpha|}\,\|H_{\alpha\beta}\|.$$

Thus, we complete the proof. ∎

Remark 4.1. When $\alpha \cup \gamma = \emptyset$, the function $\Gamma(\cdot,\cdot)$ reduces to $C(\cdot,\cdot)$ defined in (12). Then, Lemma 4.1 characterizes the following property of Clarke's Jacobian of $\Phi_{\rm FB}$ at a general point: *for any given* $X, Y \in \mathbb{S}^n$*, letting* $(\mathcal{U},\mathcal{V}) \in \partial\Phi_{\rm FB}(X,Y)$*, it holds that*

$$\mathcal{U}(G) + \mathcal{V}(H) = 0 \implies \langle G, H\rangle \le 0.$$

We achieve the main result of this section by two steps: (1) we show that the strong second order sufficient condition and constraint nondegeneracy of $(\overline{x},\overline{X})$ implies the nonsingularity of $\partial E_{\rm FB}(\overline{x},\overline{X},\overline{\mu},\overline{S},\overline{Y})$. (2) We establish the relationship between Clarke's Jacobians $\partial E_{\rm FB}(\overline{x},\overline{X},\overline{\mu},\overline{S},\overline{Y})$ and $\partial E_{\rm NR}(\overline{x},\overline{X},\overline{\mu},\overline{S},\overline{Y})$. The first step needs the following two propositions, which provide the properties of the elements in $\partial\Phi_{\rm FB}(\overline{X},\overline{Y})$.

PROPOSITION 4.1. *Let* $\overline{X},\overline{Y} \in \mathbb{S}^n$ *satisfy* (28)*, and assume that they have the spectral decomposition as in* (29)*. Then, for any* $(\mathcal{U},\mathcal{V}) \in \partial\Phi_{\rm FB}(\overline{X},\overline{Y})$*, it holds that*

$$\mathcal{U}(G) + \mathcal{V}(H) = 0 \implies \begin{cases} P_\beta^T G P_\gamma = 0,\ P_\gamma^T G P_\gamma = 0,\ P_\alpha^T H P_\alpha = 0, \\ P_\alpha^T H P_\beta = 0, P_\alpha^T G P_\gamma D_\gamma + D_\alpha P_\alpha^T H P_\gamma = 0, \\ \langle P_\beta^T G P_\beta, P_\beta^T H P_\beta\rangle \le 0. \end{cases} \tag{30}$$

Proof. Fix any $(\mathcal{U},\mathcal{V}) \in \partial\Phi_{\rm FB}(\overline{X},\overline{Y})$ and $G, H \in \mathbb{S}^n$ with $\mathcal{U}(G) + \mathcal{V}(H) = 0$. Applying Proposition 3.2, there exists a $(\mathcal{G},\mathcal{H}) \in \partial\Gamma(0,0)$ such that

$$\widetilde{G} + \widetilde{H} = \begin{bmatrix} \mathcal{L}_{D_\kappa}^{-1}\left(\mathcal{L}_{\widetilde{X}_{\kappa\kappa}}(\widetilde{G}_{\kappa\kappa}) + \mathcal{L}_{\widetilde{Y}_{\kappa\kappa}}(\widetilde{H}_{\kappa\kappa})\right) & D_\kappa^{-1}(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta}) \\ (\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D_\kappa^{-1} & \mathcal{G}(\widetilde{G}) + \mathcal{H}(\widetilde{H}) \end{bmatrix},$$

where $\kappa = \alpha \cup \gamma$, $\widetilde{X} = P^T\overline{X}P$, $\widetilde{Y} = P^T\overline{Y}P$, $\widetilde{G} = P^TGP$ and $\widetilde{H} = P^THP$. By (29), an elementary calculation shows that the last equality can be rewritten as

$$\begin{bmatrix} \mathcal{L}_{D_\alpha}(\widetilde{G}_{\alpha\alpha} + \widetilde{H}_{\alpha\alpha}) & D_\alpha(\widetilde{G}_{\alpha\gamma} + \widetilde{H}_{\alpha\gamma}) + (\widetilde{G}_{\alpha\gamma} + \widetilde{H}_{\alpha\gamma})D_\gamma & \widetilde{G}_{\alpha\beta} + \widetilde{H}_{\alpha\beta} \\ (\widetilde{G}_{\gamma\alpha} + \widetilde{H}_{\gamma\alpha})D_\alpha + D_\gamma(\widetilde{G}_{\gamma\alpha} + \widetilde{H}_{\gamma\alpha}) & \mathcal{L}_{D_\gamma}(\widetilde{G}_{\gamma\gamma} + \widetilde{H}_{\gamma\gamma}) & \widetilde{G}_{\gamma\beta} + \widetilde{H}_{\gamma\beta} \\ \widetilde{G}_{\beta\alpha} + \widetilde{H}_{\beta\alpha} & \widetilde{G}_{\beta\gamma} + \widetilde{H}_{\beta\gamma} & \widetilde{G}_{\beta\beta} + \widetilde{H}_{\beta\beta} \end{bmatrix}$$

$$= \begin{bmatrix} \mathcal{L}_{D_\alpha}(\widetilde{G}_{\alpha\alpha}) & D_\alpha\widetilde{G}_{\alpha\gamma} + \widetilde{H}_{\alpha\gamma}D_\gamma & \widetilde{G}_{\alpha\beta} \\ \widetilde{G}_{\gamma\alpha}D_\alpha + D_\gamma\widetilde{H}_{\gamma\alpha} & \mathcal{L}_{D_\gamma}(\widetilde{H}_{\gamma\gamma}) & \widetilde{H}_{\gamma\beta} \\ \widetilde{G}_{\beta\alpha} & \widetilde{H}_{\beta\gamma} & \mathcal{G}(\widetilde{G}) + \mathcal{H}(\widetilde{H}) \end{bmatrix}.$$

From this, we readily obtain the equalities in (30) as well as the following equality:

$$\widetilde{G}_{\beta\beta} + \widetilde{H}_{\beta\beta} = \mathcal{G}(\widetilde{G}) + \mathcal{H}(\widetilde{H}).$$

Using Lemma 4.1 and noting that $\widetilde{G}_{\gamma\beta} = 0, \widetilde{H}_{\alpha\beta} = 0$, we get the inequality in (30). ∎

PROPOSITION 4.2. *Let* $\overline{X},\overline{Y} \in \mathbb{S}^n$ *satisfy* (28)*, and assume that they have the spectral decomposition as in* (29)*. Then, for any* $(\mathcal{U},\mathcal{V}) \in \partial\Phi_{\rm FB}(\overline{X},\overline{Y})$*, it holds that*

$$\mathcal{U}(G) + \mathcal{V}(H) = 0 \implies \langle G, H\rangle \le \Upsilon_{\overline{X}}(-\overline{Y}, G),$$

where, for any given $B \in \mathbb{S}^n$*,* $\Upsilon_B : \mathbb{S}^n \times \mathbb{S}^n \to \mathbb{R}$ *is the linear-quadratic function*

$$\Upsilon_B(\Delta, A) \ := \ 2\langle \Delta, AB^\dagger A\rangle \quad \forall (\Delta, A) \in \mathbb{S}^n \times \mathbb{S}^n \tag{31}$$

introduced in [22] *with* $B^\dagger$ *denoting the Moore–Penrose pseudoinverse of* B*.*

Proof. The proof is direct by Proposition 4.1, the definition of $\Upsilon_{\overline{X}}(-\overline{Y}, G)$, and (29). ∎

We also need to recall from [22] the strong second order sufficient condition and constraint nondegeneracy for the NLSDP (2). Let $z \equiv (x, X) \in \mathbb{X} \times \mathbb{S}^n$. Let

$$\widetilde{f}(z) \equiv f(x), \quad \widetilde{h}(z) \equiv \begin{pmatrix} h(x) \\ g(x) - X \end{pmatrix}, \quad \text{and} \quad \widetilde{g}(z) \equiv X.$$

By (29) and [22, eq. (17)], the tangent cone $\mathcal{T}_{\mathbb{S}^n_+}(\overline{X})$ of $\mathbb{S}^n_+$ at $\overline{X}$ takes the form

$$\mathcal{T}_{\mathbb{S}^n_+}(\overline{X}) = \left\{ B \in \mathbb{S}^n : [P_\beta \ \ P_\gamma]^T B [P_\beta \ \ P_\gamma] \succeq 0 \right\}. \tag{32}$$

Let $A \equiv \overline{X} - \overline{Y}$ and $A_+ \equiv \Pi_{\mathbb{S}^n_+}(A)$. The critical cone of $\mathbb{S}^n_+$ at A is defined as

$$\mathcal{C}(A; \mathbb{S}^n_+) := \mathcal{T}_{\mathbb{S}^n_+}(A_+) \cap (A_+ - A)^\perp = \mathcal{T}_{\mathbb{S}^n_+}(\overline{X}) \cap \overline{Y}^\perp. \tag{33}$$

By the spectral decomposition of $\overline{Y}$ and the expression of $\mathcal{T}_{\mathbb{S}^n_+}(\overline{X})$, we may verify that

$$\mathcal{C}(A; \mathbb{S}^n_+) = \left\{ B \in \mathbb{S}^n : P_\beta^T B P_\beta \succeq 0, \ P_\beta^T B P_\gamma = 0, \ P_\gamma^T B P_\gamma = 0 \right\}. \tag{34}$$

From [2, 22], the critical cone $\mathcal{C}(\overline{z})$ of the NLSDP (2) at $\overline{z} = (\overline{x}, \overline{X})$ has the form

$$\mathcal{C}(\overline{z}) = \left\{ \xi \in \mathbb{X} \times \mathbb{S}^n : \mathcal{J}_z \widetilde{h}(\overline{z})\xi = 0, \ \mathcal{J}_z \widetilde{g}(\overline{z})\xi \in \mathcal{C}(A; \mathbb{S}^n_+) \right\}.$$

Since it is hard to give an explicit formula to the affine hull of $\mathcal{C}(\overline{z})$, denoted by $\mathrm{aff}(\mathcal{C}(\overline{z}))$, Sun [22] defined the following outer approximation to $\mathrm{aff}(\mathcal{C}(\overline{z}))$ with respect to $(\overline{\mu}, \overline{S}, \overline{Y})$:

$$\mathrm{app}(\overline{\mu}, \overline{S}, \overline{Y}) := \left\{ \xi \in \mathbb{X} \times \mathbb{S}^n : \mathcal{J}_z \widetilde{h}(\overline{z})\xi = 0, \ \mathcal{J}_z \widetilde{g}(\overline{z})\xi \in \mathrm{aff}(\mathcal{C}(A; \mathbb{S}^n_+)) \right\}. \tag{35}$$

For a locally optimal solution $\overline{z} = (\overline{x}, \overline{X})$ of (2), we denote by $\mathcal{M}(\overline{z})$ the set of Lagrange multipliers satisfying (3) that is nonempty under certain constraint qualifications (CQs) such as Robinson's CQ. By the definition of $\widetilde{f}, \widetilde{h}$, and $\widetilde{g}$, the strong second order sufficient condition and constraint nondegeneracy [22] for the NLSDP (2) can be stated as follows.

