

# Commutation Principles for Optimization Problems Involving Strictly Schur-Convex Functions in Euclidean Jordan Algebras

Pedro G. Massey\*

Noelia B. Rios\*

David Sossa†

**Abstract.** In this work, we establish several commutation principles for optimizers of shifts of spectral functions in the context of Euclidean Jordan Algebras (EJAs). For instance, we show that under certain assumptions, if  $\bar{x}$  is a (local) optimizer of  $F(x - a)$  for  $x \in \Omega$ , where  $\Omega \subset \mathcal{V}$  is a spectral set of an EJA  $\mathcal{V}$ ,  $a \in \mathcal{V}$  and  $F : \mathcal{V} \rightarrow \mathbb{R}$  is a strictly Schur-convex spectral function, then  $a$  and  $\bar{x}$  operator commute. We make no further assumption on the smoothness of  $F$ ; instead, we take advantage of the smoothness (Lie structure) of the automorphism group of  $\mathcal{V}$  and make use of majorization techniques for the eigenvalues of elements in EJAs. Our approach allows us to deal with several problems considered in the literature, related to strictly convex spectral functions and strictly convex spectral norms. In particular, we use our commutation principles to analyze the problem of minimizing the condition number in EJAs.

*Mathematics Subject Classification:* 17C20, 15A27, 15A42, 90C26

*Keywords:* Euclidean Jordan algebra, commutation principle, Lidskii's inequality, strictly Schur-convex function, condition number.

## 1 Introduction

Commutation principles in Euclidean Jordan algebras (EJAs) consist of a class of optimization problems involving (weakly) spectral sets and functions, where it is ensured that each (local) optimizer operator commutes with some element related to the objective function. For instance, let  $\mathcal{V}$  be an EJA and consider the map

$$x \in \Omega \mapsto \Theta(x) + F(x), \tag{1}$$

---

\*Universidad Nacional de La Plata, Facultad de Cs. Exactas, CMaLP e Instituto Argentino de Matemática, CONICET, Argentina. (e-mails: massey@mate.unlp.edu.ar and nbrios@mate.unlp.edu.ar).

†Universidad de O'Higgins, Instituto de Ciencias de la Ingeniería, Av. Libertador Bernardo O'Higgins 611, Rancagua, Chile (e-mail: david.sossa@uoh.cl).

where  $\Omega \subseteq \mathcal{V}$ , and  $\Theta, F : \mathcal{V} \rightarrow \mathbb{R}$ . In [17], Ramírez, Seeger and Sossa have proved that when  $\Omega$  is a spectral set,  $\Theta$  is differentiable, and  $F$  is a spectral function (see definitions in the next section), if  $\bar{x}$  is a local optimizer (minimizer or maximizer) of (1), then  $\nabla\Theta(\bar{x})$  and  $\bar{x}$  operator commute. Gowda and Jeong [5] further developed this commutation principle by assuming that  $\Omega$  and  $F$  are weakly spectral (i.e., they are invariant under automorphisms, which is a weaker condition than spectrality). These commutation principles were successfully applied in some variational problems like obtaining a meta-formula for generalized subdifferential of spectral functions [12], and deducing bounds for the determinant of the sum of two elements in EJAs [19]. It has also been investigated strong commutation principles for this class of problems, which consist of commutation principles that provide strong operator commutation. For instance, Gowda [4] proved that under the assumption that  $\Omega$  and  $F$  are spectrally defined, and by setting  $\Theta(x) = \langle x, a \rangle$ , with  $a \in \mathcal{V}$ ,

$$\bar{x} \text{ is a global maximizer of (1)} \quad \Rightarrow \quad a \text{ and } \bar{x} \text{ strongly operator commute.} \quad (2)$$

This class of results allows us, for instance, to simplify the problem of minimizing (1) to a minimization problem formulated in  $\mathbb{R}^n$ , with  $n$  being the rank of  $\mathcal{V}$ , via the eigenvalue map.

In this paper, we study commutation principles for the problem of optimizing the map

$$x \in \Omega \mapsto F(x - a), \quad (3)$$

where  $a \in \mathcal{V}$ ,  $\Omega \subseteq \mathcal{V}$  and  $F : \mathcal{V} \rightarrow \mathbb{R}$  is a spectral function. Observe that, in general, the maps (3) and (1) are not related; so, we cannot use directly the commutation principle discussed above to characterize the optimizers of (3). Indeed, although  $F$  is spectral, the shifted spectral function  $F(\cdot - a)$  is no longer spectral. Moreover, we are not imposing a differentiability assumption over  $F$ . Some particular cases of (3) have been studied in the literature. For instance, Jeong and Sossa [9] have recently proved that when  $\Omega$  is weakly spectral and  $F$  is a strictly convex spectral function or a strictly convex spectral norm, if  $\bar{x}$  is a local optimizer of (3) then  $a$  and  $\bar{x}$  operator commute. Their proof relies on a commutation principle developed for (1) by using non-smooth analysis. Massey, Rios and Stojanoff [14] have also obtained a commutation principle for (3) in the context of Hermitian matrices, which are examples of (simple) EJAs. Their result states that in case  $\Omega$  is an orbit of some element of  $\mathcal{V}$  and  $F$  is a strictly convex spectral norm, if  $\bar{x}$  is a local minimizer of (3) then  $a$  and  $\bar{x}$  strongly operator commute. Their approach is based on majorization techniques and on the smooth manifold structure of unitary orbits of Hermitian matrices. Similar results have also been applied in some other optimization problems [1].

Inspired by the works of Massey, Rios and Stojanoff [14], and Jeong and Sossa [9], we aim to develop commutation principles and strong commutation principles for the problem of optimizing (3) with  $\Omega$  and  $F$  as general as possible. Our proof techniques are mainly inspired by those techniques

developed in [14]; that is, we will use majorization techniques in EJAs, see [8]. The spectral function  $F$  will be assumed to be strictly Schur-convex. Observe that this class of functions is quite broad and covers classes like spectral functions that are strictly convex, strictly quasi-convex, and strictly convex spectral norms. We also introduce a new strictly Schur-convex function which is used to deal with the problem of minimizing the condition number (see Section 4).

Our main results are the following commutation principles: Let  $a, b \in \mathcal{V}$ ,  $\Omega \subseteq \mathcal{V}$  and  $F : \mathcal{V} \rightarrow \mathbb{R}$  be a strictly Schur-convex spectral function;

- (i) Suppose that  $\Omega$  is a spectral set. If  $\bar{x}$  is a global minimizer (resp. global maximizer) of (3), then  $a$  and  $\bar{x}$  (resp.  $-a$  and  $\bar{x}$ ) strongly operator commute (see Theorem 3.1).
- (ii) Suppose that  $\Omega$  is a weakly spectral set. If  $\bar{x}$  is a local optimizer (minimizer or maximizer) of (3), then  $a$  and  $\bar{x}$  operator commute (see Theorem 3.4).
- (iii) Suppose that  $\mathcal{V}$  is simple, and set  $\Omega$  as the eigenvalue orbit of  $b$ . If  $\bar{x}$  is a local minimizer (resp. local maximizer) of (3), then  $a$  and  $\bar{x}$  (resp.  $-a$  and  $\bar{x}$ ) strongly operator commute. Moreover,  $\bar{x}$  is a global optimizer of (3) (see Theorem 3.8).

Regarding item (i) above, we remark that in this case, if we replace the hypothesis that  $\bar{x}$  is a global minimizer (resp. global maximizer) by the weaker hypothesis that  $\bar{x}$  is a local minimizer (resp. local maximizer) of (3) then  $a$  and  $\bar{x}$  (resp.  $-a$  and  $\bar{x}$ ) do not necessarily strongly operator commute (see Example 3.10). We also point out that if we replace the assumption that  $\Omega$  is a spectral set for the (weaker) assumption that  $\Omega$  is a weakly spectral set and assume that  $\bar{x}$  is a global minimizer (resp. global maximizer) of (3) then,  $a$  and  $\bar{x}$  (resp.  $-a$  and  $\bar{x}$ ) do not necessarily strongly operator commute (see Example 3.2 below).

Regarding item (ii), notice that weakly spectral sets are rather small domains for a spectral function in an arbitrary EJA (as opposed to spectral sets, which are larger domains). Hence, the hypothesis that  $\bar{x}$  is a local optimizer within weakly spectral sets is a rather weak assumption. In turn, to prove this result, we will need to make use of the differential manifold structure of the automorphism group of an EJA. Since spectral sets are weakly spectral, the result also applies to local optimizers within spectral sets (see Corollary 3.7).

Regarding item (iii), we point out that if we remove the hypothesis that  $\mathcal{V}$  is simple and assume that  $\bar{x}$  is a local minimizer (resp. local maximizer) of (3) then,  $a$  and  $\bar{x}$  (resp.  $-a$  and  $\bar{x}$ ) do not necessarily strongly operator commute (see Example 3.2 below). Also, notice that we have set  $\Omega$  to be a rather minimal set, and hence, the hypothesis that  $\bar{x}$  is a local minimizer in  $\Omega$  is quite weak.

As a consequence of the previous comments, notice that the results in items (i) – (iii) are sharp.

The organization of the paper is as follows: In Section 2, we provide a background of the concepts and tools from EJAs that we require for developing our results. In Section 2.1, we review

majorization concepts and provide examples of strictly Schur-convex spectral functions. Sections 3.1, 3.2, and 3.3 are mainly devoted to proving the respective statements (given above) (i), (ii), and (iii). In Section 4, we illustrate the applications of our commutation principles to the analysis of the problem of minimizing the condition number in EJA. Finally, Section 5 provides concluding remarks.

## 2 Preliminaries

### Euclidean Jordan Algebras

Let  $(\mathcal{V}, \langle \cdot, \cdot \rangle, \circ)$  be an *Euclidean Jordan algebra* (EJA) of rank  $n$  and unit element  $e$ . For any  $x, y \in \mathcal{V}$ ,  $\langle x, y \rangle$  denotes the inner product and  $x \circ y$  denotes the Jordan product. We refer the reader to the book of Faraut and Korányi [3] as a general reference for the theory of EJAs.

Here, we recall some definitions and properties of EJAs that will be used later. The Jordan product in  $\mathcal{V}$  satisfies the commutation law ( $x \circ y = y \circ x$ , for  $x, y \in \mathcal{V}$ ), but it is not associative in general; still, the Jordan product satisfies the Jordan identity:  $(x \circ y) \circ (x \circ x) = x \circ (y \circ (x \circ x))$ , for  $x, y \in \mathcal{V}$ . Recall that an EJA is said to be *simple* if it is not a direct product of nonzero EJAs. It is well known [3, Corollary IV.1.5, Theorem V.3.7] that any EJA can be decomposed into a direct product of simple EJAs and that there are only five simple ones up to isomorphism. One of them is the algebra  $\mathcal{S}^n$  of  $n \times n$  symmetric matrices with Jordan and inner products given by

$$X \circ Y = \frac{XY + YX}{2}, \quad \langle X, Y \rangle = \text{tr}(XY) \text{ (trace inner product)}.$$

An element  $c \in \mathcal{V}$  is an *idempotent* if  $c^2 = c$ . An idempotent  $c$  is *primitive* if it is nonzero and cannot be written as a sum of two nonzero idempotents. A *Jordan frame* is a collection  $\{c_1, \dots, c_n\}$  of primitive idempotents satisfying  $\sum_{i=1}^n c_i = e$ , and  $c_i \circ c_j = 0$  when  $i \neq j$ .

The spectral decomposition theorem asserts that any element  $x \in \mathcal{V}$  can be written as

$$x = \lambda_1(x)e_1 + \lambda_2(x)e_2 + \dots + \lambda_n(x)e_n,$$

where  $\lambda_1(x) \geq \lambda_2(x) \geq \dots \geq \lambda_n(x)$  are real numbers called the eigenvalues of  $x$ , and  $\{e_1, e_2, \dots, e_n\}$  is a Jordan frame. We denote  $\lambda(x) \in \mathbb{R}^n$  (eigenvalue map) to the vector whose components are the eigenvalues of  $x$  arranged in non-increasing order. The *trace* of  $x \in \mathcal{V}$  is defined as the sum of its eigenvalues. That is,  $\text{tr}(x) := \lambda_1(x) + \dots + \lambda_n(x)$ . The inner product of  $\mathcal{V}$  is scaled so that we can assume that  $\langle x, y \rangle = \text{tr}(x \circ y)$  which is the *trace inner product*. From now on,  $\|\cdot\|$  will denote the norm on  $\mathcal{V}$  induced by the trace inner product; that is,  $\|x\| = (\text{tr}(x \circ x))^{1/2}$ . When  $\mathcal{V} = \mathbb{R}^n$ ,  $\|\cdot\|$  coincides with the Euclidean norm (2-norm) on  $\mathbb{R}^n$ .

Let  $p$  be an idempotent element of  $\mathcal{V}$ . It is known that  $\mathcal{V}$  can be decomposed as the direct sum

$$\mathcal{V} = \mathcal{V}(p, 1) \oplus \mathcal{V}(p, 0) \oplus \mathcal{V}(p, 1/2), \quad (4)$$

where  $\mathcal{V}(p, \alpha) := \{x \in \mathcal{V} : x \circ p = \alpha x\}$ , for  $\alpha \in \{1, 0, 1/2\}$ . This decomposition is called the *Peirce decomposition* of  $\mathcal{V}$  with respect to  $p$ . It is also known that  $\mathcal{V}(p, 1)$  and  $\mathcal{V}(p, 0)$  are subalgebras of  $\mathcal{V}$ , and  $\mathcal{V}(p, 1/2) \neq \{0\}$  whenever  $\mathcal{V}$  is simple and  $p \notin \{0, e\}$ .

