Math. Program., Ser. B
DOI 10.1007/s10107-017-1162-3

FULL LENGTH PAPER

Spectral operators of matrices

Chao Ding[1] · Defeng Sun[2] · Jie Sun[3] · Kim-Chuan Toh[4]

Received: 28 November 2015 / Accepted: 10 May 2017

Abstract The class of matrix optimization problems (MOPs) has been recognized in recent years to be a powerful tool to model many important applications involving structured low rank matrices within and beyond the optimization community. This trend can be credited to some extent to the exciting developments in emerging fields such as compressed sensing. The Löwner operator, which generates a matrix valued function via applying a single-variable function to each of the singular values of a matrix, has played an important role for a long time in solving matrix optimization problems. However, the classical theory developed for the Löwner operator has become inadequate in these recent applications. The main objective of this paper is to provide necessary theoretical foundations from the perspectives of designing efficient numerical methods for solving MOPs. We achieve this goal by introducing and conducting a thorough study on a new class of matrix valued functions, coined as spectral oper-

✉ Defeng Sun
matsundf@nus.edu.sg

Chao Ding
dingchao@amss.ac.cn

Jie Sun
Jie.Sun@curtin.edu.au

Kim-Chuan Toh
mattohkc@nus.edu.sg

[1] Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, People's Republic of China

[2] Department of Mathematics and Risk Management Institute, National University of Singapore, Singapore, Singapore

[3] Department of Mathematics and Statistics, Curtin University, Bentley, Australia

[4] Department of Mathematics, National University of Singapore, Singapore, Singapore

Published online: 15 May 2017

ators of matrices. Several fundamental properties of spectral operators, including the well-definedness, continuity, directional differentiability and Fréchet-differentiability are systematically studied.

Keywords Spectral operators · Directional differentiability · Fréchet differentiability · Matrix valued functions · Proximal mappings

Mathematics Subject Classification 90C25 · 90C06 · 65K05 · 49J50 · 49J52

1 Introduction

In this paper, we introduce a class of matrix valued functions, to be called *spectral operators of matrices*. This class of matrix valued functions frequently arises in various applications such as matrix optimization problems (MOPs). MOPs have recently been found to have many important applications involving matrix norm approximation, matrix completion, rank minimization, graph theory, machine learning, and etc. [2–4,6,9,14,16,17,20,26,36,42–44]. A simple class of MOPs takes the form of

$$\begin{array}{ll} \min & f_0(X) + f(X) \\ \text{s.t.} & \mathcal{A}X = b, \quad X \in \mathcal{X}, \end{array} \tag{1}$$

where $\mathcal{X}$ is the real Euclidean vector space of real/complex matrices over the scalar field of real numbers $\mathbb{R}$, $f_0 : \mathcal{X} \to \mathbb{R}$ is continuously differentiable with a Lipschitzian gradient, $f : \mathcal{X} \to (-\infty, \infty]$ is a closed proper convex function, $\mathcal{A} : \mathcal{X} \to \mathbb{R}^p$ is a linear operator, and $b \in \mathbb{R}^p$. By taking $\mathcal{X} = \mathbb{S}^m$, the real vector subspace of $m \times m$ real symmetric or complex Hermitian matrices, $f_0(X) = \langle C, X\rangle := \mathrm{Re}(\mathrm{trace}(C^{\mathbb{T}}X))$, and $f = \delta_{\mathbb{S}^m_+}$, the convex indicator function of the positive semidefinite matrix cone $\mathbb{S}^m_+$, one recovers semidefinite programming (SDP) [41]. Here $C^{\mathbb{T}}$ is either the transpose or the conjugate transpose depending on whether C is a real or complex matrix. By [37, Corollary 28.3.1] and [31], the Karush-Kuhn-Tucker (KKT) conditions of (1) are equivalent to the following Lipschitzian system of equations

$$\begin{bmatrix} \nabla f_0(X) - \mathcal{A}^* y + \Gamma \\ \mathcal{A}X - b \\ X - P_f(X + \Gamma) \end{bmatrix} = 0\,,$$

where $P_f : \mathcal{X} \to \mathcal{X}$ is the proximal mapping of f at X from convex analysis [37], i.e.,

$$P_f(X) := \mathrm{argmin}_{Y \in \mathcal{X}} \left\{ f(Y) + \frac{1}{2}\|Y - X\|^2 \right\}, \quad X \in \mathcal{X}\,. \tag{2}$$

The optimal value function (denoted by ψ_f) for the minimization problem in (2) is called the Moreau-Yosida regularization of f. It is continuously differentiable with the Lipschitzian gradient $\nabla\psi_f(X) = X - P_f(X)$. The proximal mappings form one of the most important classes of spectral operators of matrices, and the differential properties of P_f play a crucial role in the algorithmic designs of MOPs, see e.g., [7,24,48].

Proximal mappings of unitarily invariant proper closed convex functions belong to a class of matrix functions studied previously in two seminal papers by Lewis [19], and Lewis and Sendov [21]. In [19], Lewis defined a Hermitian matrix valued function by using the gradient mapping $g(\cdot) = \nabla\phi(\cdot) : \mathbb{R}^m \to \mathbb{R}^m$ of a symmetric function $\phi : \mathbb{R}^m \to (-\infty, \infty]$. The corresponding Hermitian matrix valued function $G : \mathbb{S}^m \to \mathbb{S}^m$ is defined by $G(Y) = \sum_{i=1}^m g_i(\lambda(Y)) p_i p_i^{\mathbb{T}}$, where $\{p_1, \ldots, p_m\}$ forms an orthonormal basis of $\mathbb{R}^m$ (or $\mathbb{C}^m$) and $\lambda : \mathbb{S}^m \to \mathbb{R}^m$ is the mapping of the ordered eigenvalues of a Hermitian matrix satisfying $\lambda_1(Y) \geq \lambda_2(Y) \geq \ldots \geq \lambda_m(Y)$ for $Y \in \mathbb{S}^m$. Properties of G such as conditions assuring its (continuous) differentiability are well studied in [19,21]. The (strong) semismoothness [28,35] of G is studied in [34]. Note that if the function g has the form $g(y) = (h(y_1), \ldots, h(y_m))\ \forall\ y \in \mathbb{R}^m$ for a given function $h : \mathbb{R} \to \mathbb{R}$, then the corresponding Hermitian matrix valued function G is called Löwner's (Hermitian) operator [25] (see e.g., [8,40] for more details).

In the potentially non-Hermitian case, i.e., $\mathcal{X} = \mathbb{V}^{m\times n}$, where $\mathbb{V}^{m\times n}$ is either $\mathbb{R}^{m\times n}$ or $\mathbb{C}^{m\times n}$ with $m \leq n$, the mapping g above is assumed to be the gradient mapping of an absolutely symmetric function ϕ, that is, $\phi(x) = \phi(Qx)$ for any $x \in \mathbb{R}^m$ and any *signed permutation matrix* Q, i.e., an $m \times m$ matrix each of whose rows and columns has one nonzero entry which is ± 1. In [18], Lewis studied the corresponding matrix valued function $G(Y) = \sum_{i=1}^m g_i(\sigma(Y)) u_i v_i^{\mathbb{T}}$ for $Y \in \mathbb{V}^{m\times n}$, where $\{u_1, \ldots, u_m\}$ and $\{v_1, \ldots, v_m\}$ are two orthonormal bases of $\mathbb{R}^m$ (or $\mathbb{C}^m$) and σ is the mapping of the ordered singular values of matrices (see also [22] for more details). The related properties of Löwner's (non-Hermitian) operators are studied by Yang [47]. The spectral operators of matrices considered here go well beyond proximal mappings, and so the theoretical results of this paper are not covered by the previously mentioned works [19,21,34,47]. More general spectral operators have been used and played a pivotal role in the study of the low-rank matrix completion problems with fixed basis coefficients [27], where a non-traditional spectral operator G was introduced as the rank-correction function. It is shown in [27, (24)–(26)] that this spectral operator does not arise from either a proximal mapping or gradient mapping of an absolutely symmetric function.

Our main contributions here consist of defining a new class of matrix valued functions involving both Hermitian/symmetric and non-Hermitian/non-symmetric complex/real matrices, which we call spectral operators of matrices and providing the first extensive study of their first- and second-order properties, including the well-definedness, continuity, directional differentiability, and Fréchet-differentiability. We believe that these results are fundamental for both the computational and theoretical study of the general MOPs, based on the recent exciting progress made in solving the SDP problems [5,13,30,38–40,46,48], in which the Löwner operator has played an essential role in the algorithmic design. Therefore, it is expected that the theoretical results for spectral operators established here will shed new light on both designing efficient numerical methods for solving large scale MOPs and conducting second-order variational analysis of the general MOPs.

The remaining parts of this paper are organized as follows. In Sect. 2, we give the definition of spectral operators of matrices and study their well-definedness. We study the continuity, directional and Fréchet-differentiability of spectral operators defined on the single matrix space $\mathbb{V}^{m\times n}$ in Sect. 3. In Sect. 4, we extend the corresponding results to spectral operators defined on the Cartesian product of several matrix spaces. We make some final remarks in Sect. 5.

Below are some common notations and symbols to be used:

- For any $X \in \mathbb{V}^{m\times n}$, we denote by X_{ij} the (i,j)-th entry of X and x_j the j-th column of X. Let $I \subseteq \{1,\ldots,m\}$ and $J \subseteq \{1,\ldots,n\}$ be two index sets. We use X_J to denote the sub-matrix of X obtained by removing all the columns of X not in J and X_{IJ} to denote the $|I|\times|J|$ sub-matrix of X obtained by removing all the rows of X not in I and all the columns of X not in J.
- For $X \in \mathbb{V}^{m\times m}$, $\mathrm{diag}(X)$ denotes the column vector consisting of all the diagonal entries of X being arranged from the first to the last. For $x \in \mathbb{R}^m$, $\mathrm{Diag}(x)$ denotes the $m\times m$ diagonal matrix whose i-th diagonal entry is x_i, $i=1,\ldots,m$.
- We use "$\circ$" to denote the usual Hadamard product between two matrices, i.e., for any two matrices A and B in $\mathbb{V}^{m\times n}$ the (i,j)-th entry of $Z := A\circ B \in \mathbb{V}^{m\times n}$ is $Z_{ij} = A_{ij}B_{ij}$.
- For any given vector $y \in \mathbb{R}^m$, let $|y|^{\downarrow}$ be the vector of entries of $|y| = (|y|_1,\ldots,|y|_m)$ being arranged in the non-increasing order $|y|^{\downarrow}_1 \geq \ldots \geq |y|^{\downarrow}_m$.
- Let $\mathbb{O}^p$ $(p=m,n)$ be the set of $p\times p$ orthogonal/unitary matrices. Denote $\mathbb{P}^p$ and $\pm\mathbb{P}^p$ the sets of all $p\times p$ permutation matrices and signed permutation matrices, respectively. For any $Y \in \mathbb{S}^m$ and $Z \in \mathbb{V}^{m\times n}$, we use $\mathbb{O}^m(Y)$ to denote the set of all orthogonal matrices whose columns form an orthonormal basis of eigenvectors of Y, and use $\mathbb{O}^{m,n}(Z)$ to denote the set of all pairs of orthogonal matrices (U,V), where the columns of U and V form a compatible set of orthonormal left and right singular vectors for Z, respectively.

2 Spectral operators of matrices

In this section, we will first define the spectral operators on the Cartesian product of several real or complex matrix spaces. The study of spectral operators under this general setting is not only useful but also necessary. In fact, spectral operators defined on the Cartesian product of several matrix spaces appear naturally in the study of the differentiability of spectral operators, even if they are only defined on a single matrix space (see the discussion below). Moreover, the spectral operators used in many applications are defined on the Cartesian product of several matrix spaces. See, e.g., [12,45] for more details.

Let s be a positive integer and $0 \leq s_0 \leq s$ be a nonnegative integer. For given positive integers $m_1,\ldots,m_s$ and $n_{s_0+1},\ldots,n_s$, define the real vector space $\mathcal{X}$ by $\mathcal{X} := \mathbb{S}^{m_1}\times\ldots\times\mathbb{S}^{m_{s_0}}\times\mathbb{V}^{m_{s_0+1}\times n_{s_0+1}}\times\ldots\times\mathbb{V}^{m_s\times n_s}$. Without loss of generality, we assume that $m_k \leq n_k$, $k=s_0+1,\ldots,s$. For any $X=(X_1,\ldots,X_s)\in\mathcal{X}$, we have for $1\leq k\leq s_0$, $X_k\in\mathbb{S}^{m_k}$ and $s_0+1\leq k\leq s$, $X_k\in\mathbb{V}^{m_k\times n_k}$.

Denote $\mathcal{Y} := \mathbb{R}^{m_1}\times\ldots\times\mathbb{R}^{m_{s_0}}\times\mathbb{R}^{m_{s_0}}\times\ldots\times\mathbb{R}^{m_s}$. For any $X\in\mathcal{X}$, define $\kappa(X)\in\mathcal{Y}$ by $\kappa(X) := \big(\lambda(X_1),\ldots,\lambda(X_{s_0}),\sigma(X_{s_0+1}),\ldots,\sigma(X_s)\big)$. Define the set $\mathcal{P}$ by $\mathcal{P} := \{(Q_1,\ldots,Q_s) \mid Q_k\in\mathbb{P}^{m_k},\ 1\leq k\leq s_0 \text{ and } Q_k\in\pm\mathbb{P}^{m_k},\ s_0+1\leq k\leq s\}$. Let $g:\mathcal{Y}\to\mathcal{Y}$ be a given mapping. For any $x=(x_1,\ldots,x_s)\in\mathcal{Y}$ with $x_k\in\mathbb{R}^{m_k}$, we write $g(x) \in \mathcal{Y}$ in the form $g(x) = (g_1(x), \ldots, g_s(x))$ with $g_k(x) \in \mathbb{R}^{m_k}$ for $1 \le k \le s$.