DEFINITION 4.1. *Let $\overline{z} = (\overline{x}, \overline{X})$ be a stationary point of the NLSDP* (2). *We say that the strong second order sufficient condition holds at $\overline{z}$ if, for any $\xi \in \widehat{\mathcal{C}}(\overline{z}) \backslash \{0\}$,*

$$\sup_{(\mu, S, Y) \in \mathcal{M}(\overline{z})} \left\{ \langle \xi, \mathcal{J}^2_{zz} L(\overline{x}, \overline{X}, \mu, S, Y)\xi \rangle - \Upsilon_{\widetilde{g}(\overline{z})}(-Y, \mathcal{J}_z \widetilde{g}(\overline{z})\xi) \right\} > 0, \tag{36}$$

where $\widehat{\mathcal{C}}(\overline{z}) := \bigcap_{(\mu, S, Y) \in \mathcal{M}(\overline{z})} \mathrm{app}(\mu, S, Y)$.

DEFINITION 4.2. *We say that a feasible point $\overline{z} = (\overline{x}, \overline{X})$ of the NLSDP* (2) *is constraint nondegenerate if*

$$\begin{pmatrix} \mathcal{J}_z \widetilde{h}(\overline{z}) \\ \mathcal{J}_z \widetilde{g}(\overline{z}) \end{pmatrix} \begin{pmatrix} \mathbb{X} \\ \mathbb{S}^n \end{pmatrix} + \begin{pmatrix} \{0\} \\ \mathrm{lin}\left(\mathcal{T}_{\mathbb{S}^n_+}(\widetilde{g}(\overline{z}))\right) \end{pmatrix} = \begin{pmatrix} \mathbb{R}^m \times \mathbb{S}^n \\ \mathbb{S}^n \end{pmatrix}, \tag{37}$$

where $\mathrm{lin}(\mathcal{T}_{\mathbb{S}^n_+}(\cdot))$ *denotes the largest linear space in the tangent cone* $\mathcal{T}_{\mathbb{S}^n_+}(\cdot)$.

Now Propositions 4.1 and 4.2 allow us to prove the following result by using an argument similar to that of [22, Proposition 3.2]. We include the proof for completeness.

PROPOSITION 4.3. *Let* $(\overline{x}, \overline{X}, \overline{\mu}, \overline{S}, \overline{Y})$ *be an arbitrary KKT point of the NLSDP* (2). *Suppose the strong second order sufficient condition* (36) *holds at* $\overline{z} = (\overline{x}, \overline{X})$ *and* $\overline{z}$ *is constraint nondegenerate. Then any element in* $\partial E_{\rm FB}(\overline{x}, \overline{X}, \overline{\mu}, \overline{S}, \overline{Y})$ *is nonsingular.*

Proof. Since the nondegeneracy condition (37) is assumed to hold at $\overline{z}$ and, by (32)–(34), $\mathrm{lin}(\mathcal{T}_{\mathbb{S}^n_+}(\widetilde{g}(\overline{z})) \subseteq \mathcal{T}_{\mathbb{S}^n_+}(\widetilde{g}(\overline{z})) \cap \overline{Y}^{\perp}$, we have from [22, Proposition 3.1] that

$$\mathcal{M}(\overline{z}) = \{(\overline{\mu}, \overline{S}, \overline{Y})\} \quad \text{and} \quad \mathrm{aff}(\mathcal{C}(\overline{z})) = \mathrm{app}(\overline{\mu}, \overline{S}, \overline{Y}).$$

Then, the strong second order sufficient condition (36) takes the form

$$\left\langle d, \mathcal{J}^2_{xx} l(\overline{x}, \overline{\mu}, \overline{S}) d \right\rangle - \Upsilon_{\overline{X}}(-\overline{Y}, \Lambda) > 0 \quad \forall (d, \Lambda) \in \mathrm{aff}(\mathcal{C}(\overline{z})) \backslash \{(0,0)\}, \tag{38}$$

where $l(x, \mu, S) := f(x) + \langle \mu, h(x) \rangle + \langle g(x), S \rangle$. Let $\mathcal{W}$ be an arbitrary element in $\partial E_{\rm FB}(\overline{x}, \overline{X}, \overline{\mu}, \overline{S}, \overline{Y})$. To prove that W is nonsingular, we let $(\triangle x, \triangle X, \triangle \mu, \triangle S, \triangle Y) \in \mathbb{X} \times \mathbb{S}^n \times \mathbb{R}^m \times \mathbb{S}^n \times \mathbb{S}^n$ be such that $\mathcal{W}(\triangle x, \triangle X, \triangle \mu, \triangle S, \triangle Y) = 0$. Then, by the expression of the mapping $E_{\rm FB}$, there exists a $(\mathcal{U}, \mathcal{V}) \in \partial \Phi_{\rm FB}(\overline{X}, \overline{Y})$ such that

$$\begin{bmatrix} \mathcal{J}_{xx} l(\overline{x}, \overline{\mu}, \overline{S}) \triangle x + \mathcal{J}_x h(\overline{x})^* \triangle \mu + \mathcal{J}_x g(\overline{x})^* \triangle S \\ -\triangle S - \triangle Y \\ \mathcal{J}_x h(\overline{x}) \triangle x \\ \mathcal{J}_x g(\overline{x}) \triangle x - \triangle X \\ \mathcal{U}(\triangle X) + \mathcal{V}(\triangle Y) \end{bmatrix} = 0,$$

which can be simplified as

$$\begin{bmatrix} \mathcal{J}^2_{xx} l(\overline{x}, \overline{\mu}, \overline{S}) \triangle x + \mathcal{J}_x h(\overline{x})^* \triangle \mu - \mathcal{J}_x g(\overline{x})^* \triangle Y \\ \mathcal{J}_x h(\overline{x}) \triangle x \\ \mathcal{U}(\mathcal{J}_x g(\overline{x}) \triangle x) + \mathcal{V}(\triangle Y) \end{bmatrix} = 0. \tag{39}$$

From the second and the third equations of (39) and Proposition 4.1, it follows that

$$\mathcal{J}_x h(\overline{x}) \triangle x = 0, \quad P_\beta^T (\mathcal{J}_x g(\overline{x}) \triangle x) P_\gamma = 0, \quad P_\gamma^T (\mathcal{J}_x g(\overline{x}) \triangle x) P_\gamma = 0.$$

Comparing this with the definition of $\mathrm{app}(\overline{\mu}, \overline{S}, \overline{Y})$ in (35), we have that

$$(\triangle x, \mathcal{J}_x g(\overline{x}) \triangle x) \in \mathrm{app}(\overline{\mu}, \overline{S}, \overline{Y}) = \mathrm{aff}(\mathcal{C}(\overline{z})). \tag{40}$$

By the first and the second equations of (39), we can obtain that

$$\langle \triangle x, \mathcal{J}^2_{xx} l(\overline{x}, \overline{\mu}, \overline{S}) \triangle x \rangle - \langle \mathcal{J}_x g(\overline{x}) \triangle x, \triangle Y \rangle = 0,$$

whereas the third equality of (39) and Proposition 4.2 imply that

$$\langle \mathcal{J}_x g(\overline{x}) \triangle x, \triangle Y \rangle \leq \Upsilon_{\overline{X}}(-\overline{Y}, \mathcal{J}_x g(\overline{x}) \triangle x).$$

From the last two equations, we immediately obtain

$$\langle \triangle x, \mathcal{J}^2_{xx} l(\overline{x}, \overline{\mu}, \overline{S}) \triangle x \rangle - \Upsilon_{\overline{X}}(-\overline{Y}, \mathcal{J}_x g(\overline{x}) \triangle x) \leq 0.$$

Putting this with (40) and (38), we get $\triangle x = 0$. Thus, (39) reduces to

$$\begin{bmatrix} \mathcal{J}_x h(\overline{x})^* \triangle\mu - \mathcal{J}_x g(\overline{x})^* \triangle Y \\ \mathcal{V}(\triangle Y) \end{bmatrix} = 0. \tag{41}$$

Applying Proposition 4.1 with $G = 0$ and $H = \triangle Y$, we obtain that

$$P_\alpha^T \triangle Y P_\alpha = 0, \quad P_\alpha^T \triangle Y P_\beta = 0, \text{ and } P_\alpha^T \triangle Y P_\gamma = 0. \tag{42}$$

In addition, by (37), there exist a $(\zeta, U) \in \mathbb{X} \times \mathbb{S}^n$ and a $V \in \mathrm{lin}(\mathcal{T}_{\mathbb{S}^n_+}(\widetilde{g}(\overline{z}))$ such that

$$\mathcal{J}_x h(\overline{x})\zeta = \triangle\mu, \quad \mathcal{J}_x g(\overline{x})\zeta - U = -\triangle Y, \quad U + V = -\triangle Y.$$

This, together with the first equation of (41), yields that

$$\begin{aligned} \langle \triangle\mu, \triangle\mu \rangle + 2\langle \triangle Y, \triangle Y \rangle &= \langle \mathcal{J}_x h(\overline{x})\zeta, \triangle\mu \rangle - \langle \mathcal{J}_x g(\overline{x})\zeta - U, \triangle Y \rangle - \langle U + V, \triangle Y \rangle \\ &= \langle V, \triangle Y \rangle = \langle P^T V P, P^T \triangle Y P \rangle = 0, \end{aligned}$$

where the last equality uses (42) and $V \in \mathrm{lin}(\mathcal{T}_{\mathbb{S}^n_+}(\widetilde{g}(\overline{z}))$. Thus, $\triangle\mu = 0$ and $\triangle Y = 0$. Together with $\triangle x = 0$, we show that $\mathcal{W}$ is nonsingular. ∎

Next we turn to the work of the second step, which needs the following key lemma.