Two elements  $a$  and  $b$  in  $\mathcal{V}$  are said to *operator commute* if  $L_a L_b = L_b L_a$ , where  $L_a : \mathcal{V} \rightarrow \mathcal{V}$  is a linear map defined as  $L_a(x) = a \circ x$  for all  $x \in \mathcal{V}$ . The operator commutativity property in  $\mathcal{S}^n$  coincides with the usual matrix commutativity property. That is,  $A$  and  $B$  operator commute in  $\mathcal{S}^n$  if and only if  $AB = BA$ . It is known that  $a, b \in \mathcal{V}$  operator commute if and only if both  $a$  and  $b$  can be *simultaneously decomposed* by the same Jordan frame. That is, there exist a Jordan frame  $\{e_1, e_2, \dots, e_n\}$  and real numbers  $\alpha_i, \beta_i \in \mathbb{R}$  for  $i = 1, 2, \dots, n$  such that

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n, \quad b = \beta_1 e_1 + \beta_2 e_2 + \dots + \beta_n e_n, \quad (5)$$

where  $\alpha$  and  $\beta$  are some rearrangements of  $\lambda(a)$  and  $\lambda(b)$ , respectively. We say that  $a, b \in \mathcal{V}$  *strongly operator commute* if there exists a frame  $\{e_1, e_2, \dots, e_n\}$  for  $\mathcal{V}$  such that the representations in Eq. (5) hold, in such a way that  $\alpha = \lambda(a)$  and  $\beta = \lambda(b)$ . Then, one can see that  $a$  and  $b$  strongly operator commute if and only if  $\langle a, b \rangle = \langle \lambda(a), \lambda(b) \rangle$ .

A linear map  $X : \mathcal{V} \rightarrow \mathcal{V}$  is an (*Jordan algebra*) *automorphism* if  $X$  is bijective (invertible) and  $X(a \circ b) = Xa \circ Xb$ , for all  $a, b \in \mathcal{V}$ . The (Lie) group of all automorphisms of  $\mathcal{V}$  is denoted by  $\text{Aut}(\mathcal{V})$ . As an example, the automorphism group  $\text{Aut}(\mathcal{S}^n)$  of the EJA of  $n \times n$  symmetric matrices can be identified with  $\mathcal{O}(n)$  i.e. the orthogonal group in dimension  $n$ .

We say that a set  $Q \subseteq \mathbb{R}^n$  is *symmetric* if it is permutation invariant which means that if  $u \in Q$  then  $Pu \in Q$  for every permutation matrix  $P$  of order  $n$ . A set  $\Omega \subset \mathcal{V}$  is called *spectral set* if there exists a symmetric set  $Q \subseteq \mathbb{R}^n$  such that  $\Omega = \{x \in \mathcal{V} : \lambda(x) \in Q\}$ .  $\Omega$  is called *weakly spectral set* if it is invariant under automorphisms which means that if  $x \in \Omega$  and  $A \in \text{Aut}(\mathcal{V})$  then  $Ax \in \Omega$ . In particular, for  $b \in \mathcal{V}$ ,  $[b] := \{x \in \mathcal{V} : \lambda(x) = \lambda(b)\}$  is a spectral set and  $[b]_w := \{Xb : X \in \text{Aut}(\mathcal{V})\}$  is a weakly spectral set. It is known that any spectral set is weakly spectral, because of  $\lambda(Ax) = \lambda(x)$  for all  $A \in \text{Aut}(\mathcal{V})$  and  $x \in \mathcal{V}$ . Moreover, the concepts of spectral and weakly spectral coincide whenever  $\mathcal{V}$  is simple.

We say that a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is *symmetric* if it is permutation invariant which means that if  $u \in \mathbb{R}^n$  then  $f(u) = f(Pu)$  for every permutation matrix  $P$  of order  $n$ . A function  $F : \mathcal{V} \rightarrow \mathbb{R}$  is said to be *spectral function* if there exists a symmetric function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  such that  $F(x) = f(\lambda(x))$  for all  $x \in \mathcal{V}$ .

## 2.1 Majorization and Strictly Schur-Convex Functions

For  $u \in \mathbb{R}^n$ , let  $u^\downarrow \in \mathbb{R}^n$  denote the vector obtained from  $u$  by arranging its entries in non-increasing order. We denote  $(\mathbb{R}^n)^\downarrow := \{u \in \mathbb{R}^n : u = u^\downarrow\}$ . For  $u, v \in \mathbb{R}^n$ , one says that  $u$  is *majorized* by  $v$ , in symbol  $u \prec v$ , if

$$u_1^\downarrow + \cdots + u_k^\downarrow \leq v_1^\downarrow + \cdots + v_k^\downarrow \quad \text{for every } k = 1, \dots, n, \quad (6)$$

and  $u_1^\downarrow + \cdots + u_n^\downarrow = v_1^\downarrow + \cdots + v_n^\downarrow$ . If only the conditions in (6) hold, we say that  $u$  is *submajorized* by  $v$ , and write  $u \prec_w v$ . We say that  $u$  is *strictly majorized* by  $v$  if  $u \prec v$  and  $u^\downarrow \neq v^\downarrow$  ( $u$  is not a rearrangement of  $v$ ). For  $x, y \in \mathcal{V}$ , we say that  $x$  is majorized by  $y$  in  $\mathcal{V}$  and write  $x \prec y$  if  $\lambda(x) \prec \lambda(y)$  in  $\mathbb{R}^n$ . We say that  $x$  is *strictly majorized* by  $y$  in  $\mathcal{V}$  if  $\lambda(x)$  is strictly majorized by  $\lambda(y)$  in  $\mathbb{R}^n$ .

A function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is said to be *Schur-convex* if for all  $u, v \in \mathbb{R}^n$ ,  $u \prec v \Rightarrow f(u) \leq f(v)$ . If, in addition,  $f(u) < f(v)$  whenever  $u$  is strictly majorized by  $v$ , then  $f$  is said to be *strictly Schur-convex* (cf. [13]).

**Definition 2.1.** Let  $F : \mathcal{V} \rightarrow \mathbb{R}$  be a spectral function given by  $F(x) = f(\lambda(x))$ . We say that  $F$  is a Schur-convex (resp. strictly Schur-convex) spectral function if its associated symmetric function  $f$  is Schur-convex (resp. strictly Schur-convex).

Since the commutation principles that we developed in this work require considering strictly Schur-convex spectral functions, we list some examples of them below. Other examples can be found in [13, Chapter 3] and [20].

**Example 2.2.** Any strictly convex spectral function  $F$  is strictly Schur-convex. Indeed, if  $f$  is the associated symmetric function of  $F$ , in [9] it was proved that  $F$  is strictly convex if and only if so is  $f$ . Then, the result follows by recalling that any strictly convex symmetric function is strictly Schur-convex.

**Example 2.3.** We say that  $F : \mathcal{V} \rightarrow \mathbb{R}$  is a *strictly convex spectral norm* if it is a spectral function and a strictly convex norm in  $\mathcal{V}$ . Recall that  $F(\cdot)$  is a strictly convex norm in  $\mathcal{V}$  if it is a norm in  $\mathcal{V}$  satisfying  $F(x + y) < 2$  for any distinct  $x, y \in \mathcal{V}$  such that  $F(x) = F(y) = 1$ . For instance, the Schatten  $p$ -norm of  $x \in \mathcal{V}$  given by  $\|x\|_p := (|\lambda_1(x)|^p + \cdots + |\lambda_n(x)|^p)^{1/p}$  is a strictly convex norm for  $p > 1$ . In [9], it is proved that any strictly convex spectral norm  $F$  in  $\mathcal{V}$  is strictly Schur-convex. Indeed,  $F$  is a strictly convex spectral norm in  $\mathcal{V}$  if and only if its associated symmetric function  $f$  is a strictly convex symmetric norm in  $\mathbb{R}^n$ . Moreover, any strictly convex symmetric norm in  $\mathbb{R}^n$  is strictly Schur-convex.

**Example 2.4.** Recall that  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a *strictly quasi-convex* function if  $f(\alpha u + (1 - \alpha)v) < \max\{f(u), f(v)\}$  for all  $u \neq v$  and  $\alpha \in (0, 1)$ , see [2, Definition 2.2.2]. By using induction, one can

prove that  $f$  is strictly quasi-convex if and only if for all integer  $k \geq 2$ , and for all distinct vectors  $u_1, \dots, u_k$  in  $\mathbb{R}^n$ ,  $f(\sum_{i=1}^k \alpha_i u_i) < \max_{i=1, \dots, k} f(u_i)$  for all  $\alpha_i \in (0, 1)$  such that  $\sum_{i=1}^k \alpha_i = 1$ .

Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a strictly quasi-convex symmetric function. We claim that  $f$  is strictly Schur-convex. Indeed, let  $u, v \in \mathbb{R}^n$  be such that  $u \prec v$  and  $u^\downarrow \neq v^\downarrow$ . Since  $f$  is strictly quasi-convex, then it is quasi-convex. It is known that a quasi-convex symmetric function is Schur-convex, see [13], Chapter 3, Result C.3. Then,  $f(u) \leq f(v)$ . We will prove that  $f(u) < f(v)$  when  $f$  is strictly quasi-convex. In fact, since  $u$  is strictly majorized by  $v$ , then  $u$  is in the convex hull of  $\{Pv : P \text{ is a } n \times n \text{ permutation matrix}\}$  but it is not an extreme point of that set. Hence,  $u$  can be written as  $u = \sum_{i=1}^k \alpha_i P_i v$ , where  $k \geq 2$ ,  $\alpha_1, \dots, \alpha_k$  are some positive scalars such that  $\sum_{i=1}^k \alpha_i = 1$ , and  $P_1, \dots, P_k$  are some permutation matrices such that  $P_i v \neq P_j v$  whenever  $i \neq j$ . Thus,

$$f(u) = f\left(\sum_{i=1}^k \alpha_i P_i v\right) < \max_{i=1, \dots, k} f(P_i v) = f(v),$$

where the strict inequality is because of the strict quasi-convexity of  $f$ , and the right equality is because  $f$  is symmetric. The claim follows.

Therefore, a spectral function is strictly Schur-convex whenever its associated symmetric function is strictly quasi-convex.

**Example 2.5.** Let  $\mathcal{V}$  be an EJA of rank  $n$  and let  $\mathcal{V}_+ = \{b \in \mathcal{V} : \lambda_n(b) > 0\}$ . Given  $b \in \mathcal{V}_+$ , we can consider the condition number  $\mathfrak{c}(b) = \lambda_1(b)/\lambda_n(b) > 0$ , which is a measure of the spread of the eigenvalues of  $b$  (see [18] for some optimization problems related to the condition number in EJAs). Similarly, if  $c \in \mathcal{V}$  we can consider the spectral spread of  $c$ ,  $\text{Sprd}(c) = \lambda_1(c) - \lambda_n(c) \geq 0$ , which is also a measure of the spread of the eigenvalues of  $c$ . For the particular case when  $\mathcal{V} = \mathcal{S}^n$  is the EJA of real symmetric  $n \times n$  matrices,  $\mathfrak{c}(b)$  and  $\text{Sprd}(c)$  play a role in determining the stability properties of  $b \in \mathcal{S}_+^n$  and  $c \in \mathcal{S}^n$ . It is well known that  $\mathfrak{c}(\cdot) : \mathcal{V}_+ \rightarrow \mathbb{R}$  and  $\text{Sprd}(\cdot) : \mathcal{V} \rightarrow \mathbb{R}$  are Schur-convex functions (since  $u \prec v$  implies that  $\lambda_n(v) \leq \lambda_n(u)$  and  $\lambda_1(v) \geq \lambda_1(u)$ ) but not strictly Schur-convex. As an alternative to studying the condition number and the spectral spread, we consider the following strictly Schur-convex functions:

$$\kappa_{\|\cdot\|}(b) = \left\| (\lambda_i(b)/\lambda_{n-i+1}(b))_{i=1}^{\lfloor n/2 \rfloor} \right\| \quad \text{and} \quad \text{Sprd}_{\|\cdot\|}(c) = \left\| (\lambda_i(c) - \lambda_{n-i+1}(c))_{i=1}^{\lfloor n/2 \rfloor} \right\|,$$

where  $b \in \mathcal{V}_+$ ,  $c \in \mathcal{V}$ , and  $\|\cdot\|$  is the Euclidean norm in  $\mathbb{R}^{\lfloor n/2 \rfloor}$  ( $\lfloor n/2 \rfloor$  denotes the integer part of  $n/2$ ). The case  $\kappa_{\|\cdot\|}$  is analyzed in Section 4. The analysis for the case  $\text{Sprd}_{\|\cdot\|}$  is analogous. See [10] for the introduction of the vector-valued spread of selfadjoint matrices and [15, 16] for some of its basic properties and applications in  $\mathcal{S}^n$ .

One of the main tools used to develop our analysis is the *Lidskii's inequality*. As mentioned

in [11], Lidskii's inequality is one of the central tools for studying perturbation theory for the eigenvalues of symmetric matrices. For the general case of Euclidean Jordan algebras, the proof of Lidskii's inequality can be found in [8, Theorem 5.1.]; this inequality says that for  $a$  and  $b$  in  $\mathcal{V}$ , the following majorization is satisfied:

$$\lambda(a) - \lambda(b) \prec \lambda(a - b). \quad (7)$$

We will also use the *Ky Fan's inequality* which says that for  $a$  and  $b$  in  $\mathcal{V}$ :

$$\lambda(a + b) \prec \lambda(a) + \lambda(b). \quad (8)$$

It is not difficult to see that Ky Fan's inequality can be obtained from Lidskii's inequality.

Below, we provide a characterization of the strong operator commutation property by using Lidskii's inequality.

**Proposition 2.6.** *Let  $a, b \in \mathcal{V}$ . The following statements are equivalent:*

- (a)  $a$  and  $b$  strongly operator commute.
- (b)  $\lambda(a + b) = \lambda(a) + \lambda(b)$ .
- (c)  $(\lambda(a) - \lambda(b))^\downarrow = \lambda(a - b)$ .