Definition 1 [1] The given mapping $g : \mathcal{Y} \to \mathcal{Y}$ is said to be *mixed symmetric*, with respect to $\mathcal{P}$, at $x = (x_1, \ldots, x_s) \in \mathcal{Y}$ with $x_k \in \mathbb{R}^{m_k}$, if

$$g(Q_1 x_1, \ldots, Q_s x_s) = (Q_1 g_1(x), \ldots, Q_s g_s(x)) \quad \forall\ (Q_1, \ldots, Q_s) \in \mathcal{P}. \tag{3}$$

The mapping g is said to be mixed symmetric, with respect to $\mathcal{P}$, over a set $\mathcal{D} \subseteq \mathcal{Y}$ if (3) holds for every $x \in \mathcal{D}$. We call g a *mixed symmetric* mapping, with respect to $\mathcal{P}$, if (3) holds for every $x \in \mathcal{Y}$.

Note that for each $k \in \{1, \ldots, s\}$, the function value $g_k(x) \in \mathbb{R}^{m_k}$ is dependent on all $x_1, \ldots, x_s$. When there is no danger of confusion, in later discussions we often drop "with respect to $\mathcal{P}$" from Definition 1. The following result on g can be checked directly from the definition.

Proposition 1 *Suppose that the mapping $g : \mathcal{Y} \to \mathcal{Y}$ is mixed symmetric at $x = (x_1, \ldots, x_s) \in \mathcal{Y}$ with $x_k \in \mathbb{R}^{m_k}$. Then, for all $1 \le k \le s$ and any $i, j \in \{1, \ldots, m_k\}$, $(g_k(x))_i = (g_k(x))_j$ if $(x_k)_i = (x_k)_j$ and for all $s_0 + 1 \le k \le s$ and any $i \in \{1, \ldots, m_k\}$, $(g_k(x))_i = 0$ if $(x_k)_i = 0$.*

Let $\mathcal{N}$ be a given nonempty set in $\mathcal{X}$. Define $\kappa_{\mathcal{N}} := \{\kappa(X) \in \mathcal{Y} \mid X \in \mathcal{N}\}$.

Definition 2 Suppose that $g : \mathcal{Y} \to \mathcal{Y}$ is mixed symmetric on $\kappa_{\mathcal{N}}$. The spectral operator $G : \mathcal{N} \to \mathcal{X}$ with respect to g is defined by $G(X) := (G_1(X), \ldots, G_s(X)),\quad X = (X_1, \ldots, X_s) \in \mathcal{N}$ with

$$G_k(X) := \begin{cases} P_k \mathrm{Diag}\big(g_k(\kappa(X))\big) P_k^{\mathbb{T}} & \text{if } 1 \le k \le s_0, \\ U_k \left[\mathrm{Diag}\big(g_k(\kappa(X))\big) \ \ 0\right] V_k^{\mathbb{T}} & \text{if } s_0 + 1 \le k \le s, \end{cases}$$

where $P_k \in \mathbb{O}^{m_k}(X_k)$, $1 \le k \le s_0$, $(U_k, V_k) \in \mathbb{O}^{m_k, n_k}(X_k)$, $s_0 + 1 \le k \le s$.

Before showing that spectral operators are well-defined, it is worth mentioning that for the case that $\mathcal{X} \equiv \mathbb{S}^m$ (or $\mathbb{V}^{m \times n}$) if g has the form $g(y) = (h(y_1), \ldots, h(y_m)) \in \mathbb{R}^m$ with $y_i \in \mathbb{R}$ for some given scalar valued functional $h : \mathbb{R} \to \mathbb{R}$, then the corresponding spectral operator G is called the Löwner operator [40] in recognition of Löwner's original contribution on this topic in [25] (or the Löwner non-Hermitian operator [47] if $h(0) = 0$).

Let $\overline{Y} \in \mathbb{S}^m$ be given. Let $\overline{\mu}_1 > \overline{\mu}_2 > \ldots > \overline{\mu}_r$ denote the distinct eigenvalues of $\overline{Y}$. Define the index sets $\alpha_l := \{i \mid \lambda_i(\overline{Y}) = \overline{\mu}_l,\ 1 \le i \le m\}$, $l = 1, \ldots, r$. Let $\Lambda(\overline{Y})$

[1] Note that Definition 1 is different from the property ($\mathcal{E}$) used in [29, Definition 2.2] for the special Hermitian/symmetric case, i.e., $\mathcal{X} = \mathbb{S}^{m_1}$. The conditions used in [29, Definition 2.1 & 2.2] do not seem to be proper ones for studying spectral operators. For instance, consider the function $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f(x) = x^{\downarrow}$ for $x \in \mathbb{R}^2$, where $x^{\downarrow}$ is the vector of entries of x being arranged in the non-increasing order, i.e., $x_1^{\downarrow} \ge x_2^{\downarrow}$. Clearly, f satisfies [29, Definition 2.1 & 2.2] and f is not differentiable at x with $x_1 = x_2$. However, the corresponding matrix function $F(X) = X$ is differentiable on $\mathcal{S}^2$, which implies that [29, Corollary 4.2] is incorrect.

be the $m \times m$ diagonal matrix whose i-th diagonal entry is $\lambda_i(\overline{Y})$. Then, the following elementary property on the eigenvalue decomposition of $\overline{Y}$ can be checked directly.

Proposition 2 *The matrix $Q \in \mathbb{O}^m$ satisfies $Q\Lambda(\overline{Y}) = \Lambda(\overline{Y})Q$ if and only if there exist $Q_l \in \mathbb{O}^{|\alpha_l|}$, $l = 1, \ldots, r$ such that Q is a block diagonal matrix whose l-th diagonal block is Q_l, i.e., $Q = \mathrm{Diag}(Q_1, Q_2, \ldots, Q_r)$.*

Let $\overline{Z} \in \mathbb{V}^{m\times n}$ be given. We use $\overline{\nu}_1 > \overline{\nu}_2 > \ldots > \overline{\nu}_r > 0$ to denote the nonzero distinct singular values of $\overline{Z}$. Let a_l, $l = 1, \ldots, r$, a, b and c be the index sets defined by

$$
\begin{aligned}
&a_l := \{i \mid \sigma_i(\overline{Z}) = \overline{\nu}_l,\ 1 \le i \le m\},\quad l = 1, \ldots, r,\quad a := \{i \mid \sigma_i(\overline{Z}) > 0, 1 \le i \le m\},\\
&b := \{i \mid \sigma_i(\overline{Z}) = 0,\ 1 \le i \le m\} \quad \text{and} \quad c := \{m+1, \ldots, n\}\,.
\end{aligned}
\tag{4}
$$

By combining Propositions 1 and 2 and [12, Proposition 5] with the mixed symmetric property of g, one can check the following result on the well-definedness of spectral operators readily. For simplicity, we omit the detailed proofs here.

Theorem 1 *Let $g : \mathcal{Y} \to \mathcal{Y}$ be mixed symmetric on $\kappa_{\mathcal{N}}$. Then the spectral operator $G : \mathcal{N} \to \mathcal{X}$ defined in Definition 2 with respect to g is well-defined.*

3 Continuity, directional and Fréchet differentiability

In this section, we will first focus on the study of spectral operators for the case that $\mathcal{X} \equiv \mathbb{V}^{m\times n}$. The corresponding extensions for the spectral operators defined on the general Cartesian product of several matrix spaces will be presented in Sect. 4. Let $\mathcal{N}$ be a given nonempty open set in $\mathbb{V}^{m\times n}$. Suppose that $g : \mathbb{R}^m \to \mathbb{R}^m$ is mixed symmetric, with respect to $\mathcal{P} \equiv \pm\mathbb{P}^m$ (called absolutely symmetric in this case), on an open set $\hat{\sigma}_{\mathcal{N}}$ in $\mathbb{R}^m$ containing $\sigma_{\mathcal{N}} := \{\sigma(X) \mid X \in \mathcal{N}\}$. The spectral operator $G : \mathcal{N} \to \mathbb{V}^{m\times n}$ with respect to g defined in Definition 2 then takes the form of $G(X) = U\left[\mathrm{Diag}(g(\sigma(X)))\quad 0\right]V^{\mathbb{T}}$, $X \in \mathcal{N}$, where $(U, V) \in \mathbb{O}^{m,n}(X)$. For the given $\overline{X} \in \mathcal{N}$, consider the singular value decomposition (SVD) for $\overline{X}$, i.e.,

$$
\overline{X} = \overline{U}\left[\Sigma(\overline{X})\quad 0\right]\overline{V}^{\mathbb{T}}, \tag{5}
$$

where $\Sigma(\overline{X})$ is an $m \times m$ diagonal matrix whose i-th diagonal entry is $\sigma_i(\overline{X})$, $\overline{U} \in \mathbb{O}^m$ and $\overline{V} = \left[\overline{V}_1\quad \overline{V}_2\right] \in \mathbb{O}^n$ with $\overline{V}_1 \in \mathbb{V}^{n\times m}$ and $\overline{V}_2 \in \mathbb{V}^{n\times(n-m)}$. Let $\overline{\sigma} := \sigma(\overline{X}) \in \mathbb{R}^m$. Let $a, b, c, a_l, l = 1, \ldots, r$ be the index sets defined by (4) with $\overline{Z}$ being replaced by $\overline{X}$. Denote $\bar{a} := \{1, \ldots, n\} \setminus a$. For each $i \in \{1, \ldots, m\}$, we also define $l_i(\overline{X})$ to be the number of singular values which are equal to $\sigma_i(\overline{X})$ but are ranked before i (including i), and $\tilde{l}_i(\overline{X})$ to be the number of singular values which are equal to $\sigma_i(\overline{X})$ but are ranked after i (excluding i), i.e., define $l_i(\overline{X})$ and $\tilde{l}_i(\overline{X})$ such that

$$
\begin{aligned}
\sigma_1(\overline{X}) \ge \ldots \ge \sigma_{i-l_i(\overline{X})}(\overline{X}) > \sigma_{i-l_i(\overline{X})+1}(\overline{X}) = \ldots = \sigma_i(\overline{X}) = \ldots = \sigma_{i+\tilde{l}_i(\overline{X})}(\overline{X})\\
> \sigma_{i+\tilde{l}_i(\overline{X})+1}(\overline{X}) \ge \ldots \ge \sigma_m(\overline{X})\,.
\end{aligned}
\tag{6}
$$

In later discussions, when the dependence of l_i and $\tilde{l}_i$ on $\overline{X}$ is clear from the context, we often drop $\overline{X}$ from these notations for convenience. We define two linear matrix operators $S : \mathbb{V}^{p\times p} \to \mathbb{S}^p$, $T : \mathbb{V}^{p\times p} \to \mathbb{V}^{p\times p}$ by

$$S(Y) := \frac{1}{2}(Y + Y^{\mathbb{T}}), \quad T(Y) := \frac{1}{2}(Y - Y^{\mathbb{T}}), \quad Y \in \mathbb{V}^{p\times p}. \tag{7}$$

Next, we introduce some notations which are used in later discussions. For any given $X \in \mathcal{N}$, let $\sigma = \sigma(X)$. For the mapping g, we define three matrices $\mathcal{E}_1^0(\sigma), \mathcal{E}_2^0(\sigma) \in \mathbb{R}^{m\times m}$ and $\mathcal{F}^0(\sigma) \in \mathbb{R}^{m\times(n-m)}$ (depending on $X \in \mathcal{N}$) by

$$(\mathcal{E}_1^0(\sigma))_{ij} := \begin{cases} (g_i(\sigma) - g_j(\sigma))/(\sigma_i - \sigma_j) & \text{if } \sigma_i \neq \sigma_j\,, \\ 0 & \text{otherwise}\,, \end{cases} \quad i, j \in \{1, \ldots, m\}\,, \tag{8}$$

$$(\mathcal{E}_2^0(\sigma))_{ij} := \begin{cases} (g_i(\sigma) + g_j(\sigma))/(\sigma_i + \sigma_j) & \text{if } \sigma_i + \sigma_j \neq 0\,, \\ 0 & \text{otherwise}\,, \end{cases} \quad i, j \in \{1, \ldots, m\}\,, \tag{9}$$

$$(\mathcal{F}^0(\sigma))_{ij} := \begin{cases} g_i(\sigma)/\sigma_i & \text{if } \sigma_i \neq 0\,, \\ 0 & \text{otherwise,} \end{cases} \quad i \in \{1, \ldots, m\}, \quad j \in \{1, \ldots, n-m\}\,. \tag{10}$$

When the dependence of $\mathcal{E}_1^0(\sigma)$, $\mathcal{E}_2^0(\sigma)$ and $\mathcal{F}^0(\sigma)$ on σ is clear from the context, we often drop σ from these notations. In particular, let $\overline{\mathcal{E}}_1^0, \overline{\mathcal{E}}_2^0 \in \mathbb{V}^{m\times m}$ and $\overline{\mathcal{F}}^0 \in \mathbb{V}^{m\times(n-m)}$ be the matrices defined by (8)–(10) with respect to $\overline{\sigma} = \sigma(\overline{X})$. Since g is absolutely symmetric at $\overline{\sigma}$, we know that for all $i \in a_l$, $1 \leq l \leq r$, the function values $g_i(\overline{\sigma})$ are the same (denoted by $\bar{g}_l$). Therefore, for any $X \in \mathcal{N}$, define

$$G_S(X) := \sum_{l=1}^{r} \bar{g}_l \mathcal{U}_l(X) \quad \text{and} \quad G_R(X) := G(X) - G_S(X), \tag{11}$$

where $\mathcal{U}_l(X) := \sum_{i\in a_l} u_i v_i^{\mathbb{T}}$ with $\mathbb{O}^{m,n}(X)$. The following lemma on the differentiablity of G_S follows from the derivative formula of Löwner's Hermitian operators (see e.g., [1]). By constructing a special Löwner's non-Hermitian operator and employing the relationship between the SVD of a given $X \in \mathbb{V}^{m\times n}$ and the eigenvalue decomposition of its extended symmetric counterpart $\begin{bmatrix} 0 & X \\ X^{\mathbb{T}} & 0 \end{bmatrix} \in \mathbb{S}^{m+n}$, one can derive the corresponding derivative formula of $\mathcal{U}_l$, especially the three components $\mathcal{E}_1^0(\sigma), \mathcal{E}_2^0(\sigma)$ and $\mathcal{F}^0(\sigma)$ defined by (8)–(10) (see [23, Section 5.1] for details).