LEMMA 4.2. *Let $\overline{X}$ and $\overline{Y}$ satisfy* (28). *Then, we have $\partial_B \Phi_{\rm NR}(\overline{X}, \overline{Y}) \subseteq \partial_B \Phi_{\rm FB}(\overline{X}, \overline{Y})$.*

Proof. By the eigenvalue decomposition of $\overline{X}$ and $\overline{Y}$, it is easy to verify that

$$\overline{X} = \frac{1}{2}\left[|\overline{X} - \overline{Y}| + (\overline{X} - \overline{Y})\right] \quad \text{and} \quad \overline{Y} = \frac{1}{2}\left[|\overline{X} - \overline{Y}| - (\overline{X} - \overline{Y})\right]. \tag{43}$$

From the definition of $\Phi_{\rm NR}$, it follows that $\Phi_{\rm NR}(X, Y) = (X + Y) - \Xi(X, Y)$, where

$$\Xi(X, Y) := \frac{1}{2}\left[(X + Y) + |X - Y|\right] \qquad \forall X, Y \in \mathbb{S}^n.$$

Then, comparing this with the definition of $\Phi_{\rm FB}$, it suffices to prove that

$$\partial_B \Xi(\overline{X}, \overline{Y}) \subseteq \partial_B C(\overline{X}, \overline{Y}). \tag{44}$$

Let $(\mathcal{U}, \mathcal{V}) \in \partial_B \Xi(\overline{X}, \overline{Y})$. From the definition of the elements in $\partial_B \Xi(\overline{X}, \overline{Y})$ and [14, Corollary 10], there exists a sequence $\{(X^k, Y^k)\} \subset \mathbb{S}^n \times \mathbb{S}^n$ converging to $(\overline{X}, \overline{Y})$ with $Z^k \equiv X^k - Y^k$ nonsingular such that $(\mathcal{U}, \mathcal{V}) = \lim_{k\to\infty} \mathcal{J}\Xi(X^k, Y^k)$. Also, for any $G, H \in \mathbb{S}^n$,

$$\begin{aligned} \mathcal{U}(G) + \mathcal{V}(H) &= \lim_{k\to\infty} \left[\frac{1}{2}(\mathcal{I} + \mathcal{L}_{|Z^k|}^{-1}\mathcal{L}_{Z^k})(G) + \frac{1}{2}(\mathcal{I} - \mathcal{L}_{|Z^k|}^{-1}\mathcal{L}_{Z^k})(H)\right] \\ &= \lim_{k\to\infty} \left[\mathcal{L}_{|Z^k|}^{-1}\mathcal{L}_{\frac{|Z^k| + Z^k}{2}}(G) + \mathcal{L}_{|Z^k|}^{-1}\mathcal{L}_{\frac{|Z^k| - Z^k}{2}}(H)\right]. \end{aligned} \tag{45}$$

For each k, let $\widehat{X}^k = \frac{|Z^k| + Z^k}{2}$ and $\widehat{Y}^k = \frac{|Z^k| - Z^k}{2}$. Then, by $Z^k \equiv X^k - Y^k$ and (43), it is easy to see that $\widehat{X}^k \to \overline{X}$ and $\widehat{Y}^k \to \overline{Y}$ as $k \to \infty$. Also, we have that

$$(\widehat{X}^k)^2 + (\widehat{Y}^k)^2 = (Z^k)^2 = (X^k - Y^k)^2 \succ 0.$$

This means that the function $C(\cdot,\cdot)$ is continuously differentiable at $(\widehat{X}^k, \widehat{Y}^k)$ with

$$\begin{aligned}\mathcal{J}C(\widehat{X}^k, \widehat{Y}^k)(G,H) &= \mathcal{L}^{-1}_{C(\widehat{X}^k,\widehat{Y}^k)}\mathcal{L}_{\widehat{X}^k}(G) + \mathcal{L}^{-1}_{C(\widehat{X}^k,\widehat{Y}^k)}\mathcal{L}_{\widehat{Y}^k}(H)\\ &= \mathcal{L}^{-1}_{|Z^k|}\mathcal{L}_{\frac{|Z^k|+Z^k}{2}}(G) + \mathcal{L}^{-1}_{|Z^k|}\mathcal{L}_{\frac{|Z^k|-Z^k}{2}}(H).\end{aligned}$$

Together with (45), we have $\mathcal{U}(G) + \mathcal{V}(H) = \lim_{k\to\infty} \mathcal{J}C(\widehat{X}^k, \widehat{Y}^k)(G,H)$. This, by the arbitrariness of G and H, shows that $(\mathcal{U},\mathcal{V}) \in \partial_B C(\overline{X},\overline{Y})$. Then, (44) follows. ◻

Lemma 4.2 states the relation between Clarke's Jacobian of $\Phi_{\rm NR}$ and that of $\Phi_{\rm FB}$ at a complementarity point pair, which by the expression of $E_{\rm FB}$ and $E_{\rm NR}$ implies

$$\partial E_{\rm NR}(\overline{x},\overline{X},\overline{\mu},\overline{S},\overline{Y}) \subseteq \partial E_{\rm FB}(\overline{x},\overline{X},\overline{\mu},\overline{S},\overline{Y}). \tag{46}$$

Along with Proposition 4.3 and [22, Theorem 4.1], we get the main result of this paper.

THEOREM 4.1. *Let $(\overline{x},\overline{X}) \in \mathbb{X}\times\mathbb{S}^n_+$ be a locally optimal solution to the NLSDP* (2). *Suppose that Robinson's CQ holds at this point. Let $(\overline{\mu},\overline{S},\overline{Y}) \in \mathbb{R}^m\times\mathbb{S}^n\times\mathbb{S}^n_+$ be such that $(\overline{x},\overline{X},\overline{\mu},\overline{S},\overline{Y})$ is a KKT point of* (2). *Then the following statements are equivalent:*

(a) *The strong second order sufficient condition in Definition* 4.1 *holds at $(\overline{x},\overline{X})$, and $(\overline{x},\overline{X})$ is constraint nondegenerate.*

(b) *Any element in $\partial E_{\rm FB}(\overline{x},\overline{X},\overline{\mu},\overline{S},\overline{Y})$ is nonsingular.*

(c) *Any element in $\partial E_{\rm NR}(\overline{x},\overline{X},\overline{\mu},\overline{S},\overline{Y})$ is nonsingular.*

(d) *$(\overline{x},\overline{X},\overline{\mu},\overline{S},\overline{Y})$ is a strongly regular solution to the generalized equation* (27).

Proof. By Proposition 4.3 and the inclusion in (46), we have that (a) ⇒ (b) ⇒ (c). Since the NLSDP (2) is obtained from (1) by introducing a slack variable, we know from [22, Theorem 4.1] that (a) ⇔ (c) ⇔ (d). Thus, we complete the proof. ◻

To close this section, we take a look at the relationship between the nonsingularity of Clarke's Jacobian of the FB nonsmooth mapping associated to the KKT system of (1) and the strong regularity of the KKT point. Let $F_{\rm FB}: \mathbb{X}\times\mathbb{R}^m\times\mathbb{S}^n \to \mathbb{X}\times\mathbb{R}^m\times\mathbb{S}^n$ be the FB nonsmooth mapping associated to the KKT system of (1); that is,

$$F_{\rm FB}(x,\mu,Y) := \begin{bmatrix} \mathcal{J}_x l(x,\mu,Y)\\ h(x)\\ \Phi_{\rm FB}(g(x),Y)\end{bmatrix} \quad \forall (x,\mu,Y)\in\mathbb{X}\times\mathbb{R}^m\times\mathbb{S}^n,$$

where $l: \mathbb{X}\times\mathbb{R}^m\times\mathbb{S}^n\to\mathbb{R}$ is the Lagrangian function of (1). It is easy to verify that if $(\overline{x},\overline{\mu},\overline{Y})$ is a KKT point of (1), then $(\overline{x},g(\overline{x}),\overline{\mu},-\overline{Y},\overline{Y})$ is a KKT point of (2); and conversely, if $(\overline{x},\overline{X},\overline{\mu},\overline{S},\overline{Y})$ is a KKT point of (2), then $(\overline{x},\overline{\mu},\overline{Y})$ is a KKT point of (1). Moreover, from the Kummer inverse function theorem [13] and the Thibault directional derivative of composite functions [25], it follows that the following result holds.

LEMMA 4.3. *If the mapping $E_{\rm FB}$ is a locally Lipschitz homeomorphism near a KKT point $(\overline{x},\overline{X},\overline{\mu},\overline{S},\overline{Y})$ of* (2), *then $F_{\rm FB}$ is a locally Lipschitz homeomorphism near $(\overline{x},\overline{\mu},\overline{Y})$.*

In fact, by Lemma 2.3 of [23], it is not hard to prove that the converse conclusion of Lemma 4.3 also holds if $\mathcal{J}_x g(\overline{x})\colon \mathbb{R}^m\to\mathbb{S}^n$ is surjective. Thus, combining Lemma 4.3 and Theorem 4.1 above with [22, Theorem 4.1], we obtain the following result.

THEOREM 4.2. *Let* $\overline{x} \in \mathbb{X}$ *be a locally optimal solution to the NLSDP* (1). *Suppose that Robinson's CQ holds at this point. Let* $(\overline{\mu}, \overline{Y}) \in \mathbb{R}^m \times \mathbb{S}^n_+$ *be such that* $(\overline{x}, \overline{\mu}, \overline{Y})$ *is a KKT point of* (1). *If* $(\overline{x}, \overline{\mu}, \overline{Y})$ *is a strongly regular solution to the generalized equation*

$$0 \in \begin{bmatrix} \mathcal{J}_x l(x, \mu, Y) \\ h(x) \\ g(x) \end{bmatrix} + \begin{bmatrix} \mathcal{N}_{\mathbb{X}}(x) \\ \mathcal{N}_{\mathbb{R}^m}(\mu) \\ \mathcal{N}_{\mathbb{S}^n_+}(Y) \end{bmatrix}, \tag{47}$$

then any element of $\partial F_{\rm FB}(\overline{x}, \overline{\mu}, \overline{Y})$ *is nonsingular. Conversely, if* $\mathcal{J}_x g(\overline{x})$ *is surjective and any element of* $\partial F_{\rm FB}(\overline{x}, \overline{\mu}, \overline{Y})$ *is nonsingular, then* $(\overline{x}, \overline{\mu}, \overline{Y})$ *is a strongly regular solution to the generalized equation* (47).