*Proof.* The equivalence (a)  $\Leftrightarrow$  (b) is proved in [6, Proposition 2.6]. Let us prove (a)  $\Leftrightarrow$  (c). Suppose that  $a$  and  $b$  strongly operator commute. Then there exists a Jordan frame  $\{c_1, \dots, c_n\}$  such that  $a = \sum_{i=1}^n \lambda_i(a)c_i$  and  $b = \sum_{i=1}^n \lambda_i(b)c_i$ . Hence,  $a - b = \sum_{i=1}^n (\lambda_i(a) - \lambda_i(b))c_i$  which implies the desired result  $\lambda(a - b) = (\lambda(a) - \lambda(b))^\downarrow$ . Suppose now that (c) holds. Let  $x \in [a]$ . Then, by observing that  $\lambda(x) = \lambda(a)$ , and by using Lidskii's inequality we obtain

$$\lambda(a - b) = (\lambda(x) - \lambda(b))^\downarrow \prec \lambda(x - b). \quad (9)$$

Since the 2-norm function  $\|\cdot\|$  in  $\mathbb{R}^n$  is convex and symmetric, we have that it is Schur-convex. Then, from (9) we obtain  $\|\lambda(a - b)\| \leq \|\lambda(x - b)\|$ . Hence,  $\|a - b\| \leq \|x - b\|$  for all  $x \in [a]$  where the norm on  $\mathcal{V}$  is the one induced by the trace inner product. Hence,  $a$  is a global minimizer of  $\frac{1}{2}\|x - b\|^2$  for  $x \in [a]$ . Note that this is equivalent to saying that  $a$  is a global maximizer of  $\langle x, b \rangle - \frac{1}{2}\|x\|^2$  for  $x \in [a]$ . Thus, since  $-\frac{1}{2}\|\cdot\|^2$  is a spectral function, we can use the strong commutation principle (2) to deduce that  $a$  and  $b$  strongly operator commute.  $\square$

### 3 Main Results: Commutation Principles

In this section we state and prove our main results. We include some rather technical lemmas that are proved in an Appendix (see Section 6).

#### 3.1 On Global Optimizers and Strong Commutation Principles

**Theorem 3.1.** *Let  $F : \mathcal{V} \rightarrow \mathbb{R}$  be a strictly Schur-convex spectral function with  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  as its associated symmetric function, let  $\Omega \subseteq \mathcal{V}$  be a spectral set with  $Q \subseteq \mathbb{R}^n$  as its associated symmetric set, and let  $a \in \mathcal{V}$ . Then, the following statements are satisfied:*

(a) *If  $\bar{x}$  is a global minimizer of  $F(x - a)$  for  $x \in \Omega$ , then  $a$  and  $\bar{x}$  strongly operator commute. Moreover,*

$$\min_{x \in \Omega} F(x - a) = \min_{u \in Q} f(u - \lambda(a)) = f(\lambda(\bar{x}) - \lambda(a)). \quad (10)$$

(b) *If  $\bar{x}$  is a global maximizer of  $F(x - a)$  for  $x \in \Omega$ , then  $-a$  and  $\bar{x}$  strongly operator commute. Moreover,*

$$\max_{x \in \Omega} F(x - a) = \max_{u \in Q} f(u + \lambda(-a)) = f(\lambda(\bar{x}) + \lambda(-a)). \quad (11)$$

*Proof.* (a). Let  $\bar{x} \in \Omega$  be a global minimizer of  $F(x - a)$ , for  $x \in \Omega$ . Suppose that  $a$  has the spectral decomposition  $a = \sum_{i=1}^n \lambda_i(a) c_i$  for some Jordan frame  $\{c_1, \dots, c_n\}$ . Let us take  $\hat{x} := \sum_{i=1}^n \lambda_i(\bar{x}) c_i$ . Observe that  $\lambda(\hat{x}) = \lambda(\bar{x})$  by construction. Hence,  $\hat{x} \in [\bar{x}] \subseteq \Omega$  since this last set is spectral. Then,  $\hat{x} - a = \sum_{i=1}^n (\lambda_i(\bar{x}) - \lambda_i(a)) c_i$  which implies

$$\lambda(\hat{x} - a) = (\lambda(\bar{x}) - \lambda(a))^\downarrow \prec \lambda(\bar{x} - a), \quad (12)$$

where the majorization is because of Lidskii's inequality (7). Let  $f$  be the symmetric function associated with  $F$ . Since  $F$  is a strictly Schur-convex spectral function, then  $f$  is a strictly Schur-convex symmetric function. Then, from (12), we obtain

$$F(\hat{x} - a) = f(\lambda(\hat{x} - a)) \leq f(\lambda(\bar{x} - a)) = F(\bar{x} - a). \quad (13)$$

Recall that  $\bar{x}$  is a global minimizer of  $F(x - a)$  for  $x \in \Omega$ . Then, this fact and (13) imply  $f(\lambda(\hat{x} - a)) = f(\lambda(\bar{x} - a))$ . Hence, due to the strict Schur-convexity of  $f$  we deduce that

$$\lambda(\hat{x} - a) = \lambda(\bar{x} - a). \quad (14)$$

Now, by combining the equality in (12) with (14) we deduce that  $\lambda(\bar{x} - a) = (\lambda(\bar{x}) - \lambda(a))^\downarrow$ . Thus, from Proposition 2.6, we conclude that  $a$  and  $\bar{x}$  strongly operator commute.

To prove (10), observe that since  $a$  and  $\bar{x}$  strongly operator commute, there exists a Jordan frame  $\{e_1, \dots, e_n\}$  such that  $a = \sum_{i=1}^n \lambda_i(a)e_i$  and  $\bar{x} = \sum_{i=1}^n \lambda_i(\bar{x})e_i$ . Then,

$$\min_{x \in \Omega} F(x - a) = F(\bar{x} - a) = f(\lambda(\bar{x}) - \lambda(a)). \quad (15)$$

We claim that  $f(\lambda(\bar{x}) - \lambda(a)) \leq f(u - \lambda(a))$  for all  $u \in Q$ . Indeed, reasoning by contradiction, suppose that there exists  $\hat{u} \in Q$  such that  $f(\hat{u} - \lambda(a)) < f(\lambda(\bar{x}) - \lambda(a))$ . Then we can take  $\hat{x} = \sum_{i=1}^n \hat{u}_i e_i$ , which is in  $\Omega$  since  $\lambda(\hat{x}) = \hat{u}^\downarrow \in Q$ , obtaining  $F(\hat{x} - a) = f(\hat{u} - \lambda(a)) < f(\lambda(\bar{x}) - \lambda(a)) = F(\bar{x} - a)$ , contradicting that  $\bar{x}$  is a global solution of the left-hand side problem of (15). Hence,  $f(\lambda(\bar{x}) - \lambda(a)) = \min_{u \in Q} f(u - \lambda(a))$ , and by combining with (15) we arrive to (10).

(b). Let  $\bar{x} \in \Omega$  be a global maximizer of  $F(x - a)$  for  $x \in \Omega$ . We consider the representation of  $a$  as  $a = \sum_{i=1}^n -\lambda_i(-a)c_i$ , and we set  $\hat{x} := \sum_{i=1}^n \lambda_i(\bar{x})c_i$ . Then, from the Ky Fan inequality (8) we get  $\lambda(\bar{x} - a) \prec \lambda(\bar{x}) + \lambda(-a) = \lambda(\hat{x} - a)$ . Thus,  $F(\bar{x} - a) \leq F(\hat{x} - a)$ , and since  $\bar{x}$  is a global maximizer we have  $F(\bar{x} - a) = F(\hat{x} - a)$ . Therefore, by reasoning as in item (a), we deduce that  $\lambda(\bar{x} - a) = \lambda(\bar{x}) + \lambda(-a)$ , which implies that  $-a$  and  $\bar{x}$  strongly operator commute because of Proposition 2.6. The proof of (11) is analogous to the proof of (10) with the difference that in this case we have  $a = \sum_{i=1}^n -\lambda_i(-a)e_i$  and  $\bar{x} = \sum_{i=1}^n \lambda_i(\bar{x})e_i$ , for some Jordan frame  $\{e_1, \dots, e_n\}$ , because  $-a$  and  $\bar{x}$  strongly operator commute.  $\square$

It is natural to consider whether the above Theorem is still valid if  $\Omega$  is assumed to be weakly spectral. The following example shows that this is not the case.

**Example 3.2.** Assume that  $\mathcal{S}$  is a simple EJA of rank 2, and consider the nonsimple EJA given by  $\mathcal{V} = \mathcal{S} \times \mathcal{S}$ . For  $b \in \mathcal{V}$ , recall that  $[b]_w$  is weakly spectral (but not necessarily spectral since the algebra is not simple). Below we construct  $a, b, \bar{x} \in \mathcal{V}$  such that  $\bar{x} \in [b]_w$  is a global minimizer of  $F(x - a)$  for  $x \in [b]_w$  (for any continuous function  $F : \mathcal{V} \rightarrow \mathbb{R}$ ) and such that  $a$  and  $\bar{x}$  do not strongly operator commute.

Recall from [5, Theorem 5.1] that in this case, each  $\phi \in \text{Aut}(\mathcal{V})$  has the form  $\phi = (\phi_1, \phi_2) \circ \sigma$ , with  $\phi_j \in \text{Aut}(\mathcal{S})$ ,  $j = 1, 2$  and where  $\sigma$  denotes the identity of  $\mathcal{V}$  or the transposition  $\sigma((b_1, b_2)) = (b_2, b_1)$ , for  $(b_1, b_2) \in \mathcal{V}$ . As a consequence, we conclude that given  $b = (b_1, b_2) \in \mathcal{V}$

$$[b]_w = \{x = (x_1, x_2) \in \mathcal{V} : (\lambda(x_1), \lambda(x_2)) = (\lambda(b_1), \lambda(b_2)) \text{ or } (\lambda(x_1), \lambda(x_2)) = (\lambda(b_2), \lambda(b_1))\},$$

where  $\lambda(x_j) \in \mathbb{R}^2$  denotes the eigenvalue vector of  $x_j \in \mathcal{S}$ , for  $j = 1, 2$ . Hence, the structure of  $\mathcal{V}$  imposes certain restrictions on  $a = (a_1, a_2) \in \mathcal{V}$  so that this element can be an element of the orbit  $[b]_w$ , beyond merely spectral conditions on  $\lambda(a) \in \mathbb{R}^4$ . Indeed, consider  $b = (b_1, b_2) \in \mathcal{V}$  such that  $\lambda(b_1) = (4, 1)$  and  $\lambda(b_2) = (3, 2)$ . Let  $\{e_1, e_2\}$  be a Jordan frame for  $\mathcal{S}$  and consider  $a_1 = 4e_1 + 3e_2$  and  $a_2 = 2e_1 + 1e_2$ . Hence,  $\lambda(a_1) = (4, 3)$  and  $\lambda(a_2) = (2, 1)$  so that  $a = (a_1, a_2) \notin [b]_w$ , even when

$$\lambda(b) = \lambda(a) = (4, 3, 2, 1).$$

We now show that  $a$  and  $x$  do not strongly operator commute, for any  $x \in [b]_w$ . Indeed, if there is  $x \in [b]_w$  such that  $a$  and  $x$  strongly operator commute then there exists a Jordan frame  $\{f_1, \dots, f_4\}$  for  $\mathcal{V}$  such that  $x = \sum_{j=1}^4 \lambda_j(b) f_j$  and  $a = \sum_{j=1}^4 \lambda_j(a) f_j$ . Since all eigenvalues of  $a$  are simple (have multiplicity one) then we get that  $f_1 = (e_1, 0)$ ,  $f_2 = (e_2, 0)$ ,  $f_3 = (0, e_1)$  and  $f_4 = (0, e_2)$ . Hence, we now see that  $x_1 = 4e_1 + 3e_2 = a_1$  and  $x_2 = 2e_1 + 1e_2 = a_2$  (thus,  $x_j$  and  $a_j$  strongly operator commute in  $\mathcal{S}$ , for  $j = 1, 2$ ). Notice that this last fact contradicts that  $x \in [b]_w$ .

Let  $F : \mathcal{V} \rightarrow \mathbb{R}$  be continuous. Since  $[b]_w$  is compact, we conclude that there exists  $\bar{x} \in [b]_w$  which is a minimizer of  $F(x - a)$  for  $x \in [b]_w$ ; by the previous arguments, we get that  $a$  and  $\bar{x}$  do not strongly operator commute.

The next result shows the converse of Theorem 3.1 when  $\Omega$  is an eigenvalue orbit.

**Proposition 3.3.** *Let  $F : \mathcal{V} \rightarrow \mathbb{R}$  be a Schur-convex spectral function. Let  $a, b \in \mathcal{V}$ , and let  $\bar{x} \in [b]$ . Then, the following statements are satisfied:*

- (a) *If  $a$  and  $\bar{x}$  strongly operator commute, then  $\bar{x}$  is a global minimizer of  $F(x - a)$  for  $x \in [b]$ .*
- (b) *If  $-a$  and  $\bar{x}$  strongly operator commute, then  $\bar{x}$  is a global maximizer of  $F(x - a)$  for  $x \in [b]$ .*

*Proof.* (a). Suppose that  $a$  and  $\bar{x}$  strongly operator commute. Then  $\lambda(\bar{x} - a) = (\lambda(\bar{x}) - \lambda(a))^\downarrow$ . Hence, for all  $x \in [b]$  we get that  $\lambda(\bar{x} - a) = (\lambda(\bar{x}) - \lambda(a))^\downarrow = (\lambda(x) - \lambda(a))^\downarrow$ , where the second equality is because  $\lambda(x) = \lambda(b) = \lambda(\bar{x})$  since  $x$  and  $\bar{x}$  remain in the orbit  $[b]$ . Then, from Lidskii's inequality (7), we get

$$\lambda(\bar{x} - a) = (\lambda(x) - \lambda(a))^\downarrow \prec \lambda(x - a), \quad (16)$$

for all  $x \in [b]$ . By definition, the fact that  $F$  is a Schur-convex spectral function means that its associated symmetric function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is Schur-convex. Then, from (16) we obtain

$$F(\bar{x} - a) = f(\lambda(\bar{x} - a)) \leq f(\lambda(x - a)) = F(x - a),$$

for all  $x \in [b]$ . It means that  $\bar{x}$  is a global minimizer of  $F(x - a)$  for  $x \in [b]$ .