Lemma 1 *Let $G_S : \mathcal{N} \to \mathbb{V}^{m\times n}$ be defined by (11). Then, there exists an open neighborhood $\mathcal{B}$ of $\overline{X}$ in $\mathcal{N}$ such that G_S is twice continuously differentiable on $\mathcal{B}$, and for any $\mathbb{V}^{m\times n} \ni H \to 0$, $G_S(\overline{X} + H) - G_S(\overline{X}) = G'_S(\overline{X})H + O(\|H\|^2)$ with*

$$G'_S(\overline{X})H = \overline{U}\big[\overline{\mathcal{E}}_1^0 \circ S(\overline{U}^{\mathbb{T}} H \overline{V}_1) + \overline{\mathcal{E}}_2^0 \circ T(\overline{U}^{\mathbb{T}} H \overline{V}_1) \quad \overline{\mathcal{F}}^0 \circ (\overline{U}^{\mathbb{T}} H \overline{V}_2)\big]\overline{V}^{\mathbb{T}}. \tag{12}$$

Lemma 1 says that in an open neighborhood of $\overline{X}$, G can be decomposed into a "smooth part" G_S plus a "nonsmooth part" G_R. As we will see in the later developments, this decomposition simplifies many of our proofs.

Next, we will first study the continuity of spectral operators. The following simple observation essentially follows from the absolutely symmetric property of g on $\hat{\sigma}_{\mathcal{N}}$, directly.

Proposition 3 *Let $U \in \mathbb{O}^m$ and $V = [V_1 \quad V_2] \in \mathbb{O}^n$ with $V_1 \in \mathbb{V}^{n\times m}$ and $V_2 \in \mathbb{V}^{n\times(n-m)}$ be given. Let $y \in \hat{\sigma}_{\mathcal{N}}$. Then, for $Y := U\left[\mathrm{Diag}(y) \quad 0\right]V^{\mathbb{T}}$ it always holds that $G(Y) = U\left[\mathrm{Diag}(g(y)) \quad 0\right]V^{\mathbb{T}} = U\mathrm{Diag}(g(y))V_1^{\mathbb{T}}$.*

Proof Let $P \in \pm\mathbb{P}^m$ be a signed permutation matrix such that $Py = |y|^{\downarrow}$. Then, we know that $\sigma(Y) = |y|^{\downarrow}$ and Y has the following SVD

$$Y = U\left[P^{\mathbb{T}}\mathrm{Diag}(|y|^{\downarrow})W \quad 0\right]V^{\mathbb{T}} = UP^{\mathbb{T}}\left[\mathrm{Diag}(|y|^{\downarrow}) \quad 0\right]\left[V_1W^{\mathbb{T}} \quad V_2\right]^{\mathbb{T}},$$

where $W := |P| \in \mathbb{P}^m$ is the m by m permutation matrix whose (i, j)-th element is the absolute value of the (i, j)-th element of P. Then, we know from Definition 2 that

$$G(Y) = UP^{\mathbb{T}}\left[\mathrm{Diag}(g(|y|^{\downarrow})) \ \ 0\right]\left[V_1W^{\mathbb{T}} \quad V_2\right]^{\mathbb{T}}.$$

Since g is absolutely symmetric at y, one has $\mathrm{Diag}(g(|y|^{\downarrow})) = \mathrm{Diag}(g(Py)) = \mathrm{Diag}(Pg(y)) = P\mathrm{Diag}(g(y))W^{\mathbb{T}}$. Thus, $G(Y) = UP^{\mathbb{T}}\left[P\mathrm{Diag}(g(y))W^{\mathbb{T}} \ \ 0\right]\left[V_1W^{\mathbb{T}} \quad V_2\right]^{\mathbb{T}} = U\left[\mathrm{Diag}(g(y)) \ \ 0\right]V^{\mathbb{T}}$, which proves the conclusion. $\square$

By using [12, Proposition 7], we have the following result on the continuity of the spectral operator G.

Theorem 2 *Suppose that $\overline{X} \in \mathcal{N}$ has the SVD (5). The spectral operator G is continuous at $\overline{X}$ if and only if g is continuous at $\sigma(\overline{X})$.*

Proof " $\Longleftarrow$" Let $X \in \mathcal{N}$. Denote $H = X - \overline{X}$ and $\sigma = \sigma(X)$. Let $U \in \mathbb{O}^m$ and $V \in \mathbb{O}^n$ be such that $X = \overline{X} + H = U\left[\Sigma(X) \quad 0\right]V^{\mathbb{T}}$. Then, we know from (5) that $\left[\Sigma(\overline{X}) \quad 0\right] + \overline{U}^{\mathbb{T}}H\overline{V} = \overline{U}^{\mathbb{T}}U\left[\Sigma(X) \quad 0\right]V^{\mathbb{T}}\overline{V}$. It follows from [12, (31) in Proposition 7] that for any X sufficiently close to $\overline{X}$, there exist $Q \in \mathbb{O}^{|a|}$, $Q' \in \mathbb{O}^{|b|}$ and $Q'' \in \mathbb{O}^{n-|a|}$ such that

$$\overline{U}^{\mathbb{T}}U = \begin{bmatrix} Q & 0 \\ 0 & Q' \end{bmatrix} + O(\|H\|) \quad \text{and} \quad \overline{V}^{\mathbb{T}}V = \begin{bmatrix} Q & 0 \\ 0 & Q'' \end{bmatrix} + O(\|H\|), \tag{13}$$

where $Q = \mathrm{Diag}(Q_1, Q_2, \ldots, Q_r)$, $Q_l \in \mathbb{O}^{|a_l|}$. On the other hand, from the definition of the spectral operator G one has $U^{\mathbb{T}}\big(G(X) - G(\overline{X})\big)V = \left[\mathrm{Diag}(g(\sigma)) \quad 0\right] - U^{\mathbb{T}}\overline{U}\left[\mathrm{Diag}(g(\overline{\sigma})) \quad 0\right]\overline{V}^{\mathbb{T}}V$. Thus, we obtain from (13) and Proposition 1 that for any X sufficiently close to $\overline{X}$, $U^{\mathbb{T}}\big(G(X) - G(\overline{X})\big)V = \left[\mathrm{Diag}(g(\sigma) - g(\overline{\sigma})) \ \ 0\right] + O(\|H\|)$. Therefore, since g is assumed to be continuous at $\overline{\sigma}$, we can conclude that the spectral operator G is continuous at $\overline{X}$.

" $\Longrightarrow$" Suppose that G is continuous at $\overline{X}$. Let $(\overline{U}, \overline{V}) \in \mathbb{O}^{m\times n}(\overline{X})$ be fixed. Choose any $\sigma \in \hat{\sigma}_{\mathcal{N}}$ and denote $X := \overline{U}[\mathrm{Diag}(\sigma) \quad 0]\overline{V}^{\mathbb{T}}$. Then, it follows from Proposition 3 that $G(X) = \overline{U}\mathrm{Diag}(g(\sigma))\overline{V}_1^{\mathbb{T}}$ and $\mathrm{Diag}(g(\sigma) - g(\overline{\sigma})) = \overline{U}^{\mathbb{T}}\big(G(X) - G(\overline{X})\big)\overline{V}_1$. Hence, we know from the assumption that g is continuous at $\overline{\sigma}$. $\square$

Secondly, we study the directional differentiability of spectral operators. Let $\mathcal{Z}$ and $\mathcal{Z}'$ be two finite dimensional real Euclidean spaces and $\mathcal{O}$ be an open set in $\mathcal{Z}$. A function $F : \mathcal{O} \subseteq \mathcal{Z} \to \mathcal{Z}'$ is said to be *Hadamard directionally differentiable* at $z \in \mathcal{O}$ if the limit

$$\lim_{t\downarrow 0,\ h'\to h} \frac{F(z + th') - F(z)}{t} \quad \text{exists for any } h \in \mathcal{Z}\,. \tag{14}$$

It is clear that if F is Hadamard directionally differentiable at z, then F is directionally differentiable at z, and the limit in (14) equals the directional derivative $F'(z; h)$ for any $h \in \mathcal{Z}$.

Assume that the g is directionally differentiable at $\overline{\sigma}$. Then, from the definition of directional derivative and the absolutely symmetry of g on the nonempty open set $\hat{\sigma}_{\mathcal{N}}$, it is easy to see that the directional derivative $g'(\overline{\sigma}; \cdot) : \mathbb{R}^m \to \mathbb{R}^m$ satisfies

$$g'(\overline{\sigma}; Qh) = Qg'(\overline{\sigma}; h) \quad \forall\, Q \in \pm\mathbb{P}^m_{\overline{\sigma}} \quad \text{and} \quad \forall\, h \in \mathbb{R}^m\,, \tag{15}$$

where $\pm\mathbb{P}^m_{\overline{\sigma}}$ is the subset defined with respect to $\overline{\sigma}$ by $\pm\mathbb{P}^m_{\overline{\sigma}} := \{Q \in \pm\mathbb{P}^m \mid \overline{\sigma} = Q\overline{\sigma}\}$. Since $\overline{\sigma}_i \neq \overline{\sigma}_j > 0$ if $i \in a_l$ and $j \in a_{l'}$ for all $l, l' = 1, \ldots, r$ with $l \neq l'$, we know that $Q \in \pm\mathbb{P}^m_{\overline{\sigma}}$ if and only if

$$Q = \mathrm{Diag}\big(Q_1, \ldots, Q_r, Q_{r+1}\big) \quad \text{with} \quad Q_l \in \mathbb{P}^{|a_l|}, \quad l = 1, \ldots, r \quad \text{and} \quad Q_{r+1} \in \pm\mathbb{P}^{|b|}\,. \tag{16}$$

Denote $\mathcal{V} := \mathbb{R}^{|a_1|} \times \ldots \times \mathbb{R}^{|a_r|} \times \mathbb{R}^{|b|}$. For any $h \in \mathcal{V}$, we rewrite $g'(\overline{\sigma}; h)$ in the following form $\phi(h) := g'(\overline{\sigma}; h) = (\phi_1(h), \ldots, \phi_r(h), \phi_{r+1}(h))$ with $\phi_l(h) \in \mathbb{R}^{|a_l|}$, $l = 1, \ldots, r$ and $\phi_{r+1}(h) \in \mathbb{R}^{|b|}$. Therefore, it follows from (15) and (16) that the function $\phi : \mathcal{V} \to \mathcal{V}$ is a mixed symmetric mapping, with respect to $\mathbb{P}^{|a_1|} \times \ldots \times \mathbb{P}^{|a_r|} \times \pm\mathbb{P}^{|b|}$. Let $\mathcal{W} := \mathbb{S}^{|a_1|} \times \ldots \times \mathbb{S}^{|a_r|} \times \mathbb{V}^{|b|\times(n-|a|)}$. Define the spectral operator $\Phi : \mathcal{W} \to \mathcal{W}$ with respect to the mixed symmetric mapping ϕ as follows: for any $W = (W_1, \ldots, W_r, W_{r+1}) \in \mathcal{W}$,

$$\Phi(W) := \big(\Phi_1(W), \ldots, \Phi_r(W), \Phi_{r+1}(W)\big) \tag{17}$$

with $\Phi_l(W) = \widetilde{P}_l \mathrm{Diag}(\phi_l(\kappa(W)))\widetilde{P}_l^{\mathbb{T}}$ if $1 \leq l \leq r$ and $\Phi_l(W) = \widetilde{M}\mathrm{Diag}(\phi_l(\kappa(W)))\,\widetilde{N}_1^{\mathbb{T}}$ if $l = r+1$, where $\kappa(W) := (\lambda(W_1), \ldots, \lambda(W_r), \sigma(W_{r+1})) \in \mathbb{R}^m$; $\widetilde{P}_l \in \mathbb{O}^{|a_l|}(W_l)$; and $(\widetilde{M}, \widetilde{N}) \in \mathbb{O}^{|b|, n-|a|}(W_{r+1})$, $\widetilde{N} := \big[\widetilde{N}_1 \quad \widetilde{N}_2\big]$ with $\widetilde{N}_1 \in \mathbb{V}^{(n-|a|)\times|b|}$, $\widetilde{N}_2 \in \mathbb{V}^{(n-|a|)\times(n-m)}$. From Theorem 1, we know that Φ is well defined on $\mathcal{W}$.