5. Conclusions. In this paper, for a locally optimal solution to the nonlinear SDP (2), we established the equivalence between the nonsingularity of Clarke's Jacobian of the FB system and the strong regularity of the KKT point. This provides a new characterization for the strong regularity of the nonlinear SDPs as well as extends the result of [9, Corollary 3.7] for the FB system of variational inequalities with the polyhedral cone constraints to the setting of SDCs. In addition, this result also implies that the semismooth Newton method [15, 16] applied to the FB system converges quadratically to a KKT point if the strong second order sufficient condition and constraint nondegeneracy are satisfied.

Appendix.

The proof of Lemma 3.3. When $(X, Y) = (0, 0)$, the result is clear since $\Psi_{\rm FB}(\cdot, \cdot) \equiv \Phi_{\rm FB}(\cdot, \cdot)$. Therefore, in the following arguments, we assume that $(X, Y) \neq (0, 0)$.

Step 1: *To prove that* $\partial_B \Phi_{\rm FB}(X, Y) \subseteq \partial_B \Psi_{\rm FB}(0, 0)$. Let $(\mathcal{U}, \mathcal{V}) \in \partial_B \Phi_{\rm FB}(X, Y)$. By Corollary 3.1 and the definition of the elements in $\partial_B \Phi_{\rm FB}(X, Y)$, there exists a sequence $\{(X^k, Y^k)\}$ in $\mathbb{S}^n \times \mathbb{S}^n$ converging to (X, Y) with $C^k \equiv C(X^k, Y^k) \succ 0$ such that $(\mathcal{U}, \mathcal{V}) = \lim_{k\to\infty} \mathcal{J}\Phi_{\rm FB}(X^k, Y^k)$. Fix any $G, H \in \mathbb{S}^n$. From formula (20), it follows that

$$(\mathcal{I} - \mathcal{U})(G) + (\mathcal{I} - \mathcal{V})(H) = \lim_{k\to\infty} \mathcal{L}^{-1}_{C^k} \left[\mathcal{L}_{X^k}(G) + \mathcal{L}_{Y^k}(H)\right]. \tag{48}$$

Let $C^k = P^k D^k (P^k)^T$ be the spectral decomposition of C^k, where D^k is the diagonal matrix of eigenvalues of C^k and P^k is a corresponding matrix of orthonormal eigenvectors. Writing each D^k in the same form as D, i.e., $D^k = \left[\begin{smallmatrix} D^k_\kappa & 0 \\ 0 & D^k_\beta \end{smallmatrix}\right]$, we have $\lim_{k\to\infty} D^k = D$, which implies that D^k_κ is a nonsingular matrix for sufficiently large k and $\lim_{k\to\infty} D^k_\beta = 0$. Without loss of generality, taking subsequences if necessary, we assume that $\{P^k\}$ is a convergent sequence with $\lim_{k\to\infty} P^k = P^\infty$, which means that

$$C(X, Y) = \lim_{k\to\infty} C^k = \lim_{k\to\infty} P^k D^k (P^k)^T = P^\infty D (P^\infty)^T.$$

Hence, P^∞ can be identified with P in (13). In what follows, we use P instead of P^∞. Let

$$Z^k \equiv \mathcal{L}^{-1}_{C^k} \left[\mathcal{L}_{X^k}(G) + \mathcal{L}_{Y^k}(H)\right]. \tag{49}$$

With $\widetilde{Z}^k \equiv (P^k)^T Z^k P^k$, $\widetilde{X}^k \equiv (P^k)^T X^k P^k$, $\widetilde{Y}^k \equiv (P^k)^T Y^k P^k$, $\widetilde{G}^k \equiv (P^k)^T G P^k$, and $\widetilde{H}^k \equiv (P^k)^T H P^k$, we can rewrite equality (49) in the block form

$$\begin{bmatrix} D_\kappa^k \widetilde{Z}_{\kappa\kappa}^k + \widetilde{Z}_{\kappa\kappa}^k D_\kappa^k & D_\kappa^k \widetilde{Z}_{\kappa\beta}^k + \widetilde{Z}_{\kappa\beta}^k D_\beta^k \\ D_\beta^k \widetilde{Z}_{\beta\kappa}^k + \widetilde{Z}_{\beta\kappa}^k D_\kappa^k & D_\beta^k \widetilde{Z}_{\beta\beta}^k + \widetilde{Z}_{\beta\beta}^k D_\beta^k \end{bmatrix} = \begin{bmatrix} \Xi_{\kappa\kappa}^k & \Xi_{\kappa\beta}^k \\ (\Xi_{\kappa\beta}^k)^T & \Xi_{\beta\beta}^k \end{bmatrix}, \tag{50}$$

where

$$\begin{aligned} \Xi_{\kappa\kappa}^k &= \mathcal{L}_{\widetilde{X}_{\kappa\kappa}^k}(\widetilde{G}_{\kappa\kappa}^k) + \widetilde{X}_{\kappa\beta}^k \widetilde{G}_{\beta\kappa}^k + \widetilde{G}_{\kappa\beta}^k \widetilde{X}_{\beta\kappa}^k + \mathcal{L}_{\widetilde{Y}_{\kappa\kappa}^k}(\widetilde{H}_{\kappa\kappa}^k) + \widetilde{Y}_{\kappa\beta}^k \widetilde{H}_{\beta\kappa}^k + \widetilde{H}_{\kappa\beta}^k \widetilde{Y}_{\beta\kappa}^k, \\ \Xi_{\kappa\beta}^k &= \widetilde{X}_{\kappa\kappa}^k \widetilde{G}_{\kappa\beta}^k + \widetilde{G}_{\kappa\kappa}^k \widetilde{X}_{\kappa\beta}^k + \widetilde{X}_{\kappa\beta}^k \widetilde{G}_{\beta\beta}^k + \widetilde{G}_{\kappa\beta}^k \widetilde{X}_{\beta\beta}^k \\ &\quad + \widetilde{Y}_{\kappa\kappa}^k \widetilde{H}_{\kappa\beta}^k + \widetilde{H}_{\kappa\kappa}^k \widetilde{Y}_{\kappa\beta}^k + \widetilde{Y}_{\kappa\beta}^k \widetilde{H}_{\beta\beta}^k + \widetilde{H}_{\kappa\beta}^k \widetilde{Y}_{\beta\beta}^k, \\ \Xi_{\beta\beta}^k &= \widetilde{X}_{\beta\kappa}^k \widetilde{G}_{\kappa\beta}^k + \widetilde{G}_{\beta\kappa}^k \widetilde{X}_{\kappa\beta}^k + \mathcal{L}_{\widetilde{X}_{\beta\beta}^k}(\widetilde{G}_{\beta\beta}^k) + \widetilde{Y}_{\beta\kappa}^k \widetilde{H}_{\kappa\beta}^k + \widetilde{H}_{\beta\kappa}^k \widetilde{Y}_{\kappa\beta}^k + \mathcal{L}_{\widetilde{Y}_{\beta\beta}^k}(\widetilde{H}_{\beta\beta}^k). \end{aligned}$$

Since $X^k \to X$ and $Y^k \to Y$ as $k \to \infty$, and $\widetilde{X}^2 + \widetilde{Y}^2 = D^2$, it is not hard to see that

$$\begin{aligned} &\lim_{k\to\infty} \widetilde{X}_{\kappa\kappa}^k = \widetilde{X}_{\kappa\kappa}, \quad \lim_{k\to\infty} \widetilde{X}_{\kappa\beta}^k = 0, \quad \lim_{k\to\infty} \widetilde{X}_{\beta\beta}^k = 0, \\ &\lim_{k\to\infty} \widetilde{Y}_{\kappa\kappa}^k = \widetilde{Y}_{\kappa\kappa}, \quad \lim_{k\to\infty} \widetilde{Y}_{\kappa\beta}^k = 0, \quad \lim_{k\to\infty} \widetilde{Y}_{\beta\beta}^k = 0. \end{aligned} \tag{51}$$

Using these equalities and (50), we immediately obtain that

$$\lim_{k\to\infty} \widetilde{Z}_{\kappa\kappa}^k = \mathcal{L}_{D_\kappa}^{-1}\left[\mathcal{L}_{\widetilde{X}_{\kappa\kappa}}(\widetilde{G}_{\kappa\kappa}) + \mathcal{L}_{\widetilde{Y}_{\kappa\kappa}}(\widetilde{H}_{\kappa\kappa})\right], \tag{52}$$

$$\lim_{k\to\infty} \widetilde{Z}_{\kappa\beta}^k = D_\kappa^{-1}\left(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta}\right), \tag{53}$$

with $\widetilde{G} = P^T G P$ and $\widetilde{H} = P^T H P$. Since $(\widetilde{X}^k)^2 + (\widetilde{Y}^k)^2 = (D^k)^2$, we have that

$$\widetilde{X}_{\kappa\kappa}^k \widetilde{X}_{\kappa\beta}^k + \widetilde{X}_{\kappa\beta}^k \widetilde{X}_{\beta\beta}^k + \widetilde{Y}_{\kappa\kappa}^k \widetilde{Y}_{\kappa\beta}^k + \widetilde{Y}_{\kappa\beta}^k \widetilde{Y}_{\beta\beta}^k = 0, \tag{54}$$

$$\widetilde{X}_{\beta\kappa}^k \widetilde{X}_{\kappa\beta}^k + (\widetilde{X}_{\beta\beta}^k)^2 + \widetilde{Y}_{\beta\kappa}^k \widetilde{Y}_{\kappa\beta}^k + (\widetilde{Y}_{\beta\beta}^k)^2 = (D_\beta^k)^2. \tag{55}$$