(b). The proof is analogous to the proof of the previous item, now using Ky Fan's inequality (8).  $\square$

### 3.2 A General Commutation Principle

**Theorem 3.4.** *Let  $\mathcal{V}$  be an EJA and let  $a \in \mathcal{V}$ . Let  $F : \mathcal{V} \rightarrow \mathbb{R}$  be a strictly Schur-convex spectral function, and let  $\Omega \subset \mathcal{V}$  be a weakly spectral set. Let  $\bar{x}$  be a local optimizer (minimizer or maximizer) of  $F(x - a)$  for  $x \in \Omega$ . Then  $a$  and  $\bar{x}$  operator commute.*

The proof of Theorem 3.4 is based on the following lemma that shows that if  $a$  and  $\bar{x}$  do not operator commute then  $\bar{x}$  cannot be a local optimizer of  $F(x - a)$  for  $x \in [b]_w$ . The proof of the lemma relies on the geometric structure of the automorphism group of an EJA and it is presented in an Appendix.

**Lemma 3.5.** *Let  $\mathcal{V}$  be an EJA with  $\text{rank}(\mathcal{V}) \geq 2$ , and let  $a, x \in \mathcal{V}$  be such that they do not operator commute. Then, for every  $\varepsilon > 0$ , the following statements are satisfied:*

(a) *there exists  $x_\varepsilon \in [x]$  such that  $\|x_\varepsilon - x\| \leq \varepsilon$ ,  $\lambda(x_\varepsilon - a) \prec \lambda(x - a)$  and  $\lambda(x_\varepsilon - a) \neq \lambda(x - a)$ .*

(b) *there exists  $y_\varepsilon \in [x]$  such that  $\|y_\varepsilon - x\| \leq \varepsilon$ ,  $\lambda(x - a) \prec \lambda(y_\varepsilon - a)$  and  $\lambda(y_\varepsilon - a) \neq \lambda(x - a)$ .*

*Proof.* See Section 6. □

We will also use the following elementary lemma about majorization:

**Lemma 3.6.** *If  $\gamma \prec \eta$  and  $\delta \prec \nu$  then  $(\gamma, \delta) \prec (\eta, \nu)$ , and if any of the first two majorization relations is strict, then the majorization relation  $(\gamma, \delta) \prec (\eta, \nu)$  is also strict.*

*Proof of Theorem 3.4.* If  $\text{rank}(\mathcal{V}) = 1$ , then  $V = \mathbb{R}c$ , where  $c$  is the unique primitive element in  $\mathcal{V}$ . Thus, in an EJA of rank one, any two elements (strongly) operator commute. Suppose that  $\text{rank}(\mathcal{V}) \geq 2$ . We consider the canonical decomposition  $\mathcal{V} = \prod_{i=1}^k \mathcal{V}_i$ , such that  $\mathcal{V}_i = \prod_{j=1}^{d_i} \mathcal{W}_i$ , where  $\mathcal{W}_1, \dots, \mathcal{W}_k$  are pairwise non-isomorphic simple EJAs of rank  $r_1, \dots, r_k$ , respectively. In this case, if  $c \in \mathcal{V}$  then  $c = (c_i)_{i=1}^k$  with  $c_i = (c_{i,j})_{j=1}^{d_i}$  such that  $c_{i,j} \in \mathcal{W}_i$ , for  $1 \leq j \leq d_i$ ,  $1 \leq i \leq k$ . In particular, we have that  $\lambda(c) = ((\lambda(c_{i,1}), \dots, \lambda(c_{i,d_i})))_{i=1}^k \downarrow$ .

Reasoning by contradiction, suppose that  $a$  and  $\bar{x}$  do not operator commute. Then, there exist  $1 \leq i_0 \leq k$  and  $1 \leq j_0 \leq d_{i_0}$  such that  $a_{i_0, j_0}$  and  $\bar{x}_{i_0, j_0}$  do not operator commute, which implies that  $\text{rank}(\mathcal{W}_{i_0}) \geq 2$ . From Lemma 3.5 we deduce that for every  $\varepsilon > 0$  there exists  $x_{i_0, j_0}(\varepsilon) \in [\bar{x}_{i_0, j_0}] = [\bar{x}_{i_0, j_0}]_w$ , where the (weakly) spectral orbit is considered with respect to the simple EJA  $\mathcal{W}_{i_0}$ , such that  $\|x_{i_0, j_0}(\varepsilon) - \bar{x}_{i_0, j_0}\| \leq \varepsilon$  and such that  $x_{i_0, j_0}(\varepsilon) - a_{i_0, j_0}$  is strictly majorized by  $\bar{x}_{i_0, j_0} - a_{i_0, j_0}$ ; that is,

$$\lambda(x_{i_0, j_0}(\varepsilon) - a_{i_0, j_0}) \prec \lambda(\bar{x}_{i_0, j_0} - a_{i_0, j_0}) \quad \text{and} \quad \lambda(x_{i_0, j_0}(\varepsilon) - a_{i_0, j_0}) \neq \lambda(\bar{x}_{i_0, j_0} - a_{i_0, j_0}).$$

Let  $x(\varepsilon)$  be obtained from  $\bar{x}$  by replacing the  $(i_0, j_0)$  component of  $\bar{x}$  by  $x_{i_0, j_0}(\varepsilon)$ , for  $\varepsilon > 0$ . Notice that by construction  $x(\varepsilon) \in [\bar{x}]_w \subset \Omega$  (since  $\Omega$  is weakly spectral) and  $\|x(\varepsilon) - \bar{x}\| \leq \varepsilon$ , for  $\varepsilon > 0$ . Furthermore, by using the Lemma 3.6, we get that

$$(\lambda(x_{i_0, 1}(\varepsilon) - a_{i_0, 1}), \dots, \lambda(x_{i_0, d_{i_0}}(\varepsilon) - a_{i_0, d_{i_0}})) \prec (\lambda(\bar{x}_{i_0, 1} - a_{i_0, 1}), \dots, \lambda(\bar{x}_{i_0, d_{i_0}} - a_{i_0, d_{i_0}}))$$

for  $\varepsilon > 0$ , where the majorization is strict. By using again the Lemma 3.6 and the construction of  $x(\varepsilon)$  we now see that  $\lambda(x(\varepsilon) - a) \prec \lambda(\bar{x} - a)$ , for  $\varepsilon > 0$ , with strict majorization.

Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be the symmetric function associated with  $F$ . We have that  $f$  is strictly Schur-convex because  $F$  is a strictly Schur-convex spectral function. By applying this fact to the strict majorization  $\lambda(x(\varepsilon) - a) \prec \lambda(\bar{x} - a)$  we get that  $F(x(\varepsilon) - a) < F(\bar{x} - a)$ . It means that  $\bar{x}$  is not a local minimizer of  $F(x - a)$  for  $x \in [\bar{x}]_w$ , which implies that  $\bar{x}$  is not a local minimizer of  $F(x - a)$  for  $x \in \Omega$ . We arrived at a contradiction.

Arguing as before, we can prove the analogous result for the case of a local maximizer.  $\square$

**Corollary 3.7.** *Let  $\mathcal{V}$  be an EJA and let  $a \in \mathcal{V}$ . Let  $F : \mathcal{V} \rightarrow \mathbb{R}$  be a strictly Schur-convex spectral function, and let  $\Omega \subset \mathcal{V}$  be a spectral set. Let  $\bar{x}$  be a local optimizer (minimizer or maximizer) of  $F(x - a)$  for  $x \in \Omega$ . Then  $a$  and  $\bar{x}$  operator commute.*

*Proof.* Since a spectral set  $\Omega \subset \mathcal{V}$  is a weakly spectral set, we now see that the result follows directly from Theorem 3.4.  $\square$

### 3.3 Strong Commutation Principles for Local Extrema in Simple EJAs

In this section, we present the following strong commutation principle for local minimizers in automorphism orbits of simple EJAs. We remark that, unless stated otherwise, all EJAs  $\mathcal{V}$  considered in this section will be simple; in particular,  $[b] = [b]_w$ , for  $b \in \mathcal{V}$ .

**Theorem 3.8.** *Let  $\mathcal{V}$  be a simple EJA, let  $F : \mathcal{V} \rightarrow \mathbb{R}$  be a strictly Schur-convex spectral function, and let  $a, b \in \mathcal{V}$ .*

(a) *If  $\bar{x}$  is a local minimizer of  $F(x - a)$  for  $x \in [b]$  then  $a$  and  $\bar{x}$  strongly operator commute.*

(b) *If  $\bar{x}$  is a local maximizer of  $F(x - a)$  for  $x \in [b]$  then  $-a$  and  $\bar{x}$  strongly operator commute.*

*Moreover, if  $\bar{x}$  is a local minimizer (resp. maximizer) of  $F(x - a)$  for  $x \in [b]$  then it is a global minimizer (resp. maximizer).*

The proof of Theorem 3.8 will require the following lemma.

**Lemma 3.9.** *Let  $\mathcal{V}$  be a simple Euclidean Jordan algebra of rank 2. Let  $a, x \in \mathcal{V}$  be such that they operator commute.*

(a) *Suppose that  $a$  and  $x$  do not strongly operator commute. Then, there exists a continuous curve  $\gamma : [0, \pi/2) \rightarrow [x]$  such that  $\gamma(0) = x$ , and  $\lambda(\gamma(\theta) - a) \prec \lambda(x - a)$  and  $\lambda(\gamma(\theta) - a) \neq \lambda(x - a)$ , for all  $\theta \in (0, \pi/2)$ .*

(b) *Suppose that  $-a$  and  $x$  do not strongly operator commute. Then, there exists a continuous curve  $\gamma : [0, \pi/2) \rightarrow [x]$  such that  $\gamma(0) = x$ , and  $\lambda(x - a) \prec \lambda(\gamma(\theta) - a)$  and  $\lambda(x - a) \neq \lambda(\gamma(\theta) - a)$ , for all  $\theta \in (0, \pi/2)$ .*

*Proof.* See Section 6. □

*Proof of Theorem 3.8.* We may assume that  $\text{rank}(\mathcal{V}) \geq 2$  since, as we saw in the proof of Theorem 3.4, any two elements strongly operator commute in an EJA of rank one.

(a). Observe that  $[b] = [b]_w$  because the algebra is simple. Since  $\bar{x}$  is a local minimizer of  $F(x - a)$  for  $x \in [b]$ , from the commutation principle given in Theorem 3.4 we deduce that  $a$  and  $\bar{x}$  operator commute. Then, there exists a Jordan frame  $\{e_1, \dots, e_n\}$  and real numbers  $\beta_1, \dots, \beta_n \in \mathbb{R}$  such that

$$a = \sum_{i=1}^n \lambda_i(a) e_i \quad \text{and} \quad \bar{x} = \sum_{i=1}^n \beta_i e_i. \quad (17)$$

Reasoning by contradiction, suppose that  $a$  and  $\bar{x}$  do not strongly operator commute. Then, there exist indices  $1 \leq j < k \leq n$  such that  $\lambda_j(a) > \lambda_k(a)$  and  $\beta_j < \beta_k$ <sup>1</sup>. Below we show that this last fact implies that  $\bar{x}$  is not a local minimizer of  $F(x - a)$  for  $x \in [b]$ .

Recall that every idempotent  $p \in \mathcal{V}$  induces the Peirce decomposition  $\mathcal{V} = \mathcal{V}(p, 1) \oplus \mathcal{V}(p, 0) \oplus \mathcal{V}(p, 1/2)$ , see (4). Let  $p := e_j + e_k$ , which turns out to be an idempotent in  $\mathcal{V}$ . Observe that  $\mathcal{V}(p, 1)$  is a simple Euclidean Jordan algebra of rank 2, and

$$\hat{a} := \lambda_j(a)e_j + \lambda_k(a)e_k \quad \text{and} \quad \hat{x} := \beta_j e_j + \beta_k e_k, \quad (18)$$

are elements of  $\mathcal{V}(p, 1)$ . Then,  $\hat{a}$  and  $\hat{x}$  operator commute in  $\mathcal{V}(p, 1)$  but not strongly because  $\lambda_j(a) > \lambda_k(a)$  and  $\beta_j < \beta_k$ . From Lemma 3.9, item (a), we deduce that there exists a continuous curve  $\gamma : [0, \pi/2) \rightarrow [\hat{x}]$  such that  $\gamma(0) = \hat{x}$ , and

$$\lambda(\gamma(\theta) - \hat{a}) \prec \lambda(\hat{x} - \hat{a}) \quad \text{and} \quad \lambda(\gamma(\theta) - \hat{a}) \neq \lambda(\hat{x} - \hat{a}), \quad (19)$$

for all  $\theta \in (0, \pi/2)$ , where  $\lambda(\cdot)$  is taken over  $\mathcal{V}(p, 1)$ . We now construct the continuous curve

$$\bar{x}(\theta) := \sum_{i=1, i \notin \{j, k\}}^n \beta_i e_i + \gamma(\theta) \in \mathcal{V}, \quad \text{for } \theta \in [0, \pi/2).$$

---

<sup>1</sup>In fact, the opposite statement asserts that for all indices  $1 \leq j < k \leq n$ ,  $\lambda_j(a) = \lambda_k(a)$  or  $\beta_j \geq \beta_k$ . That implies  $a$  and  $\bar{x}$  strongly operator commute. Indeed, let  $\mu_1 > \dots > \mu_p$  be the distinct eigenvalues of  $a$ ,  $I_\ell := \{i : \lambda_i(a) = \mu_\ell\}$ , and  $\mathbf{e}_\ell := \sum_{i \in I_\ell} e_i$ , for all  $\ell = 1, \dots, p$ . From (17) we have that  $a, \bar{x} \in \mathcal{V}(\mathbf{e}_1, 1) \oplus \dots \oplus \mathcal{V}(\mathbf{e}_p, 1)$ . Let  $a_\ell$  and  $\bar{x}_\ell$  denote the component of  $a$  and  $\bar{x}$  in  $\mathcal{V}(\mathbf{e}_\ell, 1)$ , respectively. Then,  $\lambda(a + \bar{x}) = (\lambda(a_1 + \bar{x}_1), \dots, \lambda(a_p + \bar{x}_p))^\downarrow$ . Since  $a_\ell = \mu_\ell \mathbf{e}_\ell$  and  $\mathbf{e}_\ell$  is the unit element of the EJA  $\mathcal{V}(\mathbf{e}_\ell, 1)$ , we have that  $a_\ell$  and  $\bar{x}_\ell$  strongly operator commute, which means  $\lambda(a_\ell + \bar{x}_\ell) = \lambda(a_\ell) + \lambda(\bar{x}_\ell)$  for all  $\ell = 1, \dots, p$ . Furthermore, for every  $1 \leq s < t \leq p$  and  $(j, k) \in I_s \times I_t$ , we have that  $\lambda_j(a) = \mu_s > \mu_t = \lambda_k(a)$ , which implies  $\beta_j \geq \beta_k$ , as it was pointed out at the beginning. Hence, for every  $1 \leq s < t \leq p$ ,  $\min_{j \in I_s} \beta_j \geq \max_{k \in I_t} \beta_k$ . Notice that  $\lambda_i(a_\ell + \bar{x}_\ell) = \mu_\ell + \beta_i$  for every  $\ell = 1, \dots, p$  and  $i \in I_\ell$ . Hence, by combining the above facts we deduce that  $\min_{i \in I_s} \lambda_i(a_s + \bar{x}_s) > \max_{i \in I_t} \lambda_i(a_t + \bar{x}_t)$  for every  $1 \leq s < t \leq p$ . We conclude that  $\lambda(a + \bar{x}) = (\lambda(a_1) + \lambda(\bar{x}_1), \dots, \lambda(a_p) + \lambda(\bar{x}_p)) = \lambda(a) + \lambda(\bar{x})$ .