In order to present the directional differentiability results for the spectral operator G, we define the following *first divided directional difference* $g^{[1]}(\overline{X}; H) \in \mathbb{V}^{m\times n}$ of g at $\overline{X}$ along the direction $H \in \mathbb{V}^{m\times n}$ by

$$g^{[1]}(\overline{X}; H) := \left[\overline{\mathcal{E}}_1^0 \circ S(\overline{U}^{\mathbb{T}} H \overline{V}_1) + \overline{\mathcal{E}}_2^0 \circ T(\overline{U}^{\mathbb{T}} H \overline{V}_1) \quad \overline{\mathcal{F}}^0 \circ \overline{U}^{\mathbb{T}} H \overline{V}_2\right] + \widehat{\Phi}(D(H)), \tag{18}$$

where $\overline{\mathcal{E}}_1^0, \overline{\mathcal{E}}_2^0, \overline{\mathcal{F}}^0$ are defined as in (8)-(10) at $\overline{\sigma} = \sigma(\overline{X})$,

$$D(H) := \left(S(\overline{U}_{a_1}^{\mathbb{T}} H \overline{V}_{a_1}), \ldots, S(\overline{U}_{a_r}^{\mathbb{T}} H \overline{V}_{a_r}), \overline{U}_b^{\mathbb{T}} H [\overline{V}_b \quad \overline{V}_2]\right) \in \mathcal{W} \tag{19}$$

and for any $W = (W_1, \ldots, W_r, W_{r+1}) \in \mathcal{W}$, $\widehat{\Phi}(W) \in \mathbb{V}^{m\times n}$ is defined by

$$\widehat{\Phi}(W) := \begin{bmatrix} \mathrm{Diag}\,(\Phi_1(W), \ldots, \Phi_r(W)) & 0 \\ 0 & \Phi_{r+1}(W) \end{bmatrix}. \tag{20}$$

For the directional differentiability of the spectral operator G, we have the following result.

Theorem 3 *Suppose that $\overline{X} \in \mathcal{N}$ has the SVD (5). The spectral operator G is Hadamard directionally differentiable at $\overline{X}$ if and only if g is Hadamard directionally differentiable at $\overline{\sigma} = \sigma(\overline{X})$. In that case, the directional derivative of G at $\overline{X}$ along any direction $H \in \mathbb{V}^{m\times n}$ is given by*

$$G'(\overline{X}; H) = \overline{U} g^{[1]}(\overline{X}; H) \overline{V}^{\mathbb{T}}. \tag{21}$$

Proof "$\Longleftarrow$" Let $H \in \mathbb{V}^{m\times n}$ be any given direction. For any $\mathbb{V}^{m\times n} \ni H' \to H$ and $\tau > 0$, denote $X := \overline{X} + \tau H'$. Consider the SVD of X, i.e.,

$$X = U[\Sigma(X) \quad 0]V^{\mathbb{T}}. \tag{22}$$

Denote $\sigma = \sigma(X)$. For τ and H' sufficiently close to 0 and H, let G_S and G_R be the mappings defined in (11). Then, by Lemma 1, we know that

$$\lim_{\tau\downarrow 0,\ H'\to H} \frac{1}{\tau}(G_S(X) - G_S(\overline{X})) = G_S'(\overline{X})H, \tag{23}$$

where $G_S'(\overline{X})H$ is given by (12). On the other hand, for τ and H' sufficiently close to 0 and H, we have $\mathcal{U}_l(X) = \sum_{i\in a_l} u_i v_i^{\mathbb{T}}$, $l = 1, \ldots, r$ and

$$G_R(X) = G(X) - G_S(X) = \sum_{l=1}^{r} \sum_{i\in a_l} [g_i(\sigma) - g_i(\overline{\sigma})] u_i v_i^{\mathbb{T}} + \sum_{i\in b} g_i(\sigma) u_i v_i^{\mathbb{T}}. \tag{24}$$

For τ and H' sufficiently close to 0 and H, denote $\Delta_l(\tau, H') = \frac{1}{\tau}\sum_{i\in a_l}[g_i(\sigma) - g_i(\overline{\sigma})]u_i v_i^{\mathbb{T}}$, $l = 1, \ldots, r$ and $\Delta_{r+1}(\tau, H') = \frac{1}{\tau}\sum_{i\in b} g_i(\sigma) u_i v_i^{\mathbb{T}}$.

Firstly, consider the case that $\overline{X} = [\Sigma(\overline{X}) \quad 0]$. Then, from the directional differentiability of the singular value functions (see e.g., [23, Section 5.1] or [12, Proposition 6]), we know that for any τ and $H' \in \mathbb{V}^{m\times n}$ sufficiently close to 0 and H,

$$\sigma(X) = \sigma(\overline{X}) + \tau\sigma'(\overline{X}; H') + O(\tau^2\|H'\|^2)\,, \tag{25}$$

where $(\sigma'(\overline{X}; H'))_{a_l} = \lambda(S(H'_{a_l a_l})), l = 1, \ldots, r$ and $(\sigma'(\overline{X}; H'))_b = \sigma([H'_{bb} \;\; H'_{bc}])$. Denote $h' := \sigma'(\overline{X}; H')$ and $h := \sigma'(\overline{X}; H)$. By using the fact that the singular value functions of a general matrix are globally Lipschitz continuous, we know that

$$\lim_{\tau\downarrow 0,\ H'\to H} (h' + O(\tau\|H'\|^2)) = h\,. \tag{26}$$

Since g is assumed to be Hadamard directionally differentiable at $\overline{\sigma}$, we have

$$\begin{aligned}\lim_{\tau\downarrow 0,\ H'\to H} \frac{g(\sigma) - g(\overline{\sigma})}{\tau} &= \lim_{\tau\downarrow 0,\ H'\to H} \frac{1}{\tau}[g(\overline{\sigma} + \tau(h' + O(\tau\|H'\|^2))) - g(\overline{\sigma})] \\ &= g'(\overline{\sigma}; h) = \phi(h)\,,\end{aligned}$$

where $\phi \equiv g'(\overline{\sigma}; \cdot) : \mathbb{R}^m \to \mathbb{R}^m$ satisfies the condition (15). By noting that $u_i v_i^{\mathbb{T}}$, $i = 1, \ldots, m$ are uniformly bounded, we know that for τ and H' sufficiently close to 0 and H, $\Delta_l(\tau, H') = U_{a_l}\mathrm{Diag}(\phi_l(h))V_{a_l}^{\mathbb{T}} + o(1)$, $l = 1, \ldots, r$ and $\Delta_{r+1}(\tau, H') = U_b\mathrm{Diag}(\phi_{r+1}(h))V_b^{\mathbb{T}} + o(1)$. By [12, (31) in Proposition 7], we obtain that there exist $Q_l \in \mathbb{O}^{|a_l|}$, $l = 1, \ldots, r$, $M \in \mathbb{O}^{|b|}$ and $N = [N_1 \quad N_2] \in \mathbb{O}^{n-|a|}$ with $N_1 \in \mathbb{V}^{(n-|a|)\times|b|}$ and $N_2 \in \mathbb{V}^{(n-|a|)\times(n-m)}$ (depending on τ and H') such that

$$\begin{aligned}&U_{a_l} = \begin{bmatrix} O(\tau\|H'\|) \\ Q_l + O(\tau\|H'\|) \\ O(\tau\|H'\|) \end{bmatrix}, \quad V_{a_l} = \begin{bmatrix} O(\tau\|H'\|) \\ Q_l + O(\tau\|H'\|) \\ O(\tau\|H'\|) \end{bmatrix} \; l = 1, \ldots, r\,, \\ &U_b = \begin{bmatrix} O(\tau\|H'\|) \\ M + O(\tau\|H'\|) \end{bmatrix}, \quad [V_b \quad V_c] = \begin{bmatrix} O(\tau\|H'\|) \\ N + O(\tau\|H'\|) \end{bmatrix}.\end{aligned}$$

Thus, we have

$$\Delta_l(\tau, H') = \begin{bmatrix} 0 & 0 & 0 \\ 0 & Q_l\mathrm{Diag}(\phi_l(h))Q_l^{\mathbb{T}} & 0 \\ 0 & 0 & 0 \end{bmatrix} + O(\tau\|H'\|) + o(1), \quad l = 1, \ldots, r, \tag{27}$$

$$\Delta_{r+1}(\tau, H') = \begin{bmatrix} 0 & 0 \\ 0 & M\mathrm{Diag}(\phi_{r+1}(h))N_1^{\mathbb{T}} \end{bmatrix} + O(\tau\|H'\|) + o(1)\,. \tag{28}$$

We know from [12, (32) and (33) in Proposition 7] that

$$S(H'_{a_la_l}) = S(H_{a_la_l}) + o(1) = \frac{1}{\tau} Q_l[\Sigma(X)_{a_la_l} - \overline{\nu}_l I_{|a_l|}]Q_l^{\mathbb{T}} + O(\tau\|H'\|^2), \quad l = 1,\ldots,r, \tag{29}$$

$$[H'_{bb} \quad H'_{bc}] = [H_{bb} \quad H_{bc}] + o(1) = \frac{1}{\tau} M[\Sigma(X)_{bb} - \overline{\nu}_{r+1} I_{|b|}]N_1^{\mathbb{T}} + O(\tau\|H'\|^2). \tag{30}$$

Since Q_l, $l = 1,\ldots,r$, M and N are uniformly bounded, by taking subsequences if necessary, we may assume that when $\tau \downarrow 0$ and $H' \to H$, Q_l, M and N converge to $\widetilde{Q}_l$, $\widetilde{M}$ and $\widetilde{N}$, respectively. Therefore, by taking limits in (29) and (30), we obtain from (25) and (26) that $S(H_{a_la_l}) = \widetilde{Q}_l \Lambda(S(H_{a_la_l}))\widetilde{Q}_l^{\mathbb{T}}$, $l = 1,\ldots,r$ and $[H_{bb} \quad H_{bc}] = \widetilde{M}[\Sigma([H_{bb} \quad H_{bc}]) \quad 0]\widetilde{N}^{\mathbb{T}} = \widetilde{M}\Sigma([H_{bb} \quad H_{bc}])\widetilde{N}_1^{\mathbb{T}}$. Hence, by using the notation (17), we know from (24), (27), (28) and (20) that

$$\lim_{\tau\downarrow 0,\ H'\to H} \frac{1}{\tau} G_R(X) = \lim_{\tau\downarrow 0,\ H'\to H} \sum_{l=1}^{r+1} \Delta_l(\tau, H') \;=\; \widehat{\Phi}(D(H)), \tag{31}$$

where $D(H) = \big(S(H_{a_1a_1}),\ldots,S(H_{a_ra_r}), H_{b\bar{a}}\big)$.

To prove the conclusion for the general case of $\overline{X}$, rewrite (22) as

$$[\Sigma(\overline{X}) \quad 0] + \overline{U}^{\mathbb{T}} H' \overline{V} = \overline{U}^{\mathbb{T}} U[\Sigma(X) \quad 0]V^{\mathbb{T}}\overline{V}.$$

Let $\widetilde{U} := \overline{U}^{\mathbb{T}} U$, $\widetilde{V} := \overline{V}^{\mathbb{T}} V$ and $\widetilde{H} = \overline{U}^{\mathbb{T}} H \overline{V}$. Denote $\widetilde{X} := [\Sigma(\overline{X}) \quad 0] + \overline{U}^{\mathbb{T}} H' \overline{V}$. Then, we obtain that $G_R(X) = \overline{U} G_R(\widetilde{X})\overline{V}^{\mathbb{T}}$. Thus, we know from (31) that

$$\lim_{\tau\downarrow 0,\ H'\to H} \frac{1}{\tau} G_R(X) = \overline{U}\widehat{\Phi}(D(\widetilde{H}))\overline{V}^{\mathbb{T}}. \tag{32}$$

Therefore, by combining (23) and (32) and noting that $G(\overline{X}) = G_S(\overline{X})$, we obtain that for any given $H \in \mathbb{V}^{m\times n}$,

$$\lim_{\tau\downarrow 0,\ H'\to H} \frac{G(X) - G(\overline{X})}{\tau} = \lim_{\tau\downarrow 0,\ H'\to H} \frac{G_S(X) - G_S(\overline{X}) + G_R(X)}{\tau} = \overline{U} g^{[1]}(\overline{X}; \widetilde{H})\overline{V}^{\mathbb{T}},$$

where $g^{[1]}(\overline{X}; \widetilde{H})$ is given by (18). This implies that G is Hadamard directionally differentiable at $\overline{X}$ and (21) holds.

"$\Longrightarrow$" Suppose that G is Hadamard directionally differentiable at $\overline{X}$. Let $(\overline{U}, \overline{V}) \in \mathbb{O}^{m\times n}(\overline{X})$ be fixed. For any given direction $h \in \mathbb{R}^m$, suppose that $\mathbb{R}^m \ni h' \to h$. Denote $H' := \overline{U}[\mathrm{Diag}(h') \quad 0]\overline{V}^{\mathbb{T}}$ and $H := \overline{U}[\mathrm{Diag}(h) \quad 0]\overline{V}^{\mathbb{T}}$. Then, we have $H' \to H$ as $h' \to h$. Since for all $\tau > 0$ and h' sufficiently close to 0 and h,

$\sigma := \overline{\sigma} + \tau h' \in \hat{\sigma}_{\mathcal{N}}$, we know from Proposition 3 that for all $\tau > 0$ and h' sufficiently close to 0 and h, $G(\overline{X} + \tau H') = \overline{U}\mathrm{Diag}(g(\overline{\sigma} + \tau h'))\overline{V}_1^{\mathbb{T}}$. This implies that $\mathrm{Diag}\big(\lim_{\tau\downarrow 0,\ h'\to h} \frac{g(\overline{\sigma}+\tau h')-g(\overline{\sigma})}{\tau}\big) = \overline{U}^{\mathbb{T}}\big(\lim_{\tau\downarrow 0,\ H'\to H} \frac{G(\overline{X}+\tau H')-G(\overline{X})}{\tau}\big)\overline{V}_1$. Thus, we know from the assumption that $\lim\limits_{\tau\downarrow 0,\ h'\to h} \dfrac{g(\overline{\sigma}+\tau h')-g(\overline{\sigma})}{\tau}$ exists and that g is Hadamard directionally differentiable at $\overline{\sigma}$. □

Remark 1 Note that for a general spectral operator G, we cannot obtain the directional differentiability at $\overline{X}$ if we only assume that g is directionally differentiable at $\sigma(\overline{X})$. In fact, a counterexample can be found in [19]. However, since $\mathbb{V}^{m\times n}$ is a finite dimensional Euclidean space, it is well-known that for locally Lipschitz continuous functions, the directional differentiability in the sense of Hadamard and Gâteaux are equivalent (see e.g. [32, Theorem 1.13], [10, Lemma 3.2], [15, p.259]). Therefore, if G and g are locally Lipschitz continuous near $\overline{X}$ and $\sigma(\overline{X})$, respectively (e.g., the proximal mapping P_f and its vector counterpart), then G is directionally differentiable at $\overline{X}$ if and only if g is directionally differentiable at $\sigma(\overline{X})$.