By (55) and Propositions 2.2 and 2.3, the sequences $\{\mathcal{L}_{D_\beta^k}^{-1}\mathcal{L}_{\widetilde{X}_{\beta\beta}^k}\}$, $\{\mathcal{L}_{D_\beta^k}^{-1}\mathcal{L}_{\widetilde{Y}_{\beta\beta}^k}\}$, $\{\mathcal{L}_{D_\beta^k}^{-1}(\widetilde{X}_{\beta\kappa}^k \widetilde{G}_{\kappa\beta}^k + \widetilde{G}_{\beta\kappa}^k \widetilde{X}_{\kappa\beta}^k)\}$, and $\{\mathcal{L}_{D_\beta^k}^{-1}(\widetilde{Y}_{\beta\kappa}^k \widetilde{H}_{\kappa\beta}^k + \widetilde{H}_{\beta\kappa}^k \widetilde{Y}_{\kappa\beta}^k)\}$ are bounded. So is $\{\widetilde{Z}_{\beta\beta}^k\}$ with

$$\widetilde{Z}_{\beta\beta}^k = \mathcal{L}_{D_\beta^k}^{-1}\left[\widetilde{X}_{\beta\kappa}^k \widetilde{G}_{\kappa\beta}^k + \widetilde{G}_{\beta\kappa}^k \widetilde{X}_{\kappa\beta}^k + \widetilde{Y}_{\beta\kappa}^k \widetilde{H}_{\kappa\beta}^k + \widetilde{H}_{\beta\kappa}^k \widetilde{Y}_{\kappa\beta}^k + \mathcal{L}_{\widetilde{X}_{\beta\beta}^k}(\widetilde{G}_{\beta\beta}^k) + \mathcal{L}_{\widetilde{Y}_{\beta\beta}^k}(\widetilde{H}_{\beta\beta}^k)\right]. \tag{56}$$

By taking a subsequence if necessary, we may assume that $\{\widetilde{Z}_{\beta\beta}^k\}$ is convergent. Then, together with (48)–(50) and (52)–(53), we obtain that

$$\begin{aligned} &(\mathcal{I} - \mathcal{U})(G) + (\mathcal{I} - \mathcal{V})(H) \\ &= P\begin{bmatrix} \mathcal{L}_{D_\kappa}^{-1}[\mathcal{L}_{\widetilde{X}_{\kappa\kappa}}(\widetilde{G}_{\kappa\kappa}) + \mathcal{L}_{\widetilde{Y}_{\kappa\kappa}}(\widetilde{H}_{\kappa\kappa})] & D_\kappa^{-1}(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta}) \\ (\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D_\kappa^{-1} & \lim_{k\to\infty} \widetilde{Z}_{\beta\beta}^k \end{bmatrix} P^T. \end{aligned} \tag{57}$$

Now with $\widetilde{X}^k$ and $\widetilde{Y}^k$ we define a sequence $\{(U^k, V^k)\}$ in $\mathbb{S}^n \times \mathbb{S}^n$ by

$$U^k := P\begin{bmatrix} 0 & \widetilde{X}_{\kappa\beta}^k \\ \widetilde{X}_{\beta\kappa}^k & \widetilde{X}_{\beta\beta}^k \end{bmatrix} P^T \quad \text{and} \quad V^k := P\begin{bmatrix} 0 & \widetilde{Y}_{\kappa\beta}^k \\ \widetilde{Y}_{\beta\kappa}^k & \widetilde{Y}_{\beta\beta}^k \end{bmatrix} P^T.$$

Clearly, $(U^k, V^k) \to (0,0)$ as $k \to \infty$. Let $\widetilde{U}^k \equiv P^T U^k P$ and $\widetilde{V}^k \equiv P^T V^k P$. We next argue that Θ is well defined at $(\widetilde{U}^k, \widetilde{V}^k)$ for large enough k. Indeed, from (54), we have

$$\widetilde{X}_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta} = \left(\widehat{X}^k_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta} + \widehat{Y}^k_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta}\right) - \left(\widetilde{X}^k_{\kappa\beta}\widetilde{X}^k_{\beta\beta} + \widetilde{Y}^k_{\kappa\beta}\widetilde{Y}^k_{\beta\beta}\right), \tag{58}$$

where $\widehat{X}^k_{\kappa\kappa} = \widetilde{X}_{\kappa\kappa} - \widetilde{X}^k_{\kappa\kappa}$ and $\widehat{Y}^k_{\kappa\kappa} = \widetilde{Y}_{\kappa\kappa} - \widetilde{Y}^k_{\kappa\kappa}$. Then, applying Lemma 2.1 yields that

$$\begin{aligned}
&(\widetilde{X}_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta})^T D_\kappa^{-2}(\widetilde{X}_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta})\\
&\quad = \left[(\widehat{X}^k_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta} + \widehat{Y}^k_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta}) - (\widetilde{X}^k_{\kappa\beta}\widetilde{X}^k_{\beta\beta} + \widetilde{Y}^k_{\kappa\beta}\widetilde{Y}^k_{\beta\beta})\right]^T\\
&D_\kappa^{-2}\left[(\widehat{X}^k_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta} + \widehat{Y}^k_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta}) - (\widetilde{X}^k_{\kappa\beta}\widetilde{X}^k_{\beta\beta} + \widetilde{Y}^k_{\kappa\beta}\widetilde{Y}^k_{\beta\beta})\right]\\
&\quad \preceq 4(\widehat{X}^k_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta})^T D_\kappa^{-2}(\widehat{X}^k_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta}) + 4(\widehat{Y}^k_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta})^T D_\kappa^{-2}(\widehat{Y}^k_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta})\\
&\qquad + 4(\widetilde{X}^k_{\kappa\beta}\widetilde{X}^k_{\beta\beta})^T D_\kappa^{-2}(\widetilde{X}^k_{\kappa\beta}\widetilde{X}^k_{\beta\beta}) + 4(\widetilde{Y}^k_{\kappa\beta}\widetilde{Y}^k_{\beta\beta})^T D_\kappa^{-2}(\widetilde{Y}^k_{\kappa\beta}\widetilde{Y}^k_{\beta\beta}).
\end{aligned} \tag{59}$$

This implies that for sufficiently large k,

$$\begin{aligned}
&(\widetilde{Y}^k_{\beta\beta})^2 + (\widetilde{X}^k_{\beta\beta})^2 + \widetilde{X}^k_{\beta\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{Y}^k_{\beta\kappa}\widetilde{Y}^k_{\kappa\beta} - (\widetilde{X}_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta})^T D_\kappa^{-2}(\widetilde{X}_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta})\\
&\quad \succeq \widetilde{X}^k_{\beta\kappa}\left(I - 4\widehat{X}^k_{\kappa\kappa}D_\kappa^{-2}\widehat{X}^k_{\kappa\kappa}\right)\widetilde{X}^k_{\kappa\beta} + \widetilde{Y}^k_{\beta\kappa}\left(I - 4\widehat{Y}^k_{\kappa\kappa}D_\kappa^{-2}\widehat{Y}^k_{\kappa\kappa}\right)\widetilde{Y}^k_{\kappa\beta} + \frac{1}{2}(D^k_\beta)^2\\
&\quad + \frac{1}{2}\widetilde{X}^k_{\beta\beta}\left(I - 8\widetilde{X}^k_{\beta\kappa}D_\kappa^{-2}\widetilde{X}^k_{\kappa\beta}\right)\widetilde{X}^k_{\beta\beta} + \frac{1}{2}\widetilde{Y}^k_{\beta\beta}\left(I - 8\widetilde{Y}^k_{\beta\kappa}D_\kappa^{-2}\widetilde{Y}^k_{\kappa\beta}\right)\widetilde{Y}^k_{\beta\beta} \succeq \frac{1}{2}(D^k_\beta)^2 \succ 0,
\end{aligned}$$

where the second inequality uses $\widetilde{X}^k_{\beta\kappa}, \widetilde{Y}^k_{\beta\kappa}, \widehat{X}^k_{\kappa\kappa}, \widehat{Y}^k_{\kappa\kappa} \to 0$ as $k \to \infty$. By the definition of Θ, this shows that for k large enough, Θ is well defined at $(\widetilde{U}^k, \widetilde{V}^k)$ and $\Theta(\widetilde{U}^k, \widetilde{V}^k) \succ 0$. By Lemma 3.2, Θ is F-differentiable at $(\widetilde{U}^k, \widetilde{V}^k)$ with $\mathcal{J}\Theta(\widetilde{U}^k, \widetilde{V}^k)(\widetilde{G}, \widetilde{H})$ equal to

$$\begin{aligned}
&\mathcal{L}^{-1}_{\Theta(\widetilde{U}^k, \widetilde{V}^k)}\Big[\widetilde{X}^k_{\beta\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{G}_{\beta\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{Y}^k_{\beta\kappa}\widetilde{H}_{\kappa\beta} + \widetilde{H}_{\beta\kappa}\widetilde{Y}^k_{\kappa\beta} + \mathcal{L}_{\widetilde{X}^k_{\beta\beta}}(\widetilde{G}_{\beta\beta}) + \mathcal{L}_{\widetilde{Y}^k_{\beta\beta}}(\widetilde{H}_{\beta\beta})\\
&\qquad - \left(\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa}\right)D_\kappa^{-2}\left(\widetilde{X}_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta}\right)\\
&\qquad - \left(\widetilde{X}^k_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{Y}^k_{\beta\kappa}\widetilde{Y}_{\kappa\kappa}\right)D_\kappa^{-2}\left(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta}\right)\Big].
\end{aligned} \tag{60}$$

Using (55) and (51), and $\widehat{X}^k_{\kappa\kappa}, \widehat{Y}^k_{\kappa\kappa} \to 0$ as $k \to \infty$, we have from (59) that

$$\Theta(\widetilde{U}^k, \widetilde{V}^k) = \left[(D^k_\beta)^2 + o(1)(D^k_\beta)^2\right]^{1/2}. \tag{61}$$

In addition, from (58) and (51), it is not hard to obtain that

$$\begin{aligned}
&\left(\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa}\right)D_\kappa^{-2}\left(\widetilde{X}_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta}\right)\\
&\quad + \left(\widetilde{X}^k_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{Y}^k_{\beta\kappa}\widetilde{Y}_{\kappa\kappa}\right)D_\kappa^{-2}\left(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta}\right)\\
&= R^k_{\beta\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{X}^k_{\beta\kappa}R^k_{\kappa\beta} + S^k_{\beta\kappa}\widetilde{Y}^k_{\kappa\beta} + \widetilde{Y}^k_{\beta\kappa}S^k_{\kappa\beta} + \mathcal{L}_{\widetilde{X}^k_{\beta\beta}}(R^k_{\beta\beta}) + \mathcal{L}_{\widetilde{Y}^k_{\beta\beta}}(S^k_{\beta\beta})
\end{aligned}$$

with $R^k_{\beta\kappa}, R^k_{\beta\beta}, S^k_{\beta\kappa}, S^k_{\beta\beta} \to 0$ as $k \to \infty$. Then, by (55) and Propositions 2.2 and 2.3, we have

$$\begin{aligned}
\lim_{k\to\infty}\mathcal{L}^{-1}_{D^k_\beta}\Big[&\left(\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa}\right)D_\kappa^{-2}\left(\widetilde{X}_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta}\right)\\
&+ \left(\widetilde{X}^k_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{Y}^k_{\beta\kappa}\widetilde{Y}_{\kappa\kappa}\right)D_\kappa^{-2}\left(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta}\right)\Big] = 0.
\end{aligned}$$