By construction,  $\lambda(\bar{x}(\theta)) = \lambda(b)$ , and hence  $\bar{x}(\theta) \in [b]$ , for  $\theta \in [0, \pi/2)$ . Also, notice that  $\bar{x}(0) = \bar{x}$ . We claim that

$$\lambda(\bar{x}(\theta) - a) \prec \lambda(\bar{x} - a) \quad \text{and} \quad \lambda(\bar{x}(\theta) - a) \neq \lambda(\bar{x} - a), \quad \text{for } \theta \in (0, \pi/2). \quad (20)$$

Indeed, we have that

$$\bar{x}(\theta) - a = \overbrace{\sum_{i=1, i \notin \{j, k\}}^n (\beta_i - \lambda_i(a)) e_i}^{\in \mathcal{V}(p, 0)} + \overbrace{(\gamma(\theta) - \hat{a})}^{\in \mathcal{V}(p, 1)}.$$

Then,  $P_1 \lambda(\bar{x}(\theta) - a) = ((\beta_i - \lambda_i(a))_{i=1, i \notin \{j, k\}}^n, \lambda(\gamma(\theta) - \hat{a}))$ , for some permutation matrix  $P_1$ . Analogously,  $P_2 \lambda(\bar{x} - a) = ((\beta_i - \lambda_i(a))_{i=1, i \notin \{j, k\}}^n, \lambda(\hat{x} - \hat{a}))$ , for some permutation matrix  $P_2$ . Now, we use Lemma 3.6 to conclude the claim: indeed, since the first  $n - 2$  terms of  $P_1 \lambda(\bar{x}(\theta) - a)$  and  $P_2 \lambda(\bar{x} - a)$  are equal, and because of the strict majorization (19), we conclude that  $P_1 \lambda(\bar{x}(\theta) - a)$  is strictly majorized by  $P_2 \lambda(\bar{x} - a)$  and the claim (20) follows from these facts.

Finally, given  $\varepsilon > 0$  we choose  $\theta \in (0, \pi/2)$  such that  $\|\bar{x} - \bar{x}(\theta)\| \leq \varepsilon$ . Furthermore, as  $F$  is a strictly Schur-convex spectral function, its associated symmetric function  $f$  is also strictly Schur-convex. Since  $\lambda(\bar{x}(\theta) - a)$  is strictly majorized by  $\lambda(\bar{x} - a)$ , see (20), we deduce that  $F(\bar{x}(\theta) - a) < F(\bar{x} - a)$ . This last fact contradicts our assumption that  $\bar{x}$  is a local minimizer of  $F(x - a)$  for  $x \in [b]$ . Hence  $a$  and  $\bar{x}$  strongly operator commute. Furthermore, from Proposition 3.3, we conclude that  $\bar{x}$  is a global minimizer of  $F(x - a)$  for  $x \in [b]$ .

(b). The proof follows a similar argument to that in the proof of the previous item. Since  $\bar{x}$  is a local maximizer of  $F(x - a)$  for  $x \in [b]$  we deduce that  $a$  and  $\bar{x}$  operator commute. Then, there exist a Jordan frame  $\{e_1, \dots, e_n\}$  and real numbers  $\beta_1, \dots, \beta_n \in \mathbb{R}$  such that (17) holds. If we assume that  $a$  and  $\bar{x}$  do not strongly operator commute then there exist indexes  $1 \leq j < k \leq n$  such that  $\lambda_j(a) > \lambda_k(a)$  and  $\beta_j > \beta_k$ , we now define  $\hat{a}, \hat{x} \in \mathcal{V}(p, 1)$  as in (18), where  $p = e_j + e_k$ . From the previous facts, we see that  $-\hat{a}$  and  $\hat{x}$  do not strongly operator commute. Hence, from Lemma 3.9, item (b), there exists a continuous curve  $\gamma : [0, \pi/2) \rightarrow [\hat{x}]$  such that  $\gamma(0) = \hat{x}$ , and  $\lambda(\hat{x} - \hat{a}) \prec \lambda(\gamma(\theta) - \hat{a})$  and  $\lambda(\hat{x} - \hat{a}) \neq \lambda(\gamma(\theta) - \hat{a})$  for all  $\theta \in (0, \pi/2)$ , where  $\lambda(\cdot)$  is taken over  $\mathcal{V}(p, 1)$ . We now construct the continuous curve

$$\bar{x}(\theta) := \sum_{i=1, i \notin \{j, k\}}^n \beta_i e_i + \gamma(\theta) \in \mathcal{V}, \quad \text{for } \theta \in [0, \pi/2).$$

Arguing as in the first part of the proof we get that  $\lambda(\bar{x} - a) \prec \lambda(\bar{x}(\theta) - a)$  and  $\lambda(\bar{x} - a) \neq \lambda(\bar{x}(\theta) - a)$ , for  $\theta \in (0, \pi/2)$ . As before, the previous facts imply that  $\bar{x}$  is not a local maximizer of  $F(x - a)$  for

$x \in [b]$ , against our assumption.  $\square$

The following remark shows that Theorem 3.8 cannot be extended for nonsimple EJAs.

*Remark 3.10.* Let  $\mathcal{V}$  be a possibly not simple EJA, and let  $b \in \mathcal{V}$ . Then, there exists a finite family  $\{b_j\}_{j=1}^m$  of elements in  $\mathcal{V}$  such that:  $b_1 = b$ ,

1.  $\lambda(b_j) = \lambda(b)$ , for  $1 \leq j \leq m$ ;
2. For  $1 \leq \ell \neq j \leq m$ ,  $[b_\ell]_w \neq [b_j]_w$  and  $[b] = \bigcup_{j=1}^m [b_j]_w$ .
3. For  $1 \leq \ell \neq j \leq m$  we have that  $\text{dist}([b_j]_w, [b_\ell]_w) = \inf \{\|c - d\| : c \in [b_j]_w, d \in [b_\ell]_w\} > 0$ .

Let us give a proof of the above assertions. Let  $b \in \mathcal{V}$  with spectral decomposition  $b = \sum_{i=1}^n \lambda_i(b) e_i$ . Let  $\Sigma_n$  denote the group of permutations of  $\{1, \dots, n\}$ . For  $\sigma \in \Sigma_n$ , we set  $b_\sigma := \sum_{i=1}^n \lambda_{\sigma(i)}(b) e_i$ . We claim that

$$\bigcup_{\sigma \in \Sigma_n} [b_\sigma]_w = [b]. \quad (21)$$

In fact, since  $\lambda(b_\sigma) = \lambda(b)$  we have that  $[b_\sigma]_w \subseteq [b]$  for all  $\sigma \in \Sigma_n$ . Then, the inclusion  $\subseteq$  in (21) is satisfied. For the reverse inclusion of (21), we consider the canonical decomposition  $\mathcal{V} = \prod_{i=1}^k \mathcal{V}_i$ , such that  $\mathcal{V}_i = \prod_{j=1}^{d_i} \mathcal{W}_i$ , where  $\mathcal{W}_1, \dots, \mathcal{W}_k$  are pairwise non-isomorphic simple EJAs of rank  $r_1, \dots, r_k$ , respectively. Then, for  $c \in \mathcal{V}$ , we write its representation as  $c = ((c_{i,j})_{j=1}^{d_i})_{i=1}^k$ , where  $c_{i,j} \in \mathcal{W}_i$  for  $1 \leq j \leq d_i$ ,  $1 \leq i \leq k$ . Let  $x \in [b]$ . Since  $\lambda(x) = \lambda(b)$ , there exists  $\sigma \in \Sigma_n$  such that  $\lambda(x_{i,j}) = \lambda((b_\sigma)_{i,j})$ ; this means that  $x_{i,j} \in [(b_\sigma)_{i,j}]$ , for every  $1 \leq j \leq d_i$  and  $1 \leq i \leq k$ . Since  $\mathcal{W}_i$  is simple, we have that  $[(b_\sigma)_{i,j}] = [(b_\sigma)_{i,j}]_w$ . Then, there exists  $X_{i,j} \in \text{Aut}(\mathcal{W}_i)$  such that  $X_{i,j} x_{i,j} = (b_\sigma)_{i,j}$ , for  $1 \leq j \leq d_i$ ,  $1 \leq i \leq k$ . Thus, since we can regard  $\prod_{i=1}^k \prod_{j=1}^{d_i} \text{Aut}(\mathcal{W}_i)$  sitting inside  $\text{Aut}(\mathcal{V})$  in a straightforward way, we see that  $X := \prod_{i=1}^k \prod_{j=1}^{d_i} X_{i,j} \in \text{Aut}(\mathcal{V})$ , and  $Xx = b_\sigma$ . This proves that  $x \in [b_\sigma]_w$  for some  $\sigma \in \Sigma_n$ , and the inclusion  $\supseteq$  in (21) is satisfied. Therefore, since  $\Sigma_n$  is finite, we can choose  $b_i = b_{\sigma_i}$  for convenient  $\sigma_i \in \Sigma_n$ , for  $1 \leq i \leq m$  and for some  $m \in \mathbb{N}$ , such that  $\{b_i\}_{i=1}^m$  satisfies items 1. and 2. above.

On the other hand, notice that for each  $c \in \mathcal{V}$ , the orbit  $[c]_w$  is a compact set (being the image of a continuous function with compact domain). If we choose  $1 \leq j \neq \ell \leq m$  then the fact that  $[b_j]_w \neq [b_\ell]_w$  implies that  $[b_j]_w \cap [b_\ell]_w = \emptyset$ ; hence,  $\text{dist}([b_j]_w, [b_\ell]_w) > 0$  and item 3. is also satisfied.

Consider now the notation of Example 3.2. Thus, we let  $\mathcal{V} = \mathcal{S} \times \mathcal{S}$ , where  $\mathcal{S}$  is a simple EJA of rank two. Recall that in this context we can construct  $a, b \in \mathcal{V}$  such that  $x$  and  $a$  do not strongly operator commute, for any  $x \in [b]_w$ . Moreover, notice that in that example,  $a \in [b]$ , since  $\lambda(a) = \lambda(b)$ . Given a strictly convex spectral function  $F : \mathcal{V} \rightarrow \mathbb{R}$ , let  $\bar{x} \in [b]_w$  be a global (and hence a local) minimizer of  $F(x - a)$  for  $x \in [b]_w$ . We now claim that  $\bar{x} \in [b]$  is a local minimizer of  $F(x - a)$  for  $x \in [b]$  (i.e., is a local minimizer for the larger domain  $[b]$ ). In fact, by item 3. above

we see that  $[b]_w = [b_1]_w$  is separated from  $\bigcup_{2 \leq j \leq m} [b_j]_w$ . That is, we can find a sufficiently small neighborhood of  $\bar{x}$  in  $[b]$  that is actually contained in  $[b_1]_w$ . This last fact implies the previous claim. Of course,  $\bar{x}$  is not a global minimizer of  $F$  in  $[b]$ , since  $a \in [b]$  and  $\lambda(a - a) = 0_n \prec \lambda(x - b)$  and  $0_n \neq \lambda(x - b)$ , for every  $x \in [b]_w$ . Hence,  $F(0) = F(a - a) < F(x - a)$ , for every  $x \in [b]_w$  (these facts follow from the majorization relation  $0_n \prec \alpha$ , for  $\alpha \in \mathbb{R}^n$  such that  $\sum_{i=1}^n \alpha_i = 0$ ). Finally, notice that by construction  $a$  and  $\bar{x} \in [b]_w$  do not strongly operator commute.

*Remark 3.11.* For simplicity of presentation, Theorems 3.1, 3.4, and 3.8 require  $F(\cdot)$  to be strictly Schur-convex on the entire space  $\mathcal{V}$ . However, those theorems are still valid if  $F(\cdot)$  is strictly Schur-convex in some spectral set  $K \subseteq \mathcal{V}$  such that  $\Omega - a \subseteq K$  (with  $\Omega = [b]$  for Theorem 3.8).

## 4 An application: Minimizing the Condition Vector in EJAs

In this section, we assume that  $\mathcal{V}$  is an EJA of rank  $n \geq 2$ . Recall that  $\lambda_i(x)$  are the eigenvalues of  $x \in \mathcal{V}$ , arranged in non-increasing order:  $\lambda_1(x) \geq \dots \geq \lambda_n(x)$ . Let  $\mathcal{V}_+$  denote the interior of the cone of square elements of  $\mathcal{V}$  (symmetric cone of  $\mathcal{V}$ ). It can be characterized as  $\mathcal{V}_+ = \{x \in \mathcal{V} : \lambda_n(x) > 0\}$ .