Finally, we shall study the Fréchet differentiability of spectral operators. For a given $X \in \mathcal{N}$, suppose that the given absolutely symmetric mapping g is F(réchet)-differentiable at $\sigma = \sigma(X)$. The following results on the Jacobian matrix $g'(\sigma)$ can be obtained directly from the assumed absolute symmetry of g on $\hat{\sigma}_{\mathcal{N}}$ and the block structure (16) for any $Q \in \pm\mathbb{P}^m_\sigma$.

Lemma 2 *For any $X \in \mathcal{N}$, suppose that g is F-differentiable at $\sigma = \sigma(X)$. Then, the Jacobian matrix $g'(\sigma)$ has the following property $g'(\sigma) = Q^{\mathbb{T}} g'(\sigma) Q$ for any $Q \in \pm\mathbb{P}^m_\sigma$.*

In particular,

$$\begin{cases} (g'(\sigma))_{ii} = (g'(\sigma))_{i'i'} & \text{if } \sigma_i = \sigma_{i'} \text{ and } i, i' \in \{1,\ldots,m\},\\ (g'(\sigma))_{ij} = (g'(\sigma))_{i'j'} & \text{if } \sigma_i = \sigma_{i'},\ \sigma_j = \sigma_{j'},\ i \neq j,\ i' \neq j' \text{ and } i, i', j, j' \in \{1,\ldots,m\},\\ (g'(\sigma))_{ij} = (g'(\sigma))_{ji} = 0 & \text{if } \sigma_i = 0,\ i \neq j \text{ and } i, j \in \{1,\ldots,m\}. \end{cases}$$

Lemma 2 is a simple extension of [21, Lemma 2.1] for symmetric mappings. But one should note that the Jacobian matrix $g'(\sigma)$ of g at the F-differentiable point σ may not be symmetric since here g is not assumed to be the gradient mapping as in [21, Lemma 2.1]. For example, the absolutely symmetric mapping g defined by [27, (26)] is differentiable at $x = (2, 1)$ by taking $m = 2$ and $\tau = \varepsilon = 1$. However, it is easy to see that the Jacobian matrix $g'(x)$ is not symmetric.

Let $\eta(\sigma) \in \mathbb{R}^m$ be the vector defined as

$$(\eta(\sigma))_i := \begin{cases} (g'(\sigma))_{ii} - (g'(\sigma))_{ij} & \text{if } \exists\, j \in \{1,\ldots,m\} \text{ and } j \neq i \text{ such that } \sigma_i = \sigma_j,\\ (g'(\sigma))_{ii} & \text{otherwise}, \end{cases} \quad i \in \{1,\ldots,m\}. \tag{33}$$

Define the corresponding *divided difference matrix* $\mathcal{E}_1(\sigma) \in \mathbb{R}^{m\times m}$, the *divided addition matrix* $\mathcal{E}_2(\sigma) \in \mathbb{R}^{m\times m}$, the *division matrix* $\mathcal{F}(\sigma) \in \mathbb{R}^{m\times(n-m)}$, respectively, by

$$(\mathcal{E}_1(\sigma))_{ij} := \begin{cases} (g_i(\sigma) - g_j(\sigma))/(\sigma_i - \sigma_j) & \text{if } \sigma_i \neq \sigma_j\,, \\ (\eta(\sigma))_i & \text{otherwise}\,, \end{cases} \quad i, j \in \{1, \ldots, m\}\,, \tag{34}$$

$$(\mathcal{E}_2(\sigma))_{ij} := \begin{cases} (g_i(\sigma) + g_j(\sigma))/(\sigma_i + \sigma_j) & \text{if } \sigma_i + \sigma_j \neq 0\,, \\ (g'(\sigma))_{ii} & \text{otherwise}\,, \end{cases} \quad i, j \in \{1, \ldots, m\}\,, \tag{35}$$

$$(\mathcal{F}(\sigma))_{ij} := \begin{cases} g_i(\sigma)/\sigma_i & \text{if } \sigma_i \neq 0\,, \\ (g'(\sigma))_{ii} & \text{otherwise}, \end{cases} \quad i \in \{1, \ldots, m\}, \quad j \in \{1, \ldots, n-m\}. \tag{36}$$

Define the matrix $\mathcal{C}(\sigma) \in \mathbb{R}^{m\times m}$ to be the difference between $g'(\sigma)$ and $\mathrm{Diag}(\eta(\sigma))$, i.e.,

$$\mathcal{C}(\sigma) := g'(\sigma) - \mathrm{Diag}(\eta(\sigma))\,. \tag{37}$$

When the dependence of η, $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{F}$ and $\mathcal{C}$ on σ is clear from the context, we often drop σ from the corresponding notations. Note that the divided difference matrix $\mathcal{E}_1(\sigma)$ is similar with that of [21, (3.1)] for the symmetric matrix case. Furthermore, the divided addition matrix $\mathcal{E}_2(\sigma)$ and the division matrix $\mathcal{F}(\sigma)$ arise naturally for general non-Hermitian matrices.

Let $\overline{X} \in \mathcal{N}$ be given and denote $\overline{\sigma} = \sigma(\overline{X})$. Denote $\overline{\eta} = \eta(\overline{\sigma}) \in \mathbb{R}^m$ to be the vector defined by (33). Let $\overline{\mathcal{E}}_1$, $\overline{\mathcal{E}}_2$, $\overline{\mathcal{F}}$ and $\overline{\mathcal{C}}$ be the real matrices defined in (34)–(37) with respect to $\overline{\sigma}$. Now, we are ready to state the result on the F-differentiability of spectral operators. It is worth to note that the following result, when reduced to the special symmetric case, is consistent with those obtained in [21].

Theorem 4 *Suppose that the given matrix $\overline{X} \in \mathcal{N}$ has the SVD (5). Then the spectral operator G is F-differentiable at $\overline{X}$ if and only if g is F-differentiable at $\overline{\sigma}$. In that case, the derivative of G at $\overline{X}$ is given by*

$$G'(\overline{X})H = \overline{U}[\overline{\mathcal{E}}_1 \circ S(A) + \mathrm{Diag}\left(\overline{\mathcal{C}}\mathrm{diag}(S(A))\right) + \overline{\mathcal{E}}_2 \circ T(A) \quad \overline{\mathcal{F}} \circ B]\overline{V}^{\mathbb{T}} \quad \forall\, H \in \mathbb{V}^{m\times n}, \tag{38}$$

where $A := \overline{U}^{\mathbb{T}} H \overline{V}_1$, $B := \overline{U}^{\mathbb{T}} H \overline{V}_2$ and for any $X \in \mathbb{V}^{m\times m}$, $\mathrm{diag}(X)$ denotes the column vector consisting of all the diagonal entries of X being arranged from the first to the last. Moreover, G is continuously differentiable at $\overline{X}$ if and only if g is continuously differentiable at $\overline{\sigma} = \sigma(\overline{X})$.

Proof By employing the decomposition G_S and G_R defined in (11), Lemma 1 and the properties of the Jacobian matrix $g'(\overline{\sigma})$ obtained in Lemma 2, one can derive the first part easily in the similar manner to Theorem 3. For brevity, we omit the detail proofs of the first part and only focus on the second part here.

"$\Longleftarrow$" By the assumption, we know from the first part that there exists an open neighborhood $\mathcal{B} \subseteq \mathcal{N}$ of $\overline{X}$ such that the spectral operator G is differentiable on $\mathcal{B}$, and for any $X \in \mathcal{B}$, the derivative $G'(X)$ is given by

$$G'(X)H = U[\mathcal{E}_1 \circ S(A) + \mathrm{Diag}\left(\mathcal{C}\mathrm{diag}(S(A))\right) + \mathcal{E}_2 \circ T(A) \quad \mathcal{F} \circ B]V^{\mathbb{T}} \quad \forall\, H \in \mathbb{V}^{m\times n}\,, \tag{39}$$

where $(U, V) \in \mathbb{O}^{m,n}(X)$, $A = U^{\mathbb{T}} H V_1$, $B = U^{\mathbb{T}} H V_2$ and η, $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{F}$ and $\mathcal{C}$ are defined by (33)–(37) with respect to $\sigma = \sigma(X)$, respectively. Next, we shall prove that

$$\lim_{X\to\overline{X}} G'(X)H \to G'(\overline{X})H \quad \forall\, H \in \mathbb{V}^{m\times n}\,. \tag{40}$$

Firstly, we will show that (40) holds for the special case that $\overline{X} = [\Sigma(\overline{X}) \quad 0]$ and $X = [\Sigma(X) \quad 0] \to \overline{X}$. Let $\{F^{(ij)}\}$ be the standard basis of $\mathbb{V}^{m\times n}$, i.e., for each $i \in \{1,\ldots,m\}$ and $j \in \{1,\ldots,n\}$, $F^{(ij)} \in \mathbb{V}^{m\times n}$ is a matrix whose entries are zeros, except the (i,j)-th entry is 1 or $\sqrt{-1}$. Therefore, we only need to show (40) holds for all $F^{(ij)}$. Note that since $\sigma(\cdot)$ is globally Lipschitz continuous, we know that for X sufficiently close to $\overline{X}$, $\sigma_i \neq \sigma_j$ if $\overline{\sigma}_i \neq \overline{\sigma}_j$. For each $i \in \{1,\ldots,m\}$ and $j \in \{1,\ldots,n\}$, write $F^{(ij)}$ in the following form $F^{(ij)} = \left[F_1^{(ij)} \quad F_2^{(ij)}\right]$ with $F_1^{(ij)} \in \mathbb{V}^{m\times m}$ and $F_2^{(ij)} \in \mathbb{V}^{m\times(n-m)}$. Let us consider the following cases.

Case 1: $i, j \in \{1,\ldots,m\}$ and $i = j$. In this case, since g' is continuous at $\overline{\sigma}$, we know that if $F^{(ij)}$ is real, then $\lim_{X\to\overline{X}} G'(X)F^{(ij)} = \lim_{X\to\overline{X}} \left[\mathrm{Diag}(g'(\sigma)e_i) \quad 0\right] = \left[\mathrm{Diag}(g'(\overline{\sigma})e_i) \quad 0\right] = G'(\overline{X})F^{(ij)}$, where e_i is the vector whose i-th entry is one, and zero otherwise; if $F^{(ij)}$ is complex, then

$$\begin{aligned}
\lim_{X\to\overline{X}} G'(X)F^{(ij)} &= \lim_{X\to\overline{X}} \left[\frac{g_i(\sigma)+g_j(\sigma)}{\sigma_i+\sigma_j} T(F_1^{(ij)}) \quad 0\right] \\
&= \left[\frac{g_i(\overline{\sigma})+g_j(\overline{\sigma})}{\overline{\sigma_i}+\overline{\sigma_j}} T(F_1^{(ij)}) \quad 0\right] = G'(\overline{X})F^{(ij)}\,.
\end{aligned}$$

Case 2: $i, j \in \{1,\ldots,m\}$, $i \neq j$, $\sigma_i = \sigma_j$ and $\overline{\sigma}_i = \overline{\sigma}_j > 0$. Therefore, we know that there exists $l \in \{1,\ldots,r\}$ such that $i, j \in a_l$. Since g' is continuous at $\overline{\sigma}$, we know from (33) that

$$\begin{aligned}
\lim_{X\to\overline{X}} G'(X)F^{(ij)} &= \left[\left((g'(\overline{\sigma}))_{ii} - (g'(\overline{\sigma}))_{ij}\right) S(F_1^{(ij)}) + \frac{g_i(\overline{\sigma})+g_j(\overline{\sigma})}{\overline{\sigma}_i+\overline{\sigma}_j} T(F_1^{(ij)}) \; 0\right] \\
&= G'(\overline{X})F^{(ij)}\,.
\end{aligned}$$

Case 3: $i, j \in \{1,\ldots,m\}$, $i \neq j$, $\sigma_i \neq \sigma_j$ and $\overline{\sigma}_i = \overline{\sigma}_j > 0$. In this case, we know that $G'(X)F^{(ij)} = \left[\frac{g_i(\sigma)-g_j(\sigma)}{\sigma_i-\sigma_j} S(F_1^{(ij)}) + \frac{g_i(\sigma)+g_j(\sigma)}{\sigma_i+\sigma_j} T(F_1^{(ij)}) \quad 0\right]$. Let $s, t \in \mathbb{R}^m$ be two vectors defined by

$$s_p := \begin{cases} \sigma_p & \text{if } p \neq i, \\ \sigma_j & \text{if } p = i \end{cases} \quad \text{and} \quad t_p := \begin{cases} \sigma_p & \text{if } p \neq i, j, \\ \sigma_j & \text{if } p = i, \\ \sigma_i & \text{if } p = j, \end{cases} \qquad p \in \{1,\ldots,m\}\,. \tag{41}$$