Together with (61) and (60), it is not hard to obtain that

$$\lim_{k\to\infty}\mathcal{J}\Theta(\widetilde{U}^k,\widetilde{V}^k)(\widetilde{G},\widetilde{H})=\lim_{k\to\infty}\mathcal{L}^{-1}_{D^k_\beta}\Big[\widetilde{X}^k_{\beta\kappa}\widetilde{G}_{\kappa\beta}+\widetilde{G}_{\beta\kappa}\widetilde{X}^k_{\kappa\beta}+\widetilde{Y}^k_{\beta\kappa}\widetilde{H}_{\kappa\beta} + \widetilde{H}_{\beta\kappa}\widetilde{Y}^k_{\kappa\beta}+\mathcal{L}_{\widetilde{X}^k_{\beta\beta}}(\widetilde{G}_{\beta\beta})+\mathcal{L}_{\widetilde{Y}^k_{\beta\beta}}(\widetilde{H}_{\beta\beta})\Big].$$

Comparing it with (56), we have that $\lim_{k\to\infty}\mathcal{J}\Theta(\widetilde{U}^k,\widetilde{V}^k)(\widetilde{G},\widetilde{H})=\lim_{k\to\infty}\widetilde{Z}^k_{\beta\beta}$. Also, by Remark 3.2, the above arguments show that $\Psi_{\rm FB}$ is F-differentiable at (U^k,V^k) with

$$\begin{aligned}&(G+H)-\lim_{k\to\infty}\mathcal{J}\Psi_{\rm FB}(U^k,V^k)(G,H)\\ &=P\begin{bmatrix}\mathcal{L}^{-1}_{D_\kappa}\big(\mathcal{L}_{\widetilde{X}_{\kappa\kappa}}(\widetilde{G}_{\kappa\kappa})+\mathcal{L}_{\widetilde{Y}_{\kappa\kappa}}(\widetilde{H}_{\kappa\kappa})\big) & D^{-1}_\kappa(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta}+\widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta})\\ (\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa}+\widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D^{-1}_\kappa & \lim_{k\to\infty}\widetilde{Z}^k_{\beta\beta}.\end{bmatrix}P^T.\end{aligned}$$

Comparing it with (57) yields that $\mathcal{U}(G)+\mathcal{V}(H)=\lim_{k\to\infty}\mathcal{J}\Psi_{\rm FB}(U^k,V^k)(G,H)$. Since (G,H) is arbitrary in $\mathbb{S}^n\times\mathbb{S}^n$, this shows that $(\mathcal{U},\mathcal{V})\in\partial_B\Psi_{\rm FB}(0,0)$. The result follows.

Step 2: *To prove that* $\partial_B\Psi_{\rm FB}(0,0)\subseteq\partial_B\Phi_{\rm FB}(X,Y)$. Let $(\mathcal{U},\mathcal{V})\in\partial_B\Psi_{\rm FB}(0,0)$. By the definition of the elements in $\partial_B\Psi_{\rm FB}(0,0)$ and Remark 3.2, there exists a sequence of matrices $\{(M^k,N^k)\}$ in $\mathbb{S}^n\times\mathbb{S}^n$ converging to $(0,0)$ with $\Theta(\widetilde{M}^k,\widetilde{N}^k)\succ 0$ such that $(\mathcal{U},\mathcal{V})=\lim_{k\to\infty}\mathcal{J}\Psi_{\rm FB}(M^k,N^k)$, where $\widetilde{M}^k=P^TM^kP$ and $\widetilde{N}^k=P^TN^kP$. Fix any $G,H\in\mathbb{S}^n$ with $\widetilde{G}:=P^TGP$ and $\widetilde{H}:=P^THP$. From the definition of $\Psi_{\rm FB}$, we have

$$\begin{aligned}&(\mathcal{I}-\mathcal{U})(G)+(\mathcal{I}-\mathcal{V})(H)\\ &=P\begin{bmatrix}\mathcal{L}^{-1}_{D_\kappa}[\mathcal{L}_{\widetilde{X}_{\kappa\kappa}}(\widetilde{G}_{\kappa\kappa})+\mathcal{L}_{\widetilde{Y}_{\kappa\kappa}}(\widetilde{H}_{\kappa\kappa})] & D^{-1}_\kappa(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta}+\widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta})\\ (\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa}+\widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D^{-1}_\kappa & \lim_{k\to\infty}\mathcal{J}\Theta(\widetilde{M}^k,\widetilde{N}^k)(\widetilde{G},\widetilde{H})\end{bmatrix}P^T,\end{aligned}\tag{62}$$

where, by Lemma 3.2, $\mathcal{J}\Theta(\widetilde{M}^k,\widetilde{N}^k)(\widetilde{G},\widetilde{H})$ has the following expression:

$$\begin{aligned}\mathcal{L}^{-1}_{\Theta(\widetilde{M}^k,\widetilde{N}^k)}\Big[&\widetilde{M}^k_{\beta\kappa}\widetilde{G}_{\kappa\beta}+\widetilde{G}_{\beta\kappa}\widetilde{M}^k_{\kappa\beta}+\widetilde{N}^k_{\beta\kappa}\widetilde{H}_{\kappa\beta}+\widetilde{H}_{\beta\kappa}\widetilde{N}^k_{\kappa\beta}+\mathcal{L}_{\widetilde{M}^k_{\beta\beta}}(\widetilde{G}_{\beta\beta})\\ &+\mathcal{L}_{\widetilde{N}^k_{\beta\beta}}(\widetilde{H}_{\beta\beta})-\Big(\widetilde{M}^k_{\beta\kappa}\widetilde{X}_{\kappa\kappa}+\widetilde{N}^k_{\beta\kappa}\widetilde{Y}_{\kappa\kappa}\Big)D^{-2}_\kappa\Big(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta}+\widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta}\Big)\\ &-\Big(\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa}+\widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa}\Big)D^{-2}_\kappa\Big(\widetilde{X}_{\kappa\kappa}\widetilde{M}^k_{\beta\kappa}+\widetilde{Y}_{\kappa\kappa}\widetilde{N}^k_{\beta\kappa}\Big)\Big].\end{aligned}$$

With $\widetilde{M}^k$ and $\widetilde{N}^k$ we define a sequence $\{(X^k,Y^k)\}$ in $\mathbb{S}^n\times\mathbb{S}^n$ by

$$X^k:=P\begin{bmatrix}\widetilde{X}_{\kappa\kappa} & \widetilde{S}^k_{\kappa\beta}\\ (\widetilde{S}^k_{\kappa\beta})^T & \widetilde{M}^k_{\beta\beta}\end{bmatrix}P^T\quad\text{and}\quad Y^k:=P\begin{bmatrix}\widetilde{Y}_{\kappa\kappa} & \widetilde{T}^k_{\kappa\beta}\\ (\widetilde{T}^k_{\kappa\beta})^T & \widetilde{N}^k_{\beta\beta}\end{bmatrix}P^T,\tag{63}$$

where

$$\begin{aligned}\widetilde{S}^k_{\kappa\beta}&:=\widetilde{M}^k_{\kappa\beta}-\widetilde{X}_{\kappa\kappa}D^{-2}_\kappa(\widetilde{X}_{\kappa\kappa}\widetilde{M}^k_{\kappa\beta}+\widetilde{Y}_{\kappa\kappa}\widetilde{N}^k_{\kappa\beta}),\\ \widetilde{T}^k_{\kappa\beta}&:=\widetilde{N}^k_{\kappa\beta}-\widetilde{Y}_{\kappa\kappa}D^{-2}_\kappa(\widetilde{X}_{\kappa\kappa}\widetilde{M}^k_{\kappa\beta}+\widetilde{Y}_{\kappa\kappa}\widetilde{N}^k_{\kappa\beta}).\end{aligned}$$

Clearly, $X^k\to X$ and $Y^k\to Y$ as $k\to\infty$. Let $C^k\equiv C(X^k,Y^k)$ and $\widehat{C}^k\equiv P^T(C^k)^2P$. We next show that $\widehat{C}^k\succ 0$ for large enough k. A simple computation yields that

$$\widehat{C}^k_{\kappa\kappa}=D^2_\kappa+\widetilde{S}^k_{\kappa\beta}\widetilde{S}^k_{\beta\kappa}+\widetilde{T}^k_{\kappa\beta}\widetilde{T}^k_{\beta\kappa},\ \widehat{C}^k_{\kappa\beta}=\widetilde{S}^k_{\kappa\beta}\widetilde{M}^k_{\beta\beta}+\widetilde{T}^k_{\kappa\beta}\widetilde{N}^k_{\beta\beta},\ \widehat{C}^k_{\beta\beta}=\Theta^2(\widetilde{M}^k,\widetilde{N}^k).$$