The condition number of  $x \in \mathcal{V}_+$  is defined as  $\mathfrak{c}(x) = \frac{\lambda_1(x)}{\lambda_n(x)}$ . Given  $a \in \mathcal{V}_+$ , there has been some interest in studying the problem of perturbing  $a$  in such a way as to diminish its condition number as much as possible, see [7, 18]. This problem can be formulated

$$\min_{x \in \Omega} \mathfrak{c}(x + a), \quad (22)$$

where  $\Omega$  is some set of admissible perturbations. It is known that  $\mathfrak{c}(\cdot)$  is Schur-convex on  $\mathcal{V}_+$ , but not necessarily strictly Schur-convex. Thus, we cannot apply our results directly to study this problem since we require strict Schur-convexity.

To overcome the above drawback, we consider a vector containing not only the condition number but also the ratio between the  $i$ th eigenvalue and the  $(n - i + 1)$ th eigenvalue, for  $i = 1, \dots, \lfloor n/2 \rfloor$ . This is specified in the next definition.

**Definition 4.1.** *Let  $x \in \mathcal{V}_+$ , and let*

$$\kappa_i(x) := \frac{\lambda_i(x)}{\lambda_{n-i+1}(x)} \quad \text{for } i = 1, \dots, \lfloor n/2 \rfloor.$$

*The condition vector of  $x$  is defined as the vector  $\kappa(x) = (\kappa_1(x), \dots, \kappa_{\lfloor n/2 \rfloor}(x)) \in \mathbb{R}^{\lfloor n/2 \rfloor}$ .*

It is straightforward to see that the entries of  $\kappa(x)$  are arranged in non-increasing order and that  $\kappa_1(x) = \mathfrak{c}(x)$ . Because of those facts, it can be proved that  $\lfloor n/2 \rfloor^{-1/2} \|\kappa(x)\| \leq \mathfrak{c}(x) \leq \|\kappa(x)\|$ , for every  $x \in \mathcal{V}_+$ , where  $\|\kappa(x)\|$  denotes the Euclidean norm of  $\kappa(x)$ . This shows that the behavior of

$\|\kappa(x)\|$  is comparable to the behavior of  $\mathbf{c}(x)$ . That is, if  $\|\kappa(x)\|$  is small or large, then  $\mathbf{c}(x)$  is small or large, respectively. Thus, we can analyze the problem (22) by solving the problem

$$\min_{x \in \Omega} \|\kappa(x + a)\|. \quad (23)$$

Our goal is to describe some fundamental properties of the solutions of (23) by relying on the tools we developed in the previous sections. Under some natural hypotheses, our approach also allows us to obtain structural information about the local solutions of (23). For  $x \in \mathcal{V}_+$ , let us denote  $\kappa_{\|\cdot\|}(x) := \|\kappa(x)\|$ . The next proposition shows that  $\kappa_{\|\cdot\|}$  is a strictly Schur-convex spectral function on  $\mathcal{V}_+$ . Let  $\mathbb{R}_+^n$  denote the interior of the positive orthant of  $\mathbb{R}^n$ .

**Proposition 4.2.** *The following statements are satisfied.*

- (a)  $\kappa_{\|\cdot\|} : \mathcal{V}_+ \rightarrow \mathbb{R}$  is a spectral function with associated symmetric function  $\phi_{\|\cdot\|} : \mathbb{R}_+^n \rightarrow \mathbb{R}$  given by  $\phi_{\|\cdot\|}(u) = \|\phi(u)\|$ , with

$$\phi(u) = \left( \frac{u_1^\downarrow}{u_n^\downarrow}, \frac{u_2^\downarrow}{u_{n-1}^\downarrow}, \dots, \frac{u_{\lfloor n/2 \rfloor}^\downarrow}{u_{n-\lfloor n/2 \rfloor+1}^\downarrow} \right) \in \mathbb{R}_+^{\lfloor n/2 \rfloor}.$$

- (b)  $\kappa_{\|\cdot\|}$  is a strictly Schur-convex spectral function on  $\mathcal{V}_+$ .

*Proof.* (a) It is clear from the definition of  $\kappa_{\|\cdot\|}$ .

(b) We first show that  $\phi_{\|\cdot\|}$  is Schur-convex on  $\mathbb{R}_+^n$ . Indeed, let  $u, v \in \mathbb{R}_+^n$  be such that  $u \prec v$ . Without loss of generality, we assume that  $u = u^\downarrow$  and  $v = v^\downarrow$ . Hence, by the relations between majorization and convex functions (see [13, Theorem 5.A.1]), we get that

$$(\mathbb{R}_+^n)^\downarrow \ni (u_n^{-1}, \dots, u_1^{-1}) \prec_w (v_n^{-1}, \dots, v_1^{-1}) \in (\mathbb{R}_+^n)^\downarrow.$$

As a consequence, we get the following relations for the entry-wise product vectors (see [13, Theorem 5.A.4.f])  $(\mathbb{R}_+^n)^\downarrow \ni (u_i/u_{n-i+1})_{i=1}^n \prec_w (v_i/v_{n-i+1})_{i=1}^n \in (\mathbb{R}_+^n)^\downarrow$ . In particular,

$$(\mathbb{R}_+^{\lfloor n/2 \rfloor})^\downarrow \ni (u_i/u_{n-i+1})_{i=1}^{\lfloor n/2 \rfloor} \prec_w (v_i/v_{n-i+1})_{i=1}^{\lfloor n/2 \rfloor} \in (\mathbb{R}_+^{\lfloor n/2 \rfloor})^\downarrow, \quad (24)$$

which implies that  $\phi_{\|\cdot\|}(u) \leq \phi_{\|\cdot\|}(v)$ .

We now show that  $\phi_{\|\cdot\|}$  is strictly Schur-convex on  $\mathbb{R}_+^n$ . As before, we consider  $u, v \in \mathbb{R}_+^n$  such that  $u \prec v$ , and assume that  $u = u^\downarrow$  and  $v = v^\downarrow$ . To prove that  $\phi_{\|\cdot\|}$  is strictly Schur-convex, it is enough to show that  $u = v$  whenever  $\phi_{\|\cdot\|}(u) = \phi_{\|\cdot\|}(v)$ . Suppose that  $\phi_{\|\cdot\|}(u) = \phi_{\|\cdot\|}(v)$ . Notice that  $\|\cdot\|$  is a strictly Schur-convex function, and it is a strictly increasing function on  $\mathbb{R}_+^{\lfloor n/2 \rfloor}$  (for the component-wise ordering in  $\mathbb{R}_+^{\lfloor n/2 \rfloor}$ ). Now, from [13, Theorem 3.A.8.a] we deduce that:

For a real-valued function  $g$  that is strictly increasing and strictly Schur-convex on  $\mathbb{R}_+^{\lfloor n/2 \rfloor}$ , if

$$a \prec_w b \text{ on } \mathbb{R}_+^{\lfloor n/2 \rfloor} \text{ and } g(a) = g(b), \text{ then } a^\downarrow = b^\downarrow.$$

By applying the above result to the function  $\|\cdot\|$ , and by recalling the submajorization (24) and the assumption  $\phi_{\|\cdot\|}(u) = \phi_{\|\cdot\|}(v)$ , we deduce that  $(u_i/u_{n-i+1})_{i=1}^{\lfloor n/2 \rfloor} = (v_i/v_{n-i+1})_{i=1}^{\lfloor n/2 \rfloor}$ . In particular, we get that  $u_1/u_n = v_1/v_n$ ; now, using that  $u \prec v$  implies that  $v_n \leq u_n$  and  $v_1 \geq u_1$ , we see that  $u_i = v_i$  for  $i \in \{1, n\}$ . We can continue in this way and show that  $u_i = v_i$ , for  $i \in \{j, n-j+1\}$  and  $1 \leq j \leq \lfloor n/2 \rfloor$ . Thus, we get  $u = v$  when  $n$  is even. When  $n$  is odd, we combine the previous fact with the normalization condition  $\text{tr}(u) = \text{tr}(v)$ , which comes from  $u \prec v$ , to deduce  $u = v$ .  $\square$

Finally, the following theorem characterizes the solutions of (23) and, in some cases, reduces (23) to a minimization problem formulated in  $\mathbb{R}^n$ . The proof of each item follows by a direct application of Theorems 3.4, 3.1 and 3.8, respectively (see also Remark 3.11).

**Theorem 4.3.** *Let  $a \in \mathcal{V}_+$  and let  $\Omega \subseteq \mathcal{V}$  be such that  $\Omega + a \subseteq \mathcal{V}_+$ . Consider the problem (23). Then, the following statements are satisfied:*

- (a) *Suppose that  $\Omega$  is a weakly spectral set. If  $\bar{x}$  is a local solution of (23), then  $a$  and  $\bar{x}$  operator commute.*
- (b) *Suppose that  $\Omega$  is a spectral set with  $Q \subseteq \mathbb{R}^n$  as its associated symmetric set. If  $\bar{x}$  is a solution of (23), then  $-a$  and  $\bar{x}$  strongly operator commute. Furthermore,*

$$\min_{x \in \Omega} \|\kappa(x + a)\| = \min_{u \in Q} \|\phi(u - \lambda(-a))\|.$$

- (c) *Let  $b \in \mathcal{V}$  be such that  $[b] + a \subseteq \mathcal{V}_+$ . Suppose that  $\mathcal{V}$  is simple. If  $\bar{x}$  is a local solution of*

$$\min_{x \in [b]} \|\kappa(x + a)\|, \tag{25}$$

*then  $-a$  and  $\bar{x}$  strongly operator commute. Furthermore,  $\bar{x}$  is a global solution of (25), and*

$$\min_{x \in [b]} \|\kappa(x + a)\| = \|\phi(\lambda(b) - \lambda(-a))\|.$$

$\square$

## 5 Conclusions

In this paper, we established several commutation principles, in the setting of Euclidean Jordan algebras, that describe the global and local extrema of functions of the form  $F(x - a)$  for  $x \in \Omega$ ,

provided that  $F$  is a strictly Schur-convex spectral function, and  $\Omega$  is a spectral or a weakly spectral set. Essentially, our results ensure that a local extremum operator commutes with  $a$ . Also, strongly operator commutation is obtained for a global extremum under the assumption that  $\Omega$  is spectral. Furthermore, we prove that when the algebra is simple and  $\Omega$  is the orbit of some element, any local extremum should be a global extremum since it satisfies a strongly operator commutation with  $a$ . As an application, we study the problem of minimizing the condition number of an element affected by a perturbation. This is done by defining a new strictly Schur-convex function that measures the spread of the eigenvalues of an element.

Let us point out that there are at least two possible directions for future research: on the one hand, our results allow us to reduce optimization problems involving an objective function of the form  $F(x - a)$ , for  $x \in \Omega \subset \mathcal{V}$ , to the case when  $a$  and  $x$  (strongly) operator commute. Still, some further work is needed to obtain a complete solution to the initial optimization problem. For example: for a weakly spectral set  $\Omega \subset \mathcal{V}_+$  and for  $a \in \mathcal{V}_+$ , we would like to obtain a complete description of the optimizers  $x_0 \in \Omega$  such that  $\min_{x \in \Omega} \kappa(x + a) = \kappa(x_0 + a)$ , depending on the set  $\Omega$ , or on the set  $\{\lambda(x) : x \in \Omega\}$  in case  $\Omega$  is an spectral set (here  $\kappa(x)$  denotes the condition vector of  $x \in \mathcal{V}_+$ , see Definition 4.1). On the other hand, it would be interesting to adapt our approach and techniques to tackle some related optimization problems e.g., to obtain operator commutation relations between  $x_0 \in \Omega$  and  $a$ , for those  $x_0$  such that  $\min_{x \in \Omega} F(x \circ a) = F(x_0 \circ a)$ , where  $a$  and  $\Omega$  are as before, and  $F$  is a strictly Schur-convex spectral function. We will consider these and other related problems elsewhere.

**Acknowledgments** This work was partially supported by FONDECYT (Chile) through grant 11220268, MATH-AMSUD through AMSUD230018 MORA-DataS project, CONICET (Argentina) through PIP 11220200103209CO, and Universidad Nacional de La Plata (Argentina) through UNLP 11X974. We would like to thank to the editor and the reviewers for their valuable comments and suggestions.

## 6 Appendix: Several Technical Proofs

We begin by considering some preliminary results needed for the proof of Lemma 6.3. These are results of a geometric nature, mostly.

Let  $a, b \in \mathcal{V}$ . Recall that the property that  $a$  and  $b$  operator commute is equivalent to the property that  $a$  and  $b$  can be simultaneously (spectrally) decomposed; that is, there exists a Jordan frame  $\{c_1, \dots, c_n\}$  such that  $a, b \in \mathcal{V}(c_1, 1) \oplus \dots \oplus \mathcal{V}(c_n, 1)$ . If  $a$  and  $b$  do not operator commute, in some cases, they can still be simultaneously decomposed but in blocks; that is, there exist non-zero idempotents  $\{p_1, \dots, p_k\}$  (for some  $2 \leq k \leq n$ ) such that  $p_i \circ p_j = 0$  if  $i \neq j$ ,  $p_1 + \dots + p_k = e$

(which means that  $\{p_1, \dots, p_k\}$  is a complete system of orthogonal idempotents) and such that  $a, b \in \mathcal{V}(p_1, 1) \oplus \dots \oplus \mathcal{V}(p_k, 1)$ . If this is not the case, we say that  $a$  and  $b$  are totally non-commuting. This concept will be key in our forthcoming lemmas. Below, we give a precise definition.

**Definition 6.1.** *Let  $\mathcal{V}$  be an EJA and let  $a, b \in \mathcal{V}$ . We say that  $a$  and  $b$  are simultaneously block decomposable if there exists an idempotent  $p \notin \{0, e\}$  such that  $a, b \in \mathcal{V}(p, 1) \oplus \mathcal{V}(p, 0)$ . We say that  $a$  and  $b$  are totally (operator) non-commuting if they are not simultaneously block decomposable.*

We also need to recall some tools from differential geometry. Let  $\mathcal{N}$  and  $\mathcal{M}$  be smooth manifolds embedded into finite-dimensional real inner product spaces  $\mathcal{U}$  and  $\mathcal{W}$ , respectively. For  $u \in \mathcal{N}$ , we denote by  $T_u(\mathcal{N})$  the tangent space of  $\mathcal{N}$  at  $u$ . A smooth map  $\Psi : \mathcal{N} \rightarrow \mathcal{M}$  is a *submersion* at  $u \in \mathcal{N}$  if its differential  $D\Psi_u : T_u(\mathcal{N}) \rightarrow T_{\Psi(u)}(\mathcal{M})$  is a surjective linear map.