It is clear that both s and t converge to $\overline{\sigma}$ as $X \to \overline{X}$. By noting that g is absolutely symmetric on $\hat{\sigma}_{\mathcal{N}}$, we know from (3) that $g_j(\sigma) = g_i(t)$, since the vector t is obtained from σ by swapping the i-th and the j-th components. By the mean value theorem (cf. e.g., [33, Page 68–69]), we have

$$
\begin{aligned}
\frac{g_i(\sigma)-g_j(\sigma)}{\sigma_i-\sigma_j} &= \frac{g_i(\sigma)-g_i(s)+g_i(s)-g_j(\sigma)}{\sigma_i-\sigma_j} \\
&= \frac{\dfrac{\partial g_i(\xi)}{\partial \mu_i}(\sigma_i-\sigma_j)+g_i(s)-g_j(\sigma)}{\sigma_i-\sigma_j} \\
&= \frac{\partial g_i(\xi)}{\partial \mu_i} + \frac{\dfrac{\partial g_i(\hat{\xi})}{\partial \mu_j}(\sigma_j-\sigma_i)+g_i(t)-g_j(\sigma)}{\sigma_i-\sigma_j} = \frac{\partial g_i(\xi)}{\partial \mu_i} - \frac{\partial g_i(\hat{\xi})}{\partial \mu_j},
\end{aligned} \tag{42}
$$

where $\xi \in \mathbb{R}^m$ lies between σ and s and $\hat{\xi} \in \mathbb{R}^m$ is between s and t. Consequently, we have $\xi \to \overline{\sigma}$ and $\widehat{\xi} \to \overline{\sigma}$ as $X \to \overline{X}$. By the continuity of g', we obtain that $\lim_{X\to\overline{X}} \frac{g_i(\sigma)-g_j(\sigma)}{\sigma_i-\sigma_j} = (g'(\overline{\sigma}))_{ii} - (g'(\overline{\sigma}))_{ij}$ and $\lim_{X\to\overline{X}} \frac{g_i(\sigma)+g_j(\sigma)}{\sigma_i+\sigma_j} = \frac{g_i(\overline{\sigma})+g_j(\overline{\sigma})}{\overline{\sigma}_i+\overline{\sigma}_j}$. Therefore, we have

$$
\begin{aligned}
\lim_{X\to\overline{X}} G'(X)F^{(ij)} &= \left[\left((g'(\overline{\sigma}))_{ii}-(g'(\overline{\sigma}))_{ij}\right) S(F_1^{(ij)}) + \frac{g_i(\overline{\sigma})+g_j(\overline{\sigma})}{\overline{\sigma}_i+\overline{\sigma}_j} T(F_1^{(ij)}) \quad 0\right] \\
&= G'(\overline{X})F^{(ij)}\,.
\end{aligned}
$$

Case 4: $i, j \in \{1,\ldots,m\}$, $i \neq j$, $\sigma_i > 0$ or $\sigma_j > 0$ and $\overline{\sigma}_i \neq \overline{\sigma}_j$. Then, we have $\sigma_i > 0$ or $\sigma_j > 0$ and $\sigma_i \neq \sigma_j$. Since g' is continuous at $\overline{\sigma}$, we know that

$$
\begin{aligned}
\lim_{X\to\overline{X}} G'(X)F^{(ij)} &= \left[\frac{g_i(\overline{\sigma})-g_j(\overline{\sigma})}{\overline{\sigma}_i-\overline{\sigma}_j} S(F_1^{(ij)}) + \frac{g_i(\overline{\sigma})+g_j(\overline{\sigma})}{\overline{\sigma}_i+\overline{\sigma}_j} T(F_1^{(ij)}) \quad 0\right] \\
&= G'(\overline{X})F^{(ij)}\,.
\end{aligned}
$$

Case 5: $j \in \{m+1,\ldots,n\}$ and $\overline{\sigma}_i > 0$. Since g' is continuous at $\overline{\sigma}$, we have $\lim_{X\to\overline{X}} G'(X)F^{(ij)} = \lim_{X\to\overline{X}} \left[0 \quad \frac{g_i(\sigma)}{\sigma_i} F_2^{(ij)}\right] = \left[0 \quad \frac{g_i(\overline{\sigma})}{\overline{\sigma}_i} F_2^{(ij)}\right] = G'(\overline{X})F^{(ij)}$.

Case 6: $i, j \in \{1,\ldots,m\}$, $i \neq j$, $\overline{\sigma}_i = \overline{\sigma}_j = 0$ and $\sigma_i = \sigma_j > 0$. Therefore, we know that

$$
G'(X)F^{(ij)} = \left[\left((g'(\sigma))_{ii}-(g'(\sigma))_{ij}\right) S(F_1^{(ij)}) + \frac{g_i(\sigma)+g_j(\sigma)}{\sigma_i+\sigma_j} T(F_1^{(ij)}) \quad 0\right].
$$

We know from (33) and Lemma 2 that

$$
\lim_{X\to\overline{X}} (g'(\sigma))_{ii} = (g'(\overline{\sigma}))_{ii} = \overline{\eta}_i \quad \text{and} \quad \lim_{X\to\overline{X}} (g'(\sigma))_{ij} = (g'(\overline{\sigma}))_{ij} = 0\,. \tag{43}
$$

Let $\hat{s}, \hat{t} \in \mathbb{R}^m$ be two vectors defined by

$$
\hat{s}_p := \begin{cases} \sigma_p & \text{if } p \neq i, \\ -\sigma_j & \text{if } p = i \end{cases} \quad \text{and} \quad \hat{t}_p := \begin{cases} \sigma_p & \text{if } p \neq i, j, \\ -\sigma_j & \text{if } p = i\,, \\ -\sigma_i & \text{if } p = j\,, \end{cases} \qquad p \in \{1,\ldots,m\}. \tag{44}
$$

Also, it clear that both $\hat{s}$ and $\hat{t}$ converge to $\overline{\sigma}$ as $X \to \overline{X}$. Again, by noting that g is absolutely symmetric on $\hat{\sigma}_{\mathcal{N}}$, we know from (3) that $g_i(\sigma) = -g_j(\hat{t})$ and $g_j(\sigma) = -g_i(\hat{t})$. By using similar arguments for deriving (42), we have

$$\frac{g_i(\sigma) + g_j(\sigma)}{\sigma_i + \sigma_j} = \frac{\partial g_i(\zeta)}{\partial \mu_i} + \frac{\partial g_i(\hat{\zeta})}{\partial \mu_j}\,, \tag{45}$$

where $\zeta \in \mathbb{R}^m$ is between σ and $\hat{s}$ and $\hat{\zeta} \in \mathbb{R}^m$ is between $\hat{s}$ and $\hat{t}$. Consequently, we know that $\zeta, \hat{\zeta} \to \overline{\sigma}$ as $X \to \overline{X}$. By the continuity of g', we know from (33) that

$$\lim_{X \to \overline{X}} \frac{g_i(\sigma) + g_j(\sigma)}{\sigma_i + \sigma_j} = (g'(\overline{\sigma}))_{ii} = \overline{\eta}_i\,. \tag{46}$$

Therefore, from (43) and (46), we have $\lim_{X \to \overline{X}} G'(X)F^{(ij)} = \begin{bmatrix}\overline{\eta}_i F_1^{(ij)} & 0\end{bmatrix} = G'(\overline{\sigma})F^{(ij)}$.

Case 7: $i, j \in \{1, \ldots, m\}$, $i \neq j$, $\overline{\sigma}_i = \overline{\sigma}_j = 0$, $\sigma_i \neq \sigma_j$ and $\sigma_i > 0$ or $\sigma_j > 0$. Let s, t and $\hat{s}, \hat{t}$ be defined by (41) and (44), respectively. By the continuity of g', we know from (42) and (45) that $\lim_{X \to \overline{X}} G'(X)F^{(ij)} = \lim_{X \to \overline{X}} \left[\frac{g_i(\sigma) - g_j(\sigma)}{\sigma_i - \sigma_j} S(F_1^{(ij)}) + \frac{g_i(\sigma)+g_j(\sigma)}{\sigma_i+\sigma_j} T(F_1^{(ij)}) \;\; 0\right] = \begin{bmatrix}\overline{\eta}_i F_1^{(ij)} & 0\end{bmatrix} = G'(\overline{X})F^{(ij)}$.

Case 8: $i \neq j \in \{1, \ldots, m\}$, $\overline{\sigma}_i = \overline{\sigma}_j = 0$ and $\sigma_i = \sigma_j = 0$. By the continuity of g', we obtain that

$$\lim_{X \to \overline{X}} G'(X)F^{(ij)} = \lim_{X \to \overline{X}} \begin{bmatrix}(g'(\sigma))_{ii} F_1^{(ij)} & 0\end{bmatrix} = \begin{bmatrix}\overline{\eta}_i F_1^{(ij)} & 0\end{bmatrix} = G'(\overline{X})F^{(ij)}\,.$$

Case 9: $j \in \{m+1, \ldots, n\}$, $\overline{\sigma}_i = 0$ and $\sigma_i > 0$. We know that $G'(X)F^{(ij)} = \begin{bmatrix}0 & \frac{g_i(\sigma)}{\sigma_i} F_2^{(ij)}\end{bmatrix}$. Let $\tilde{s} \in \mathbb{R}^m$ be a vector given by $\tilde{s}_p := \begin{cases}\sigma_p & \text{if } p \neq i, \\ 0 & \text{if } p = i,\end{cases}$ $p \in \{1, \ldots, m\}$. Therefore, we have $\tilde{s}$ converges to $\overline{\sigma}$ as $X \to \overline{X}$. Since g is absolutely symmetric on $\hat{\sigma}_{\mathcal{N}}$, we know that $g_i(\tilde{s}) = 0$. Also, by the mean value theorem, we have $g_i(\sigma)/\sigma_i = (g_i(\sigma) - g_i(\tilde{s}))/\sigma_i = \frac{\partial g_i(\rho)}{\partial \mu_i}$, where $\rho \in \mathbb{R}^m$ is between σ and $\tilde{s}$. Consequently, we have ρ converges to $\overline{\sigma}$ as $X \to \overline{X}$. By the continuity of g', we know from (33) that $\lim_{X \to \overline{X}} \frac{g_i(\sigma)}{\sigma_i} = (g'(\overline{\sigma}))_{ii} = \overline{\eta}_i$. Thus, $\lim_{X \to \overline{X}} G'(X)F^{(ij)} = \lim_{X \to \overline{X}} \begin{bmatrix}0 & \frac{g_i(\sigma)}{\sigma_i} F_2^{(ij)}\end{bmatrix} = \begin{bmatrix}0 & \overline{\eta}_i F_2^{(ij)}\end{bmatrix} = G'(\overline{X})F^{(ij)}$.

Case 10: $j \in \{m+1, \ldots, n\}$, $\overline{\sigma}_i = 0$ and $\sigma_i = 0$. By the continuity of g', we know that

$$\lim_{X \to \overline{X}} G'(X)F^{(ij)} = \begin{bmatrix}0 & (g'(\overline{\sigma}))_{ii} F_2^{(ij)}\end{bmatrix} = G'(\overline{X})F^{(ij)}\,.$$

Finally, for the general case that $X = U\,[\Sigma(X) \;\; 0]\,V^{\mathbb{T}}$ and $\overline{X} = \overline{U}\begin{bmatrix}\Sigma(\overline{X}) & 0\end{bmatrix}\overline{V}^{\mathbb{T}}$, it follows from the first part of this theorem that G is F-differential at X if and only if G is F-differential at $[\Sigma(X) \;\; 0]$ and for any $H \in \mathbb{V}^{m \times n}$, $G'(X)H =$

$U\left(G'([\Sigma(X)\quad 0])(U^T HV)\right)V^T$. Thus, we know from the above analysis that (40) holds, which implies that G is continuously differentiable at $\overline{X}$.

"$\Longrightarrow$" Suppose that G is continuously differentiable at $\overline{X}$. Let $(\overline{U},\overline{V})\in\mathbb{O}^{m\times n}(\overline{X})$ be fixed. For any $\sigma\in\mathbb{R}^m$, define $X:=\overline{U}[\mathrm{Diag}(\sigma)\quad 0]\overline{V}^{\mathbb{T}}$. For any $h\in\mathbb{R}^m$, let $H:=\overline{U}[\mathrm{Diag}(h)\quad 0]\overline{V}^{\mathbb{T}}$. By the derivative formula (38), we know from the assumption that for all σ sufficiently close to $\overline{\sigma}$, $\mathrm{Diag}(g'(\sigma)h)=\overline{U}^{\mathbb{T}}(G'(X)H)\overline{V}_1$ for all $h\in\mathbb{R}^m$. Consequently, g is continuously differentiable at $\overline{\sigma}$. □

Remark 2 In order to compute (38), it appears that one needs to compute and store $\overline{V}_2\in\mathbb{V}^{n\times(n-m)}$ explicitly, which would incur huge memory costs if $n\gg m$. Fortunately, due to the special form of $\overline{\mathcal{F}}$, the explicit computation of $\overline{V}_2$ can be avoided as we shall show next. Let $\bar{f}=(\bar{f}_1,\ldots,\bar{f}_m)^T$ be defined by $\bar{f}_i=g_i(\bar{\sigma})/\bar{\sigma}_i$ if $\bar{\sigma}_i\neq 0$ and $\bar{f}_i=(g'(\bar{\sigma}))_{ii}$ otherwise. Observe that the term in (38) involving $\overline{V}_2$ is given by

$$\begin{aligned}\overline{U}(\overline{\mathcal{F}}\circ(\overline{U}^{\mathbb{T}}H\overline{V}_2))\overline{V}_2^{\mathbb{T}}&=\overline{U}\mathrm{Diag}(\bar{f})\overline{U}^{\mathbb{T}}H(I_n-\overline{V}_1\overline{V}_1^{\mathbb{T}})\\&=\overline{U}\mathrm{Diag}(\bar{f})\overline{U}^{\mathbb{T}}(H-(H\overline{V}_1)\overline{V}_1^{\mathbb{T}}).\end{aligned}$$

Thus, in numerical implementations, the large matrix $\overline{V}_2$ is not needed.