Since $\widehat{C}^k_{\kappa\kappa} \succ 0$ and $\widehat{C}^k_{\beta\beta} \succ 0$ for each k, by [10, Theorem 7.7.6] we need only argue that $\Gamma^k_{\beta\beta} \equiv \widehat{C}^k_{\beta\beta} - \widehat{C}^k_{\beta\kappa}(\widehat{C}^k_{\kappa\kappa})^{-1}\widehat{C}^k_{\kappa\beta} \succ 0$ for large enough k. By computation, $\Gamma^k_{\beta\beta}$ equals

$$\begin{aligned}
&\Theta^2(\widetilde{M}^k, \widetilde{N}^k) - (\widetilde{S}^k_{\kappa\beta}\widetilde{M}^k_{\beta\beta} + \widetilde{T}^k_{\kappa\beta}\widetilde{N}^k_{\beta\beta})^T \left(D^2_\kappa + \widetilde{S}^k_{\kappa\beta}\widetilde{S}^k_{\beta\kappa} + \widetilde{T}^k_{\kappa\beta}\widetilde{T}^k_{\beta\kappa}\right)^{-1} (\widetilde{S}^k_{\kappa\beta}\widetilde{M}^k_{\beta\beta} + \widetilde{T}^k_{\kappa\beta}\widetilde{N}^k_{\beta\beta})\\
&\quad \succeq \Theta^2(\widetilde{M}^k, \widetilde{N}^k) - (\widetilde{S}^k_{\kappa\beta}\widetilde{M}^k_{\beta\beta} + \widetilde{T}^k_{\kappa\beta}\widetilde{N}^k_{\beta\beta})^T D_\kappa^{-2} (\widetilde{S}^k_{\kappa\beta}\widetilde{M}^k_{\beta\beta} + \widetilde{T}^k_{\kappa\beta}\widetilde{N}^k_{\beta\beta})\\
&\quad \succeq \Theta^2(\widetilde{M}^k, \widetilde{N}^k) - 2\widetilde{M}^k_{\beta\beta}\widetilde{S}^k_{\beta\kappa}D_\kappa^{-2}\widetilde{S}^k_{\kappa\beta}\widetilde{M}^k_{\beta\beta} - 2\widetilde{N}^k_{\beta\beta}\widetilde{T}^k_{\beta\kappa}D_\kappa^{-2}\widetilde{T}^k_{\kappa\beta}\widetilde{N}^k_{\beta\beta}\\
&\quad = \frac{1}{2}[(\widetilde{M}^k_{\beta\beta})^2 + (\widetilde{N}^k_{\beta\beta})^2] + \widetilde{M}^k_{\beta\kappa}\widetilde{M}^k_{\kappa\beta} + \widetilde{N}^k_{\beta\kappa}\widetilde{N}^k_{\kappa\beta} + \frac{1}{2}\widetilde{M}^k_{\beta\beta}(I - 4\widetilde{S}^k_{\beta\kappa}D_\kappa^{-2}\widetilde{S}^k_{\kappa\beta})\widetilde{M}^k_{\beta\beta}\\
&\qquad + \frac{1}{2}\widetilde{N}^k_{\beta\beta}(I - 4\widetilde{T}^k_{\beta\kappa}D_\kappa^{-2}\widetilde{T}^k_{\kappa\beta})\widetilde{N}^k_{\beta\beta} - (\widetilde{M}^k_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{N}^k_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D_\kappa^{-2}(\widetilde{X}_{\kappa\kappa}\widetilde{M}^k_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{N}^k_{\kappa\beta})\\
&\quad \succeq \frac{1}{2}[(\widetilde{M}^k_{\beta\beta})^2 + (\widetilde{N}^k_{\beta\beta})^2] + \widetilde{M}^k_{\beta\kappa}\widetilde{M}^k_{\kappa\beta} + \widetilde{N}^k_{\beta\kappa}\widetilde{N}^k_{\kappa\beta}\\
&\qquad - (\widetilde{M}^k_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{N}^k_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D_\kappa^{-2}(\widetilde{X}_{\kappa\kappa}\widetilde{M}^k_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{N}^k_{\kappa\beta}),
\end{aligned}$$

where the second inequality is due to Lemma 2.1, and the third inequality uses $\widetilde{S}^k_{\beta\kappa} \to 0$ and $\widetilde{T}^k_{\beta\kappa} \to 0$. Now applying Lemma 2.2 with $A = \widetilde{M}^k_{\kappa\beta}$ and $B = \widetilde{N}^k_{\kappa\beta}$, we know that the second term on the right-hand side of last equation is positive semidefinite, which implies that for sufficiently large k, $\Gamma^k_{\beta\beta} \succeq \frac{1}{2}\Theta^2(\widetilde{M}^k, \widetilde{N}^k) \succ 0$. Thus, we show that $\widehat{C}^k \succ 0$, and consequently $C^k \succ 0$ for sufficiently large k. By Corollary 3.1, $\Phi_{\rm FB}$ is F-differentiable at (X^k, Y^k) for sufficiently large k, and for any $G, H \in \mathbb{S}^n$,

$$(G + H) - \lim_{k\to\infty} \mathcal{J}\Phi_{\rm FB}(X^k, Y^k)(G, H) = \lim_{k\to\infty} \mathcal{L}^{-1}_{C^k}\left(\mathcal{L}_{X^k}(G) + \mathcal{L}_{Y^k}(H)\right). \tag{64}$$

Let

$$Z^k \equiv \mathcal{L}^{-1}_{C^k}\left(\mathcal{L}_{X^k}(G) + \mathcal{L}_{Y^k}(H)\right).$$

Then, with $\widetilde{X}^k \equiv P^T X^k P, \widetilde{Y}^k \equiv P^T Y^k P, \widetilde{C}^k \equiv P^T C^k P$, and $\widetilde{Z}^k \equiv P^T Z^k P$, we have

$$\widetilde{C}^k\widetilde{Z}^k + \widetilde{Z}^k\widetilde{C}^k = \widetilde{X}^k\widetilde{G} + \widetilde{G}\widetilde{X}^k + \widetilde{Y}^k\widetilde{H} + \widetilde{H}\widetilde{Y}^k. \tag{65}$$

Note that

$$\widetilde{C}^k = (\widehat{C}^k)^{1/2} = \left(\begin{bmatrix} \widehat{C}^k_{\kappa\kappa} & 0 \\ 0 & \Theta^2(\widetilde{M}^k, \widetilde{N}^k) \end{bmatrix} + \widetilde{W}^k\right)^{1/2},$$

where

$$\widetilde{W}^k = \begin{bmatrix} 0 & \widetilde{S}^k_{\kappa\beta}\widetilde{M}^k_{\beta\beta} + \widetilde{T}^k_{\kappa\beta}\widetilde{N}^k_{\beta\beta} \\ \widetilde{M}^k_{\beta\beta}\widetilde{S}^k_{\beta\kappa} + \widetilde{N}^k_{\beta\beta}\widetilde{T}^k_{\beta\kappa} & 0 \end{bmatrix}.$$

Applying [24, Lemma 6.2] and noting that $\widehat{C}^k_{\kappa\kappa} = D^2_\kappa + \widetilde{S}^k_{\kappa\beta}\widetilde{S}^k_{\beta\kappa} + \widetilde{T}^k_{\kappa\beta}\widetilde{T}^k_{\beta\kappa}$, we have

$$\widetilde{C}^k = \begin{bmatrix} (\widehat{C}^k_{\kappa\kappa})^{1/2} + o(\|\widetilde{W}^k\|) & D_\kappa^{-1}\widetilde{W}^k_{\kappa\beta} + o(\|\widetilde{W}^k\|) \\ \widetilde{W}^k_{\beta\kappa}D_\kappa^{-1} + o(\|\widetilde{W}^k\|) & \Theta(\widetilde{M}^k, \widetilde{N}^k) + o(\|\widetilde{W}^k\|) \end{bmatrix}.$$

By this expression of $\widetilde{C}^k$ and (65), we may calculate that

$$
\begin{aligned}
\mathcal{L}_{(\widehat{C}^k_{\kappa\kappa})^{1/2}}(\widetilde{Z}^k_{\kappa\kappa}) &= \Big(\widetilde{X}^k_{\kappa\kappa}\widetilde{G}_{\kappa\kappa} + \widetilde{G}_{\kappa\kappa}\widetilde{X}^k_{\kappa\kappa} + \widetilde{X}^k_{\kappa\beta}\widetilde{G}_{\beta\kappa} + \widetilde{G}_{\kappa\beta}\widetilde{X}^k_{\beta\kappa}\Big)\\
&\quad + \Big(\widetilde{Y}^k_{\kappa\kappa}\widetilde{H}_{\kappa\kappa} + \widetilde{H}_{\kappa\kappa}\widetilde{Y}^k_{\kappa\kappa} + \widetilde{Y}^k_{\kappa\beta}\widetilde{H}_{\beta\kappa} + \widetilde{H}_{\kappa\beta}\widetilde{Y}^k_{\beta\kappa}\Big) + \widetilde{R}^k_{\kappa\kappa},\\
(\widehat{C}^k_{\kappa\kappa})^{1/2}\widetilde{Z}^k_{\kappa\beta} &= \Big(\widetilde{X}^k_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{G}_{\kappa\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{X}^k_{\kappa\beta}\widetilde{G}_{\beta\beta} + \widetilde{G}_{\kappa\beta}\widetilde{X}^k_{\beta\beta}\Big)\\
&\quad + \Big(\widetilde{Y}^k_{\kappa\kappa}\widetilde{H}_{\kappa\beta} + \widetilde{H}_{\kappa\kappa}\widetilde{Y}^k_{\kappa\beta} + \widetilde{Y}^k_{\kappa\beta}\widetilde{H}_{\beta\beta} + \widetilde{H}_{\kappa\beta}\widetilde{Y}^k_{\beta\beta}\Big) + \widetilde{R}^k_{\kappa\beta},\\
\mathcal{L}_{\Theta(\widetilde{M}^k,\widetilde{N}^k)}(\widetilde{Z}^k_{\beta\beta}) &= \Big(\widetilde{X}^k_{\beta\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{G}_{\beta\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{X}^k_{\beta\beta}\widetilde{G}_{\beta\beta} + \widetilde{G}_{\beta\beta}\widetilde{X}^k_{\beta\beta}\Big)\\
&\quad + \Big(\widetilde{Y}^k_{\beta\kappa}\widetilde{H}_{\kappa\beta} + \widetilde{H}_{\beta\kappa}\widetilde{Y}^k_{\kappa\beta} + \widetilde{Y}^k_{\beta\beta}\widetilde{H}_{\beta\beta} + \widetilde{H}_{\beta\beta}\widetilde{Y}^k_{\beta\beta}\Big)\\
&\quad - \widetilde{W}^k_{\beta\kappa}D_\kappa^{-1}\widetilde{Z}^k_{\kappa\beta} - \widetilde{Z}^k_{\beta\kappa}D_\kappa^{-1}\widetilde{W}^k_{\kappa\beta} + o(\|\widetilde{W}^k\|),
\end{aligned} \tag{66}
$$

where $\widetilde{R}^k_{\kappa\kappa} \to 0$ and $\widetilde{R}^k_{\kappa\beta} \to 0$ as $k \to \infty$. From the first two equalities in (66), we have

$$
\begin{aligned}
\lim_{k\to\infty} \widetilde{Z}^k_{\kappa\kappa} &= \mathcal{L}^{-1}_{D_\kappa}(\mathcal{L}_{\widetilde{X}_{\kappa\kappa}}(\widetilde{G}_{\kappa\kappa}) + \mathcal{L}_{\widetilde{Y}_{\kappa\kappa}}(\widetilde{H}_{\kappa\kappa})),\\
\lim_{k\to\infty} \widetilde{Z}^k_{\kappa\beta} &= D_\kappa^{-1}(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta}).
\end{aligned} \tag{67}
$$