It is known that  $\text{Aut}(\mathcal{V})$  has a smooth manifold (Lie group) structure whose tangent space at the identity (Lie algebra) is  $\text{Der}(\mathcal{V})$ . Recently, in [9, Proposition 2], it has been proved that the operator commutativity property can be characterized as follows:

$$a \text{ and } b \text{ operator commute} \quad \Leftrightarrow \quad \langle Da, b \rangle = 0 \text{ for all } D \in \text{Der}(\mathcal{V}). \quad (26)$$

In what follows we consider  $\mathcal{V}_t := \{x \in \mathcal{V} : \text{tr}(x) = t\}$ , for some  $t \in \mathbb{R}$ . Notice that  $\mathcal{V}$  is a translation of a hyperplane in  $\mathcal{V}$  and has a natural sub-manifold structure; moreover, we have that  $T_a(\mathcal{V}_t) = \mathcal{V}_0 = \{x \in \mathcal{V} : \text{tr}(x) = 0\}$ , for every  $a \in \mathcal{V}_t$ .

**Lemma 6.2.** *Let  $\mathcal{V}$  be an EJA. Fix  $a, b \in \mathcal{V}$  and set  $t = \text{tr}(a - b)$ . Let  $\Psi = \Psi_{(a,b)} : \text{Aut}(\mathcal{V}) \times \text{Aut}(\mathcal{V}) \rightarrow \mathcal{V}_t$  be given by  $\Psi(X, Y) = Xa - Yb$ . Then  $\Psi$  is a submersion at  $(I, I)$  if and only if  $a$  and  $b$  are totally non-commuting.*

*Proof.* Observe that  $T_{(I,I)}(\text{Aut}(\mathcal{V}) \times \text{Aut}(\mathcal{V})) = \text{Der}(\mathcal{V}) \times \text{Der}(\mathcal{V})$  and  $T_{\Psi(I,I)}(\mathcal{V}_t) = T_{a-b}(\mathcal{V}_t) = \mathcal{V}_0$ . Then, the differential map of  $\Psi$  at  $(I, I)$  is given by  $D\Psi : \text{Der}(\mathcal{V}) \times \text{Der}(\mathcal{V}) \rightarrow \mathcal{V}_0$ ,  $D\Psi(D, E) = Da - Eb$ . By definition,  $\Psi$  is a submersion at  $(I, I)$  if and only if the differential map  $D\Psi$  is surjective. Hence,  $\Psi$  is not a submersion if and only if there exists  $f \in \mathcal{V}_0$ ,  $f \neq 0$  such that  $f$  is orthogonal to the range of  $D\Psi$ . In turn, this last condition is equivalent to the fact that

$$\langle Da, f \rangle = \langle Db, f \rangle = 0 \quad \text{for every } D \in \text{Der}(\mathcal{V}), \quad (27)$$

which means, because of (26), that  $f$  operator commutes with  $a$  and  $b$  simultaneously. Let  $\gamma_1, \dots, \gamma_k$  be the distinct eigenvalues of  $f$ . We know from the first version of the spectral theorem [3, Theorem III.1.1] that there exists a unique complete system of orthogonal idempotents  $c_1, \dots, c_k$  such that

$$f = \sum_{i=1}^k \gamma_i c_i.$$

Notice that since  $f \neq 0$  and  $\text{tr}(f) = 0$  then  $f$  should have at least two different eigenvalues which implies that  $c_i \neq e$  for all  $i$ . Thus, since  $f$  operator commutes with  $a$  and  $b$ , we deduce that  $c_1$  also operator commutes with  $a$  and  $b$ . Therefore,  $a, b \in \mathcal{V}(c_1, 1) \oplus \mathcal{V}(c_1, 0)$  with  $c_1 \neq 0$  and  $c_1 \neq e$ . This proves that  $a$  and  $b$  are simultaneously block decomposable, which means that they are not totally non-commuting.

Conversely, suppose that  $a$  and  $b$  are simultaneously block decomposable. Then, there is an idempotent  $p \notin \{0, e\}$  such that  $a, b \in \mathcal{V}(p, 1) \oplus \mathcal{V}(p, 0)$ . If we let  $f = \text{tr}(p)^{-1}p - \text{tr}(e-p)^{-1}(e-p)$  then  $f$  is a nonzero element in  $\mathcal{V}_0$ . Furthermore, relation (27) holds because  $\langle Da, p \rangle = \langle Db, p \rangle = 0$ , since  $p$  operator commutes with  $a$  and  $b$ , and  $\langle Da, e \rangle = \langle Db, e \rangle = 0$ , since  $e$  operator commutes with each element in  $\mathcal{V}$ . Therefore,  $f$  is orthogonal to  $D\Psi(\text{Der}(\mathcal{V}) \times \text{Der}(\mathcal{V}))$ . The converse now follows from these facts.  $\square$

**Lemma 6.3.** *Let  $\mathcal{V}$  be an EJA, and let  $a, x \in \mathcal{V}$ . Suppose that  $a$  and  $x$  are totally non-commuting. Then, for every  $\epsilon > 0$ , the following statements are satisfied:*

- (a) *there exists  $x_\epsilon \in [x]_w$  such that  $\|x_\epsilon - x\| \leq \epsilon$ ,  $\lambda(x_\epsilon - a) \prec \lambda(x - a)$  and  $\lambda(x_\epsilon - a) \neq \lambda(x - a)$ ;*
- (b) *there exists  $y_\epsilon \in [x]_w$  such that  $\|y_\epsilon - x\| \leq \epsilon$ ,  $\lambda(x - a) \prec \lambda(y_\epsilon - a)$  and  $\lambda(y_\epsilon - a) \neq \lambda(x - a)$ .*

*Proof.* (a) Assume that  $a$  and  $x$  are totally non-commuting. Then, it is clear that  $a$  and  $x$  do not operator commute. In particular, by item (c) in Proposition 2.6 we get that  $(\lambda(x) - \lambda(a))^\downarrow \neq \lambda(x - a)$ .

By Lemma 6.2 we see that the map  $\Psi = \Psi_{(x,a)} : \text{Aut}(\mathcal{V}) \times \text{Aut}(\mathcal{V}) \rightarrow \mathcal{V}_t$  given by  $\Psi(X, Y) = Xx - Ya$  is a submersion at  $(I, I)$ , where  $t = \text{tr}(x - a) = \text{tr}(x) - \text{tr}(a)$ . Let  $\{e_1, \dots, e_n\}$  be a Jordan frame such that  $x - a = \sum_{i=1}^n \lambda_i(x - a) e_i$ . Consider the continuous curve  $\alpha(\cdot) : [0, 1] \rightarrow (\mathbb{R}^n)^\downarrow$  given by  $\alpha(s) = (1 - s)\lambda(x - a) + s(\lambda(x) - \lambda(a))^\downarrow \in (\mathbb{R}^n)^\downarrow$ , for  $s \in [0, 1]$ . We now define a continuous curve

$$d(\cdot) : [0, 1] \rightarrow \mathcal{V}_t, \quad d(s) = \sum_{i=1}^n \alpha_i(s) e_i, \quad s \in [0, 1].$$

Notice that by construction  $d(0) = x - a$  and  $\lambda(d(s)) = \alpha(s)$ , for  $s \in [0, 1]$ . Furthermore, we have that  $\alpha(s) \prec \lambda(x - a)$  and  $\alpha(s) \neq \lambda(x - a)$ , for  $s \in (0, 1]$ . Indeed, the majorization is because  $(\lambda(x) - \lambda(a))^\downarrow$  is in the convex hull of the permutations of  $\lambda(x - a)$ , thanks to Lidskii's inequality, and  $\alpha(s)$  is a convex combination of  $\lambda(x - a)$  and  $(\lambda(x) - \lambda(a))^\downarrow$ . The strict majorization holds because  $(\lambda(x) - \lambda(a))^\downarrow \neq \lambda(x - a)$  and  $s > 0$ . Since  $\Psi$  is a submersion at  $(I, I)$  then there exists  $\eta > 0$  such that the restriction  $\Psi : O_\eta \rightarrow \Psi(O_\eta)$  to the open set  $O_\eta := \{(X, Y) \in \text{Aut}(\mathcal{V}) \times \text{Aut}(\mathcal{V}) : \|X - I\|, \|Y - I\| \leq \eta\} \ni (I, I)$  is an open mapping onto the (relative) open set  $\Psi(O_\eta) \subset \mathcal{V}_t$  (where, as before, we are considering the product structure on  $\text{Aut}(\mathcal{V}) \times \text{Aut}(\mathcal{V})$ ). Let  $\epsilon > 0$  and assume, without loss of generality, that  $\epsilon < \eta$ ; then, there exists an  $s_0 = s_0(\epsilon) \in (0, 1)$  such that  $d(s) \in \Psi(O_\epsilon)$ , for  $0 \leq s \leq s_0$ , where  $\hat{\epsilon} := \min\{\frac{\epsilon}{2}, \frac{\epsilon}{2\|x\|}\}$ ; notice that we have used that  $\Psi(O_{\hat{\epsilon}}) \subset \mathcal{V}_t$

is a (relative) open set and that  $d(s)$  is a continuous curve such that  $d(0) = x - a = \Psi(I, I)$ . In particular, there exists  $(X, Y) \in O_\varepsilon$  such that

$$\Psi(X, Y) = Xx - Ya = d(s_0) = \sum_{i=1}^n \alpha_i(s_0) e_i.$$

Hence,  $Y^{-1}d(s_0) = Y^{-1}Xx - a$  is such that  $\lambda(Y^{-1}Xx - a) = \lambda(d(s_0)) = \alpha(s_0) \prec \lambda(x - a)$ ,  $\lambda(Y^{-1}Xx - a) \neq \lambda(x - a)$  and  $\|Y^{-1}Xx - x\| \leq \|Y^{-1}(Xx - x)\| + \|Y^{-1}x - x\| = \|(X - I)x\| + \|(Y - I)x\| \leq \varepsilon$ . Hence, we can take  $x_\varepsilon = Y^{-1}Xx \in [x]_w$ .

(b) We argue as in the proof of the previous item. Assume that  $a$  and  $x$  are totally non-commuting so that  $a$  and  $x$  do not operator commute. In particular, by Ky Fan's inequality (8) and Proposition 2.6 we get that  $\lambda(x - a) \prec \lambda(x) + \lambda(-a)$  and  $\lambda(x - a) \neq \lambda(x) + \lambda(-a)$ . Then, by Lemma 6.2 we can consider the map  $\Psi = \Psi_{(x,a)} : \text{Aut}(\mathcal{V}) \times \text{Aut}(\mathcal{V}) \rightarrow \mathcal{V}_t$  given by  $\Psi(X, Y) = Xx - Ya$  as before, which is a submersion at  $(I, I)$ . Let  $\{e_1, \dots, e_n\}$  be a Jordan frame such that  $x - a = \sum_{i=1}^n \lambda_i(x - a) e_i$ . Now, we define the continuous curve  $\beta(\cdot) : [0, 1] \rightarrow (\mathbb{R}^n)^\downarrow$  given by  $\beta(s) = (1 - s)\lambda(x - a) + s(\lambda(x) - \lambda(a)^\uparrow)^\downarrow \in (\mathbb{R}^n)^\downarrow$ , for  $s \in [0, 1]$ , and consider the continuous curve

$$d(\cdot) : [0, 1] \rightarrow \mathcal{V}_t, \quad d(s) = \sum_{i=1}^n \beta_i(s) e_i, \quad s \in [0, 1].$$

Notice that by construction  $d(0) = x - a$  and  $\lambda(d(s)) = \beta(s)$ , for  $s \in [0, 1]$ . Hence  $\lambda(x - a) \prec \beta(s)$  and  $\beta(s) \neq \lambda(x - a)$ , for  $s \in (0, 1]$ . Indeed, the majorization is because  $\lambda(x - a)$  is in the convex hull of the permutations of  $\lambda(x) + \lambda(-a)$ , thanks to Ky Fan inequality (8), and  $\beta(s)$  is a convex combination of  $\lambda(x - a)$  and  $\lambda(x) + \lambda(-a)$ . The strict majorization holds because  $\lambda(x) + \lambda(-a) \neq \lambda(x - a)$  and  $s > 0$ . Following the same steps as item (a), given  $\varepsilon > 0$  there exist  $(X, Y) \in O_\varepsilon$  such that

$$\Psi(X, Y) = Xx - Ya = d(s_0) = \sum_{i=1}^n \beta_i(s_0) e_i.$$

Hence,  $Y^{-1}d(s_0) = Y^{-1}Xx - a$  is such that  $\lambda(x - a) \prec \beta(s_0) = \lambda(d(s_0)) = \lambda(Y^{-1}Xx - a)$ ,  $\lambda(Y^{-1}Xx - a) \neq \lambda(x - a)$  and  $\|Y^{-1}Xx - x\| \leq \|Y^{-1}(Xx - x)\| + \|Y^{-1}x - x\| = \|(X - I)x\| + \|(Y - I)x\| \leq \varepsilon$ . Hence, we can take  $y_\varepsilon = Y^{-1}Xx \in [x]_w$ .  $\square$

*Proof of Lemma 3.5.* (a) Consider the class  $\mathcal{P} = \mathcal{P}_{(a,x)}(\mathcal{V})$  consisting of all nonzero idempotents  $p \in \mathcal{V}$  such that  $a, x \in \mathcal{V}(p, 1) \oplus \mathcal{V}(p, 0)$  and  $a_1$  and  $x_1$  do not operator commute, where  $a_1$  and  $x_1$  are the components (projection) of  $a$  and  $x$  in  $\mathcal{V}(p, 1)$ , respectively. Notice that for  $p \in \mathcal{P}$ ,  $\text{rank}(p) \geq 2$ . In fact, if  $\text{rank}(p) = 1$  then  $p$  is a primitive idempotent, which implies  $\mathcal{V}(p, 1) = \mathbb{R}p$  (see [3], pp. 65). Hence,  $a_1$  and  $x_1$  operator commute because they are scalar multiples of  $p$ , which is

a contradiction. Observe also that the unit element  $e$  is always in  $\mathcal{P}$ . Under our present assumptions, we claim that there exists a minimal element  $q \in \mathcal{P}$  (that is, an element of  $\mathcal{P}$  with the minimum rank) such that  $\text{rank}(q) \geq 2$ . We show this claim by induction on  $\text{rank}(V)$ . If  $\text{rank}(V) = 2$ ,  $e$  is minimal: otherwise, we can deduce that  $a$  and  $x$  operator commute, against our assumption.