4 Extensions

In this section, we consider the spectral operators defined on the Cartesian product of several real or complex matrices. The corresponding properties, including continuity, directional differentiability and (continuous) differentiability, can be studied in the same fashion as those in Sect. 3 though the analysis for the general case is more involved. For simplicity, we omit the proofs here. For readers who are interested in seeking the complete proofs, we refer them to the PhD thesis of Ding [11] for worked out details.

Without loss of generality, from now on, we assume that $\mathcal{X}=\mathbb{S}^{m_1}\times\mathbb{V}^{m_2\times n_2}$ and $\mathcal{Y}=\mathbb{R}^{m_1}\times\mathbb{R}^{m_2}$ with $m=m_1+m_2$. For any $X=(X_1,X_2)\in\mathbb{S}^{m_1}\times\mathbb{V}^{m_2\times n_2}$, denote $\kappa(X)=(\lambda(X_1),\sigma(X_2))\in\mathcal{Y}$. Let $\mathcal{N}$ be a given nonempty open set in $\mathcal{X}$. Suppose that $g:\mathcal{Y}\to\mathcal{Y}$ is mixed symmetric, with respect to $\mathcal{P}\equiv\mathbb{P}^{m_1}\times\pm\mathbb{P}^{m_2}$, on an open set $\hat{\kappa}_{\mathcal{N}}$ in $\mathbb{R}^m$ containing $\kappa_{\mathcal{N}}=\{\kappa(X)\in\mathcal{Y}\mid X\in\mathcal{N}\}$. Let $G:\mathcal{X}\to\mathcal{X}$ be the corresponding spectral operator defined in Definition 2.

Let $\overline{X}=(\overline{X}_1,\overline{X}_2)\in\mathcal{N}$ be given. Suppose the given $\overline{X}_1\in\mathbb{S}^{m_1}$ and $\overline{X}_2\in\mathbb{V}^{m_2\times n_2}$ have the following decompositions

$$\overline{X}_1=\overline{P}\mathrm{Diag}(\lambda(\overline{X}_1))\overline{P}^{\mathbb{T}}\quad\text{and}\quad\overline{X}_2=\overline{U}[\mathrm{Diag}(\sigma(\overline{X}_2))\quad 0]\overline{V}^{\mathbb{T}},\tag{47}$$

where $\overline{P}\in\mathbb{O}^{m_1}$, $\overline{U}\in\mathbb{O}^{m_2}$ and $\overline{V}=\begin{bmatrix}\overline{V}_1 & \overline{V}_2\end{bmatrix}\in\mathbb{O}^{n_2}$ with $\overline{V}_1\in\mathbb{V}^{n_2\times m_2}$ and $\overline{V}_2\in\mathbb{V}^{n_2\times(n_2-m_2)}$. Denote $\overline{\lambda}:=\lambda(\overline{X}_1)$, $\overline{\sigma}:=\sigma(\overline{X}_2)$ and $\overline{\kappa}:=(\overline{\lambda},\overline{\sigma})$. We use $\overline{\nu}_1>\ldots>\overline{\nu}_{r_1}$ to denote the distinct eigenvalues of $\overline{X}_1$ and $\overline{\nu}_{r_1+1}>\ldots>\overline{\nu}_{r_1+r_2}>0$ to denote the distinct nonzero singular values of $\overline{X}_2$. Define the index sets $a_l:=$

$\{i \mid \overline{\lambda}_i = \overline{\nu}_l,\ 1 \le i \le m_1\}$, $l = 1, \ldots, r_1$, $a_l := \{i \mid \overline{\sigma}_i = \overline{\nu}_l,\ 1 \le i \le m_2\}$, $l = r_1 + 1, \ldots, r_1 + r_2$ and $b := \{i \mid \overline{\sigma}_i = 0,\ 1 \le i \le m_2\}$.

First, we have the following result on the continuity of spectral operators.

Theorem 5 *Let $\overline{X} = (\overline{X}_1, \overline{X}_2) \in \mathcal{N}$ be given. Suppose that $\overline{X}_1$ and $\overline{X}_2$ have the decompositions (47). The spectral operator G is continuous at $\overline{X}$ if and only if g is continuous at $\kappa(\overline{X})$.*

In order to present the results on the directional differentiability of spectral operators of matrices, we introduce some notations. For the given mixed symmetric mapping $g = (g_1, g_2) : \mathcal{Y} \to \mathcal{Y}$, let $\overline{\mathcal{E}}_1^0 \in \mathbb{S}^{m_2}$, $\overline{\mathcal{E}}_2^0 \in \mathbb{V}^{m_2 \times m_2}$ and $\overline{\mathcal{F}}^0 \in \mathbb{V}^{m_2 \times (n_2 - m_2)}$ be the matrices given by (8)-(10) with respect to $\overline{\kappa} = (\overline{\lambda}, \overline{\sigma})$, and $\overline{\mathcal{A}}^0 \in \mathbb{S}^{m_1}$ be the matrix defined by

$$(\overline{\mathcal{A}}^0)_{ij} := \begin{cases} ((g_1(\overline{\kappa}))_i - (g_1(\overline{\kappa}))_j)/(\overline{\lambda}_i - \overline{\lambda}_j) & \text{if } \overline{\lambda}_i \ne \overline{\lambda}_j, \\ 0 & \text{otherwise,} \end{cases} \quad i, j \in \{1, \ldots, m_1\}.$$

Suppose that g is directionally differentiable at $\overline{\kappa}$. Then, we know that the directional derivative $g'(\overline{\kappa}; \cdot) = \big(g_1'(\overline{\kappa}; \cdot), g_2'(\overline{\kappa}; \cdot)\big) : \mathcal{Y} \to \mathcal{Y}$ satisfies that for any $(Q_1, Q_2) \in \mathcal{P}_{\overline{\kappa}}$ and any $(h_1, h_2) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$,

$$\begin{aligned} &\Big(g_1'(\overline{\kappa}; (Q_1 h_1, Q_2 h_2)), g_2'(\overline{\kappa}; (Q_1 h_1, Q_2 h_2))\Big) \\ &\quad = \Big(Q_1 g_1'(\overline{\kappa}; (h_1, h_2)), Q_2 g_2'(\overline{\kappa}; (h_1, h_2))\Big), \end{aligned} \tag{48}$$

where $\mathcal{P}_{\overline{\kappa}}$ is the subset of $\mathcal{P} \equiv \mathbb{P}^{m_1} \times \pm\mathbb{P}^{m_2}$ defined with respect to $\overline{\kappa}$ by $\mathcal{P}_{\overline{\kappa}} := \{(Q_1, Q_2) \in \mathbb{P}^{m_1} \times \pm\mathbb{P}^{m_2} \mid (\overline{\lambda}, \overline{\sigma}) = (Q_1 \overline{\lambda}, Q_2 \overline{\sigma})\}$. Note that $\overline{\lambda}_i \ne \overline{\lambda}_j$ if $i \in a_l$ and $j \in a_{l'}$ for all $l, l' = 1, \ldots, r_1$ with $l \ne l'$ and $\overline{\sigma}_i \ne \overline{\sigma}_j > 0$ if $i \in a_l$ and $j \in a_{l'}$ for all $l, l' = r_1 + 1, \ldots, r_1 + r_2$ with $l \ne l'$. Therefore, we have $(Q_1, Q_2) \in \mathcal{P}_{\overline{\kappa}}$ if and only if there exist $Q_1^l \in \mathbb{P}^{|a_l|}$, $l = 1, \ldots, r_1$, $Q_2^l \in \mathbb{P}^{|a_l|}$, $l = r_1 + 1, \ldots, r_1 + r_2$ and $Q_2^{r_1+r_2+1} \in \pm\mathbb{P}^{|b|}$ such that

$$Q_1 = \mathrm{Diag}\big(Q_1^1, \ldots, Q_1^{r_1}\big) \in \mathbb{P}^{m_1} \quad \text{and} \quad Q_2 = \mathrm{Diag}\big(Q_2^{r_1+1}, \ldots, Q_2^{r_1+r_2}, Q_2^{r_1+r_2+1}\big) \in \pm\mathbb{P}^{m_2}. \tag{49}$$

Denote $\mathcal{V} := \mathbb{R}^{|a_1|} \times \ldots \times \mathbb{R}^{|a_{r_1+r_2}|} \times \mathbb{R}^{|b|}$. For any $h \in \mathcal{V}$, rewrite $g'(\overline{\kappa}; h) =: \phi(h) \in \mathcal{V}$ as $\phi(h) = \big(\phi_1(h), \ldots, \phi_{r_1+r_2+1}(h)\big)$ with $\phi_l(h) \in \mathbb{R}^{|a_l|}$ for $l = 1, \ldots, r_1 + r_2$ and $\phi_{r_1+r_2+1}(h) \in \mathbb{R}^{|b|}$. Therefore, we know from (48) and (49) that the directional derivative ϕ is mixed symmetric mapping, with respect to $\mathbb{P}^{|a_1|} \times \ldots \times \mathbb{P}^{|a_{r_1+r_2}|} \times \pm\mathbb{P}^{|b|}$. Denote $\mathcal{W} := \mathbb{S}^{|a_1|} \times \ldots \times \mathbb{S}^{|a_{r_1+r_2}|} \times \mathbb{V}^{|b| \times (|b| + n_2 - m_2)}$. Let $\Phi : \mathcal{W} \to \mathcal{W}$ be the corresponding spectral operator defined in Definition 2 with respect to the mixed symmetric mapping ϕ, i.e., for any $W = \big(W_1, \ldots, W_{r_1+r_2}, W_{r_1+r_2+1}\big) \in \mathcal{W}$, $\Phi(W) = \big(\Phi_1(W), \ldots, \Phi_{r_1+r_2}(W), \Phi_{r_1+r_2+1}(W)\big)$ with

$$\Phi_l(W) = \begin{cases} \widetilde{R}_l \mathrm{Diag}(\phi_l(\kappa(W))) \widetilde{R}_l^{\mathbb{T}} & \text{if } l = 1, \ldots, r_1 + r_2, \\ \widetilde{M} \mathrm{Diag}(\phi_{r_1+r_2+1}(\kappa(W))) \widetilde{N}_1^{\mathbb{T}} & \text{if } l = r_1 + r_2 + 1, \end{cases}$$

where $\widetilde{R}_l \in \mathbb{O}^{|a_l|}(W_l)$, $(\widetilde{M}, \widetilde{N}) \in \mathbb{O}^{|b|,|b|+n_2-m_2}(W_{r_1+r_2+1})$ and

$$\kappa(W) = \big(\lambda(W_1), \ldots, \lambda(W_{r_1+r_2}), \sigma(W_{r_1+r_2+1})\big) \in \mathbb{R}^m .$$

Then, the first divided directional difference $g^{[1]}(\overline{X}; H) \in \mathcal{X}$ of g at $\overline{X}$ along the direction $H = (H_1, H_2) \in \mathcal{X}$ is defined by $g^{[1]}(\overline{X}; H) := \Big(g_1^{[1]}(\overline{X}; H),\ g_2^{[1]}(\overline{X}; H)\Big)$ with

$$g_1^{[1]}(\overline{X}; H) = \overline{\mathcal{A}}^0 \circ \overline{P}^{\mathbb{T}} H_1 \overline{P} + \mathrm{Diag}\Big(\Phi_1(D(H)), \ldots, \Phi_{r_1}(D(H))\Big) \in \mathbb{S}^{m_1}$$

and

$$\begin{aligned} g_2^{[1]}(\overline{X}; H) = &\Big[\overline{\mathcal{E}}_1^0 \circ S(\overline{U}^{\mathbb{T}} H_2 \overline{V}_1) + \overline{\mathcal{E}}_2^0 \circ T(\overline{U}^{\mathbb{T}} H_2 \overline{V}_1) \quad \overline{\mathcal{F}}^0 \circ \overline{U}^{\mathbb{T}} H_2 \overline{V}_2\Big] \\ &+ \begin{bmatrix} \mathrm{Diag}\big(\Phi_{r_1+1}(D(H)), \ldots, \Phi_{r_1+r_2}(D(H))\big) & 0 \\ 0 & \Phi_{r_1+r_2+1}(D(H)) \end{bmatrix} \in \mathbb{V}^{m_2 \times n_2} , \end{aligned}$$

where

$$\begin{aligned} D(H) = \Big(&\overline{P}_{a_1}^{\mathbb{T}} H_1 \overline{P}_{a_1}, \ldots, \overline{P}_{a_{r_1}}^{\mathbb{T}} H_1 \overline{P}_{a_{r_1}}, S(\overline{U}_{a_{r_1+1}}^{\mathbb{T}} H_2 \overline{V}_{a_{r_1+1}}), \ldots, S(\overline{U}_{a_{r_1+r_2}}^{\mathbb{T}} \\ &H_2 \overline{V}_{a_{r_1+r_2}}), \overline{U}_b^{\mathbb{T}} H_2 [\overline{V}_b \overline{V}_2]\Big) \in \mathcal{W} . \end{aligned}$$

Now, we are ready to state the results on the directional differentiability of the spectral operator G.