In addition, applying Lemma 2.2 with $A = \widetilde{M}^k_{\kappa\beta}$ and $B = \widetilde{N}^k_{\kappa\beta}$, we know that

$$\widetilde{M}^k_{\beta\kappa}\widetilde{M}^k_{\kappa\beta} + \widetilde{N}^k_{\beta\kappa}\widetilde{N}^k_{\kappa\beta} - (\widetilde{M}^k_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{N}^k_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D_\kappa^{-2}(\widetilde{X}_{\kappa\kappa}\widetilde{M}^k_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{N}^k_{\kappa\beta}) \succeq 0,$$

which, by the definition of Θ, means that $\Theta(\widetilde{M}^k, \widetilde{N}^k) \succeq [(\widetilde{M}^k_{\beta\beta})^2 + (\widetilde{N}^k_{\beta\beta})^2]^{1/2}$. Note that

$$
\begin{aligned}
\|\mathcal{L}^{-1}_{\Theta(\widetilde{M}^k,\widetilde{N}^k)}(\widetilde{M}^k_{\beta\beta})\| &= \frac{1}{2}\|\mathcal{L}^{-1}_{\Theta(\widetilde{M}^k,\widetilde{N}^k)}\mathcal{L}_{\widetilde{M}^k_{\beta\beta}}(I_{|\beta|})\|\\
&\le \frac{1}{2}\|\mathcal{L}^{-1}_{\Theta(\widetilde{M}^k,\widetilde{N}^k)}\mathcal{L}_{\widetilde{M}^k_{\beta\beta}}\|_2\|I_{|\beta|}\| \le \frac{1}{2}\sqrt{|\beta|},
\end{aligned}
$$

where $I_{|\beta|}$ is a $|\beta|\times|\beta|$ unit matrix, and the last inequality is by Proposition 2.2. Hence, $\{\mathcal{L}^{-1}_{\Theta(\widetilde{M}^k,\widetilde{N}^k)}(\widetilde{M}^k_{\beta\beta})\}$ is bounded. Similarly, $\{\mathcal{L}^{-1}_{\Theta(\widetilde{M}^k,\widetilde{N}^k)}(\widetilde{N}^k_{\beta\beta})\}$ is bounded. Thus,

$$\lim_{k\to\infty}\mathcal{L}^{-1}_{\Theta(\widetilde{M}^k,\widetilde{N}^k)}(\widetilde{W}^k_{\beta\kappa}D_\kappa^{-1}\widetilde{Z}^k_{\kappa\beta} + \widetilde{Z}^k_{\beta\kappa}D_\kappa^{-1}\widetilde{W}^k_{\kappa\beta}) = 0 \text{ and } \lim_{k\to\infty}\mathcal{L}^{-1}_{\Theta(\widetilde{M}^k,\widetilde{N}^k)}(o(\|\widetilde{W}^k\|)) = 0.$$

From the third equality of (66), the definition of $\widetilde{X}^k, \widetilde{Y}^k$, and (63), it then follows that

$$
\begin{aligned}
\lim_{k\to\infty}\widetilde{Z}^k_{\beta\beta} &= \lim_{k\to\infty}\mathcal{L}^{-1}_{\Theta(\widetilde{M}^k,\widetilde{N}^k)}\Big(\widetilde{X}^k_{\beta\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{G}_{\beta\kappa}\widetilde{X}^k_{\kappa\beta} + \widetilde{X}^k_{\beta\beta}\widetilde{G}_{\beta\beta} + \widetilde{G}_{\beta\beta}\widetilde{X}^k_{\beta\beta}\\
&\qquad + \widetilde{Y}^k_{\beta\kappa}\widetilde{H}_{\kappa\beta} + \widetilde{H}_{\beta\kappa}\widetilde{Y}^k_{\kappa\beta} + \widetilde{Y}^k_{\beta\beta}\widetilde{H}_{\beta\beta} + \widetilde{H}_{\beta\beta}\widetilde{Y}^k_{\beta\beta}\Big)\\
&= \lim_{k\to\infty}\mathcal{J}\Theta(\widetilde{M}^k,\widetilde{N}^k)(\widetilde{G},\widetilde{H}).
\end{aligned}
$$

Combining this equality with (64)–(65) and (67), we obtain that

$$
\begin{aligned}
&(G+H) - \lim_{k\to\infty}\mathcal{J}\Phi_{\rm FB}(X^k,Y^k)(G,H)\\
&= P\begin{bmatrix} \mathcal{L}^{-1}_{D_\kappa}(\mathcal{L}_{\widetilde{X}_{\kappa\kappa}}(\widetilde{G}_{\kappa\kappa}) + \mathcal{L}_{\widetilde{Y}_{\kappa\kappa}}(\widetilde{H}_{\kappa\kappa})) & D_\kappa^{-1}(\widetilde{X}_{\kappa\kappa}\widetilde{G}_{\kappa\beta} + \widetilde{Y}_{\kappa\kappa}\widetilde{H}_{\kappa\beta})\\ (\widetilde{G}_{\beta\kappa}\widetilde{X}_{\kappa\kappa} + \widetilde{H}_{\beta\kappa}\widetilde{Y}_{\kappa\kappa})D_\kappa^{-1} & \lim_{k\to\infty}\mathcal{J}\Theta(\widetilde{M}^k,\widetilde{N}^k)(\widetilde{G},\widetilde{H}) \end{bmatrix} P^T.
\end{aligned}
$$

Comparing it with (62) yields that $\mathcal{U}(G) + \mathcal{V}(H) = \lim_{k\to\infty} \mathcal{J}\Phi_{\rm FB}(X^k, Y^k)(G, H)$. Since (G, H) is arbitrary in $\mathbb{S}^n \times \mathbb{S}^n$, this shows that $(\mathcal{U}, \mathcal{V}) \in \partial_B \Phi_{\rm FB}(X, Y)$. The result follows.

Combining Step 2 with Step 1, we complete the proof of Lemma 3.3. ∎

Acknowledgments. We would like to thank Professor Defeng Sun at National University of Singapore for pointing out an error in the original proof of Lemma 3.3 and giving us many helpful suggestions for the revision and reorganization of the manuscript. Without his help, we would not have the current version of the manuscript. Thanks also go to Professor Liwei Zhang at Dalian University of Technology for sending his manuscript to us, and Dr. Lu Nan at TianJin University for pointing out an error in the original proof of Lemma 3.3.

REFERENCES

[1] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.
[2] J.-F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.
[3] Z.-X. Chan and D. Sun, *Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming*, SIAM J. Optim., 19 (2008), pp. 370–396.
[4] X. Chen and P. Tseng, *Non-interior continuation methods for solving semidefinite complementarity problems*, Math. Program., 95 (2003), pp. 431–474.
[5] X. Chen, H. Qi, and P. Tseng, *Analysis of nonsmooth symmetric-matrix-valued functions with applications to semidefinite complementarity problems*, SIAM J. Optim., 13 (2003), pp. 960–985.
[6] F. H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983 (reprinted as Classics Appl. Math. 5, SIAM, Philadelphia, 1990).
[7] A. Fischer, *Solution of monotone complementarity problems with locally Lipschitzian functions*, Math. Program., 76 (1997), pp. 513–532.
[8] M. Fukushima, Z.-Q. Luo, and P. Tseng, *Smoothing functions for second-order-cone complementarity problems*, SIAM J. Optim., 12 (2002), pp. 436–460.
[9] F. Facchinei, A. Fischer, and C. Kanzow, *Regularity properties of a semismooth reformulation of variational inequalities*, SIAM J. Optim., 8 (1998), pp. 850–869.
[10] R. A. Horn and C. R. Johnson, Matrix Analysis. Cambridge University Press, Cambridge, UK, 1986.
[11] C. Kanzow and C. Nagel, *Semidefinite programs: New search directions, smoothing-type methods, and numerical results*, SIAM J. Optim., 13 (2002), pp. 1–23.
[12] C. Kanzow and C. Nagel, *Corrigendum: Semidefinite programs: New search directions, smoothing-type methods, and numerical results*, SIAM J. Optim., 14 (2004), pp. 936–937.
[13] B. Kummer, *An implicit-function theorem for $C^{0,1}$-equations and parametric $C^{1,1}$-optimization*, J. Math. Anal. Appl., 158 (1991), pp. 35–46.
[14] J.-S. Pang, D. F. Sun, and J. Sun, *Semismooth homemorphisms and strong stability of semidefinite and Lorentz complementarity problems*, Math. Oper. Res., 28 (2003), pp. 39–63.
[15] J.-S. Pang and L. Qi, *Nonsmooth equations: Motivation and algorithms*, SIAM J. Optim., 3 (1993), pp. 443–465.
[16] L. Qi and J. Sun, *A nonsmooth version of Newton's method*, Math. Program., 58 (1993), pp. 353–367.
[17] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
[18] S. M. Robinson, *Strong regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43–62.
[19] J. Sun, D. Sun, and L. Qi, *A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems*, SIAM J. Optim., 14 (2004), pp. 783–806.
[20] D. F. Sun and J. Sun, *Semismooth matrix-valued functions*, Math. Oper. Res., 27 (2002), pp. 150–169.
[21] D. F. Sun and J. Sun, *Strong semismoothness of the Fischer-Burmeister SDC and SOC complementarity functions*, Math. Program., 103 (2005), pp. 575–581.
[22] D. F. Sun, *The strong second order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications*, Math. Oper. Res., 31 (2006), pp. 761–776.

[23] D. F. Sun, L. W. Zhang, and Y. Wang, *On the Locally Lipschitz Homeomorphism of Nonsmooth Mappings for Variational Inequalities*, manuscript, 2011.

[24] P. Tseng, *Merit functions for semidefinite complementarity problems*, Math. Program., 83 (1998), pp. 159–185.

[25] L. Thibault, *On generalized differentials and subdifferentials of Lipschitz vector-valued functions*, Nonlinear Anal., 6 (1982), pp. 1037–1053.

[26] L. W. Zhang, N. Zhang, and L. P. Pang, *Differential properties of the symmetric matrix-valued Fischer-Burmeister function*, J. Optim. Theory Appl., submitted.