Theorem 6 *Let $\overline{X} = (\overline{X}_1, \overline{X}_2) \in \mathcal{N}$ be given. Suppose that $\overline{X}_1$ and $\overline{X}_2$ have the decompositions (47). The spectral operator G is Hadamard directionally differentiable at $\overline{X}$ if and only if g is Hadamard directionally differentiable at $\kappa(\overline{X})$. In that case, G is directionally differentiable at $\overline{X}$ and the directional derivative at $\overline{X}$ along any direction $H \in \mathcal{X}$ is given by $G'(\overline{X}; H) = \Big(\overline{P} g_1^{[1]}(\overline{X}; H) \overline{P}^{\mathbb{T}},\ \overline{U} g_2^{[1]}(\overline{X}; H) \overline{V}^{\mathbb{T}}\Big)$.*

In order to present the derivative formulas of spectral operators, we introduce the following notation. For the given $\overline{X} = (\overline{X}_1, \overline{X}_2) \in \mathcal{N}$, suppose that g is F-differentiable at $\overline{\kappa}$. Denote by $g'(\overline{\kappa}) \in \mathbb{R}^{m \times m}$ the Jacobian matrix of g at $\overline{\kappa}$. Let $\overline{\eta}_1 \in \mathbb{R}^{m_1}$ and $\overline{\eta}_2 \in \mathbb{R}^{m_2}$ be the vectors defined by for each $i \in \{1, \ldots, m_1\}$,

$$\left(\overline{\eta}_1\right)_i := \begin{cases} (g_1'(\overline{\kappa}))_{ii} - (g_1'(\overline{\kappa}))_{i(i+1)} & \text{if } \exists\, j \in \{1, \ldots, m_1\} \text{ and } j \neq i \text{ such that } \overline{\lambda}_i = \overline{\lambda}_j , \\ (g_1'(\overline{\kappa}))_{ii} & \text{otherwise} , \end{cases}$$

and for each $i \in \{1, \ldots, m_2\}$,

$$\left(\overline{\eta}_2\right)_i := \begin{cases} (g_2'(\overline{\kappa}))_{ii} - (g_2'(\overline{\kappa}))_{i(i+1)} & \text{if } \exists\, j \in \{1, \ldots, m_2\} \text{ and } j \neq i \text{ such that } \overline{\sigma}_i = \overline{\sigma}_j , \\ (g_2'(\overline{\kappa}))_{ii} & \text{otherwise} . \end{cases}$$

Define the corresponding *divided difference matrices* $\overline{\mathcal{A}} \in \mathbb{R}^{m_1 \times m_1}$ by

$$(\overline{\mathcal{A}})_{ij} := \begin{cases} ((g_1(\overline{\kappa}))_i - (g_1(\overline{\kappa}))_j)/(\overline{\lambda}_i - \overline{\lambda}_j) & \text{if } \overline{\lambda}_i \neq \overline{\lambda}_j, \\ (\eta_1(\overline{\kappa}))_i & \text{otherwise,} \end{cases} \quad i, j \in \{1, \ldots, m_1\}.$$

Let $\overline{\mathcal{E}}_1 \in \mathbb{R}^{m_2 \times m_2}$, $\overline{\mathcal{E}}_2 \in \mathbb{R}^{m_2 \times m_2}$ and $\overline{\mathcal{F}} \in \mathbb{R}^{m_2 \times (n_2 - m_2)}$ by the matrices defined by (34)–(37) with respect to $\overline{\kappa}$. Moreover, define the matrices $\overline{\mathcal{C}}_1 \in \mathbb{R}^{m_1 \times m}$ and $\overline{\mathcal{C}}_2 \in \mathbb{R}^{m_2 \times m}$ by $\overline{\mathcal{C}}_1 = g_1'(\overline{\kappa}) - \begin{bmatrix}\mathrm{Diag}(\overline{\eta}_1) & 0\end{bmatrix}$ and $\overline{\mathcal{C}}_2 = g_2'(\overline{\kappa}) - \begin{bmatrix}0 & \mathrm{Diag}(\overline{\eta}_2)\end{bmatrix}$. Then, we have the following results on the F-differentiability of spectral operators.

Theorem 7 *Let $\overline{X} = (\overline{X}_1, \overline{X}_2) \in \mathcal{N}$ be given. Suppose that $\overline{X}_1$ and $\overline{X}_2$ have the decompositions (47). The spectral operator G is (continuously) differentiable at $\overline{X}$ if and only if g is (continuously) differentiable at $\overline{\kappa} = \kappa(\overline{X})$. In that case, the derivative of G at $\overline{X}$ is given by for any $H = (H_1, H_2) \in \mathcal{X}$,*

$$\begin{aligned} G'(\overline{X})(H) = \Big(& \overline{P}[\overline{\mathcal{A}} \circ \overline{P}^{\mathbb{T}} H_1 \overline{P} + \mathrm{Diag}\left(\overline{\mathcal{C}}_1 h\right)]\overline{P}^{\mathbb{T}}, \\ & \overline{U}\left[\overline{\mathcal{E}}_1 \circ S(\overline{U}^{\mathbb{T}} H_2 \overline{V}_1) + \mathrm{Diag}\left(\overline{\mathcal{C}}_2 h\right) + \overline{\mathcal{E}}_2 \circ T(\overline{U}^{\mathbb{T}} H_2 \overline{V}_1) \quad \overline{\mathcal{F}} \circ \overline{U}^{\mathbb{T}} H_2 \overline{V}_2\right] \overline{V}^{\mathbb{T}} \Big), \end{aligned}$$

where $h := \left(\mathrm{diag}\left(\overline{P}^{\mathbb{T}} H_1 \overline{P}\right), \mathrm{diag}\left(S(\overline{U}^{\mathbb{T}} H_2 \overline{V}_1)\right)\right) \in \mathbb{R}^m$.

5 Conclusions

In this paper, we have introduced a class of matrix-valued functions, termed spectral operators of matrices and have systematically studied several fundamental properties of spectral operators, including the well-definedness, continuity, directional differentiability and Fréchet-differentiability. These results provide the necessary theoretical foundations for both the computational and theoretical aspects of many applications such as MOPs. Consequently, one is able to use these results to design some efficient numerical methods for solving large-scale MOPs arising from various applications. For instance, Chen et al. [7] proposed an efficient and robust semismooth Newton-CG dual proximal point algorithm for solving large scale matrix spectral norm approximation problems. The work done in this paper on spectral operators of matrices is by no means complete. Due to the rapid advances in the applications of matrix optimization in different fields, spectral operators of matrices will become even more important and many other properties of spectral operators are waiting to be explored.

Acknowledgements We would like to thank the referees as well as the editors for their constructive comments that have helped to improve the quality of the paper. The research of C. Ding was supported by the National Natural Science Foundation of China under projects No. 11301515, No. 11671387 and No. 11531014.

References

1. Bhatia, R.: Matrix Analysis. Springer, New York (1997)
2. Candès, E.J., Recht, B.: Exact matrix completion via convex optimization. Found. Comput. Math. **9**, 717–772 (2008)

3. Candès, E.J., Tao, T.: The power of convex relaxation: near-optimal matrix completion. IEEE Trans. Inf. Theory **56**, 2053–2080 (2009)
4. Candès, E.J., Li, X., Ma, Y., Wright, J.: Robust principal component analysis? J. ACM **58**, 11 (2011)
5. Chan, Z.X., Sun, D.F.: Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming. SIAM J. Optim. **19**, 370–396 (2008)
6. Chandrasekaran, V., Sanghavi, S., Parrilo, P.A., Willsky, A.: Rank-sparsity incoherence for matrix decomposition. SIAM J. Optim. **21**, 572–596 (2011)
7. Chen, C.H., Liu, Y.J., Sun, D.F., Toh, K.C.: A semismooth Newton-CG dual proximal point algorithm for matrix spectral norm approximation problems. Math. Program. **155**, 435–470 (2016)
8. Chen, X., Qi, H.D., Tseng, P.: Analysis of nonsmooth symmetric-matrix-valued functions with applications to semidefinite complement problems. SIAM J. Optim. **13**, 960–985 (2003)
9. Chu, M., Funderlic, R., Plemmons, R.: Structured low rank approximation. Linear Algebra Appl. **366**, 157–172 (2003)
10. Demyanov, V.F., Rubinov, A.M.: On quasidifferentiable mappings. Optimization **14**, 3–21 (1983)
11. Ding, C.: An introduction to a class of matrix optimization problems. PhD thesis, National University of Singapore. http://www.math.nus.edu.sg/~matsundf/DingChao_Thesis_final.pdf (2012)
12. Ding, C., Sun, D.F., Toh, K.C.: An introduction to a class of matrix cone programming. Math. Program. **144**, 141–179 (2014)
13. Ding, C., Sun, D.F., Ye, J.J.: First order optimality conditions for mathematical programs with semidefinite cone complementarity constraints. Math. Program. **147**, 539–579 (2014)
14. Dobrynin, V.: On the rank of a matrix associated with a graph. Discrete Math. **276**, 169–175 (2004)
15. Flett, T.M.: Differential Analysis. Cambridge University Press, Cambridge (1980)
16. Greenbaum, A., Trefethen, L.N.: GMRES/CR and Arnoldi/Lanczos as matrix approximation problems. SIAM J. Sci. Comput. **15**, 359–368 (1994)
17. Kotlov, A., Lovász, L., Vempala, S.: The Colin de Verdière number and sphere representations of a graph. Combinatorica **17**, 483–521 (1997)
18. Lewis, A.S.: The convex analysis of unitarily invariant matrix functions. J. Convex Anal. **2**, 173–183 (1995)
19. Lewis, A.S.: Derivatives of spectral functions. Math. Oper. Res. **21**, 576–588 (1996)
20. Lewis, A.S., Overton, M.L.: Eigenvalue optimization. Acta Numer. **5**, 149–190 (1996)
21. Lewis, A.S., Sendov, H.S.: Twice differentiable spectral functions. SIAM J. Matrix Anal. Appl. **23**, 368–386 (2001)
22. Lewis, A.S., Sendov, H.S.: Nonsmooth analysis of singular values. Part I: theory. Set-Valued Anal. **13**, 213–241 (2005)
23. Lewis, A.S., Sendov, H.S.: Nonsmooth analysis of singular values. Part II: application. Set-Valued Anal. **13**, 243–264 (2005)
24. Liu, Y.J., Sun, D.F., Toh, K.C.: An implementable proximal point algorithmic framework for nuclear norm minimization. Math. Program. **133**, 399–436 (2012)
25. Löwner, K.: Über monotone matrixfunktionen. Math. Z. **38**, 177–216 (1934)
26. Lovász, L.: On the Shannon capacity of a graph. IEEE Trans. Inf. Theory **25**, 1–7 (1979)
27. Miao, W.M., Sun, D.F., Pan, S.H.: A rank-corrected procedure for matrix completion with fixed basis coefficients. Math. Program. **159**, 289–338 (2016)
28. Mifflin, R.: Semismooth and semiconvex functions in constrained optimization. SIAM J. Control Optim. **15**, 959–972 (1977)
29. Mohebi, H., Salemi, A.: Analysis of symmetric matrix valued functions. Numer. Funct. Anal. Optim. **28**, 691–715 (2007)
30. Mordukhovich, B.S., Nghia, T.T.A., Rockafellar, R.T.: Full stability in finite-dimensional optimization. Math. Oper. Res. **40**, 226–252 (2015)
31. Moreau, J.-J.: Proximité et dualité dans un espace hilbertien. Bull. Soc. Math. Fr. **93**, 1067–1070 (1965)
32. Nashed, M.Z.: Differentiability and related properties of nonlinear operators: some aspects of the role of differentials in nonlinear functional analysis. In: Rall, L.B. (ed.) Nonlinear Functional Analysis and Applications, pp. 103–309. Academic Press, New York (1971)
33. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. SIAM, Philadelphia (1970)
34. Qi, H.D., Yang, X.Q.: Semismoothness of spectral functions. SIAM J. Matrix Anal. Appl. **25**, 766–783 (2003)
35. Qi, L., Sun, J.: A nonsmooth version of Newton's method. Math. Program. **58**, 353–367 (1993)

36. Recht, B., Fazel, M., Parrilo, P.A.: Guaranteed minimum rank solutions to linear matrix equations via nuclear norm minimization. SIAM Rev. **52**, 471–501 (2010)
37. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
38. Sun, D.F.: The strong second order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications. Math. Oper. Res. **31**, 761–776 (2006)
39. Sun, D.F., Sun, J.: Semismooth matrix-valued functions. Math. Oper. Res. **27**, 150–169 (2002)
40. Sun, D.F., Sun, J.: Löwner's operator and spectral functions in Euclidean Jordan algebras. Math. Oper. Res. **33**, 421–445 (2008)
41. Todd, M.J.: Semidefinite optimization. Acta Numer. **10**, 515–560 (2001)
42. Toh, K.C.: GMRES vs. ideal GMRES. SIAM J. Matrix Anal. Appl. **18**, 30–36 (1997)
43. Toh, K.C., Trefethen, L.N.: The Chebyshev polynomials of a matrix. SIAM J. Matrix Anal. Appl. **20**, 400–419 (1998)
44. Wright, J., Ma, Y., Ganesh, A., Rao, S.: Robust principal component analysis: exact recovery of corrupted low-rank matrices via convex optimization. In: Bengio, Y., Schuurmans, D., Lafferty, J., Williams, C. (eds.), Advances in Neural Information Processing Systems **22** (2009)
45. Wu, B., Ding, C., Sun, D.F., Toh, K.C.: On the Moreau–Yosida regularization of the vector k-norm related functions. SIAM J. Optim. **24**, 766–794 (2014)
46. Yang, L.Q., Sun, D.F., Toh, K.C.: SDPNAL+: a majorized semismooth Newton-CG augmented Lagrangian method for semidefinite programming with nonnegative constraints. Math. Program. Comput. **7**, 331–366 (2015)
47. Yang, Z.: A study on nonsymmetric matrix-valued functions. Master's Thesis, National University of Singapore. http://www.math.nus.edu.sg/~matsundf/Main_YZ.pdf (2009)
48. Zhao, X.Y., Sun, D.F., Toh, K.C.: A Newton-CG augmented Lagrangian method for semidefinite programming. SIAM J. Optim. **20**, 1737–1765 (2010)