Assume that our claim is true for EJAs with rank at most  $n - 1 \geq 2$  and let  $\mathcal{V}$  be an EJA with  $\text{rank}(\mathcal{V}) = n$ . Then, if  $e$  is minimal in  $\mathcal{P}$  we are done. Otherwise, there exists  $p \in \mathcal{P}$  such that  $0 \neq p \neq e$ . Recall that  $\text{rank}(p) \geq 2$ , then we can consider the EJA  $\mathcal{V}(p, 1)$  where  $a_1$  and  $x_1$  do not operator commute. Since  $2 \leq \text{rank}(\mathcal{V}(p, 1)) < n$ , we can apply the induction hypothesis and conclude that there exists a minimal element  $q \in \mathcal{P}_{(a_1, x_1)}(\mathcal{V}(p, 1))$ . It follows that  $q$  has the required properties.

Now, let  $q$  be a minimal element in  $\mathcal{P}$ , and consider the decomposition of  $a$  and  $x$  in the direct sum  $\mathcal{V}(q, 1) \oplus \mathcal{V}(q, 0)$  as  $a = \tilde{a} + \hat{a}$ ,  $x = \tilde{x} + \hat{x}$ . Since  $q$  is minimal, we have that  $\tilde{a}$  and  $\tilde{x}$  are totally non-commuting in  $\mathcal{V}(q, 1)$ . Indeed, otherwise,  $\tilde{a}$  and  $\tilde{x}$  are simultaneously block decomposable, which allows us to construct an element in  $\mathcal{P}$  with rank smaller than  $\text{rank}(q)$ . Take  $\varepsilon > 0$ ; by item (a) in Lemma 6.3 there exists  $\tilde{x}_\varepsilon \in [\tilde{x}]_w$ , where the weak orbit is relative to  $\mathcal{V}(q, 1)$ , such that  $\|\tilde{x}_\varepsilon - \tilde{x}\| \leq \varepsilon$ ,  $\lambda(\tilde{x}_\varepsilon - \tilde{a}) \prec \lambda(\tilde{x} - \tilde{a})$  and  $\lambda(\tilde{x}_\varepsilon - \tilde{a}) \neq \lambda(\tilde{x} - \tilde{a})$ . Finally, we take  $x_\varepsilon = \tilde{x}_\varepsilon + \hat{x} \in \mathcal{V}$ ; notice that by construction  $\lambda(x_\varepsilon) = \lambda(x)$  and hence  $x_\varepsilon \in [x]$ . On the other hand,  $\|x_\varepsilon - x\| = \|\tilde{x}_\varepsilon - \tilde{x}\| \leq \varepsilon$ . Finally, using Lemma 3.6 we now get that

$$\lambda(x_\varepsilon - a) = (\lambda(\tilde{x}_\varepsilon - \tilde{a}), \lambda(\hat{x} - \hat{a}))^\downarrow \prec (\lambda(\tilde{x} - \tilde{a}), \lambda(\hat{x} - \hat{a}))^\downarrow = \lambda(x - a).$$

Arguing as above we also see that  $\lambda(x_\varepsilon - a) \neq \lambda(x - a)$ .

(b) The proof is analogous to the previous item, using item (b) in Lemma 6.3.  $\square$

We now prove Lemma 3.9. Our approach will use the following elementary fact: given  $u = (u_1, u_2)$ ,  $v = (v_1, v_2) \in \mathbb{R}^2$  such that  $u_1 + u_2 = v_1 + v_2$ , if we assume that  $\|u\|^2 < \|v\|^2$  then  $u$  is strictly majorized by  $v$ ; that is,  $u \prec v$  and  $u^\downarrow \neq v^\downarrow$ . The idea of the proof of the above fact follows a geometric argument: Let  $L$  be the line of points in  $\mathbb{R}^2$  whose components sum to  $u_1 + u_2 = v_1 + v_2$ . Hence,  $u, v \in L$ . Notice that  $v_1 \neq v_2$  because, otherwise,  $v$  will be the vector on  $L$  of minimal norm contradicting  $\|u\| < \|v\|$ . Now,  $L$  is the line that passes through  $(v_1, v_2)$  and  $(v_2, v_1)$ . The fact that  $\|u\| < \|v\|$  will force that  $u$  is in the relative interior of the convex hull of  $(v_1, v_2)$  and  $(v_2, v_1)$ . This implies the desired result.

*Proof of Lemma 3.9.* That  $a$  and  $x$  operator commute means that there exist a Jordan frame  $\{e_1, e_2\}$  and scalars  $\lambda_1, \lambda_2, \beta_1, \beta_2$  such that  $a = \lambda_1 e_1 + \lambda_2 e_2$  and  $x = \beta_1 e_1 + \beta_2 e_2$ . We first prove item (a). Since  $a$  and  $x$  do not strongly operator commute we may assume that  $\lambda_1 > \lambda_2$  and  $\beta_1 < \beta_2$ . The fact that  $\mathcal{V}$  is simple ensures that there exists  $w \in \mathcal{V}(e_1, 1/2) \cap \mathcal{V}(e_2, 1/2)$  with  $\|w\|^2 = 2$ . In

this case, we have that  $w^2 = e_1 + e_2$  (see [3, Proposition IV.1.4.]). Let  $\mathcal{D}$  be the vector subspace generated by  $e_1, e_2$  and  $w$ . That is,  $\mathcal{D} := \mathbb{R}e_1 + \mathbb{R}e_2 + \mathbb{R}w$ . Thus,  $\mathcal{D}$  is a 3 dimensional vector space and constitutes a simple Euclidean Jordan algebra of rank 2 equipped with the inner and Jordan products inherited from  $\mathcal{V}$ . Furthermore,  $\mathcal{D}$  is isomorphic to  $\mathcal{S}^2$  (space of  $2 \times 2$  symmetric matrices) in terms of the isomorphism  $\phi : \mathcal{S}^2 \rightarrow \mathcal{D}$  given by

$$\phi \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = e_1, \quad \phi \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = e_2, \quad \phi \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = w. \quad (28)$$

For  $0 \leq \theta < \pi/2$  we set

$$\begin{aligned} e_1(\theta) &= \cos(\theta)^2 e_1 + \cos(\theta) \sin(\theta) w + \sin(\theta)^2 e_2, \\ e_2(\theta) &= \sin(\theta)^2 e_1 - \cos(\theta) \sin(\theta) w + \cos(\theta)^2 e_2. \end{aligned}$$

Notice that  $e_1(\theta), e_2(\theta) : [0, \pi/2) \rightarrow \mathcal{V}$  are continuous functions such that  $\{e_1(\theta), e_2(\theta)\}$  is a Jordan frame for  $\mathcal{D}$ , for  $\theta \in [0, \pi/2)$ . Moreover, notice that  $(e_1(0), e_2(0)) = (e_1, e_2)$ . We now construct the continuous curve

$$\gamma(\theta) := \beta_1 e_1(\theta) + \beta_2 e_2(\theta), \quad (29)$$

for all  $\theta \in [0, \pi/2)$ . Observe that  $\gamma(0) = x$ . Let us prove that item (a) is satisfied. To see this, let  $D(\theta) = \phi^{-1}(\gamma(\theta) - a) \in \mathcal{S}^2$  so that  $\lambda(\gamma(\theta) - a) = \lambda(D(\theta))$ , for  $\theta \in [0, \pi/2)$ , since  $\phi$  is an algebra isomorphism. By construction, we have that

$$D(\theta) = V(\theta) \begin{pmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix} V(\theta)^\top - \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \text{with} \quad V(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

for all  $\theta \in [0, \pi/2)$ . Consider

$$R(\theta) := \begin{pmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix} - V(\theta)^\top \begin{pmatrix} \lambda_1 - \lambda_2 & 0 \\ 0 & 0 \end{pmatrix} V(\theta). \quad (30)$$

Then,

$$D(\theta) = V(\theta) R(\theta) V(\theta)^\top - \lambda_2 I_2. \quad (31)$$

Hence, using the elementary fact mentioned right before this proof, we can deduce that  $\lambda(R(\theta)) \prec \lambda(R(0))$  and  $\lambda(R(\theta)) \neq \lambda(R(0))$ , for  $\theta \in (0, \pi/2)$ . Indeed, observe that  $\text{tr}(R(\theta))$  is constant for every  $\theta \in [0, \pi/2)$ , we only need to show that the function  $[0, \pi/2) \ni \theta \mapsto \|R(\theta)\|_2^2$  is strictly decreasing on  $[0, \pi/2)$ , where  $\|X\|_2 = (\lambda_1(X)^2 + \lambda_2(X)^2)^{1/2}$  stands for the Frobenius norm of  $X \in \mathcal{S}^2$ . Indeed,

since  $\lambda_1 - \lambda_2 > 0$ , we have that

$$V(\theta)^\top \begin{pmatrix} \lambda_1 - \lambda_2 & 0 \\ 0 & 0 \end{pmatrix} V(\theta) = g(\theta)g(\theta)^\top,$$

where  $g(\theta) = (\lambda_1 - \lambda_2)^{1/2}(\cos(\theta), \sin(\theta))^\top$ . If  $M \in \mathcal{S}^2$  is the diagonal matrix with main diagonal  $(\beta_1, \beta_2)$  then  $R(\theta) = M - g(\theta)g(\theta)^\top$ . Thus,  $\|R(\theta)\|_2^2 = \gamma - 2\langle M g(\theta), g(\theta) \rangle$ , where  $\gamma := \|g(\theta)\|^4 + \beta_1^2 + \beta_2^2 = (\lambda_1 - \lambda_2)^2 + \beta_1^2 + \beta_2^2 \in \mathbb{R}$  is a constant and

$$h(\theta) := \langle M g(\theta), g(\theta) \rangle = (\lambda_1 - \lambda_2) (\cos^2(\theta) \beta_1 + \sin^2(\theta) \beta_2),$$

is strictly increasing in  $[0, \pi/2)$ . In fact,  $h'(\theta) = (\lambda_1 - \lambda_2)(\beta_2 - \beta_1) \sin(2\theta) > 0$  for  $\theta \in (0, \pi/2)$ , since the function  $\sin(2\theta)$  is positive for  $\theta \in (0, \pi/2)$ ,  $\lambda_1 - \lambda_2 > 0$  and  $\beta_2 - \beta_1 > 0$ . Therefore,  $\|R(\theta)\|_2^2 < \|R(0)\|_2^2$  for every  $\theta \in (0, \pi/2)$ ; hence, using the elementary fact mentioned right before this proof, we get that  $\lambda(R(\theta)) \prec \lambda(R(0))$  and  $\lambda(R(\theta)) \neq \lambda(R(0))$  for every  $\theta \in (0, \pi/2)$ . The previous facts and Eq. (31) imply that

$$\lambda(D(\theta)) = \lambda(R(\theta)) - \lambda_2(1, 1)^\top \implies \lambda(D(\theta)) \prec \lambda(D(0)), \lambda(D(\theta)) \neq \lambda(D(0)), \theta \in (0, \pi/2).$$

The previous facts imply that  $\lambda(\gamma(\theta) - a) \prec \lambda(x - a)$  and  $\lambda(\gamma(\theta) - a) \neq \lambda(x - a)$  for every  $\theta \in (0, \pi/2)$ .

To prove item (b) we notice that since  $-a$  and  $x$  do not strongly operator commute we may assume that  $\lambda_1 > \lambda_2$  and  $\beta_1 > \beta_2$ . We now argue as in the first part of the proof and construct the isomorphism  $\phi : \mathcal{S}^2 \rightarrow \mathcal{D}$  according to Eq. (28), the curve  $\gamma(\theta)$  according to (29) and similarly  $D(\theta) = \phi^{-1}(\gamma(\theta) - a)$ , for  $\theta \in [0, \pi/2)$ . Consider  $R(\theta)$  as in Eq. (30). Arguing as before, to check item (b) we only need to show that  $[0, \pi/2) \ni \theta \mapsto \|R(\theta)\|_2^2$  is strictly increasing on  $[0, \pi/2)$ . As before, we only need to check that  $h(\theta) := \langle M g(\theta), g(\theta) \rangle = (\lambda_1 - \lambda_2) (\cos^2(\theta) \beta_1 + \sin^2(\theta) \beta_2)$ , is strictly decreasing in  $[0, \pi/2)$ ; but this fact follows from the inequality  $h'(\theta) = (\lambda_1 - \lambda_2)(\beta_2 - \beta_1) \sin(2\theta) < 0$  for  $\theta \in (0, \pi/2)$ , where we used that the function  $\sin(2\theta)$  is positive for  $\theta \in (0, \pi/2)$ ,  $\lambda_1 - \lambda_2 > 0$  and  $\beta_2 - \beta_1 < 0$  in the present case. □

## Data availability

No data was used for the research described in the article.

## References

- [1] Benac, M.J., Massey, P., Rios, N. B., Ruiz, M.: Local minimizers of the distances to the majorization flows. *J. Phys. A: Math. Theor.*, 56, 505203 (2023)
- [2] Cambini, A., Martein, L.: Generalized convexity and optimization: theory and applications. *Lecture Notes in Economics and Mathematical Systems*, vol. 616, Springer, Berlin, (2009)
- [3] Faraut, J., Korányi, A.: *Analysis on symmetric cones*. Clarendon Press, Oxford, UK (1994)
- [4] Gowda, M. S.: Commutation principles for optimization problems on spectral sets in Euclidean Jordan algebras. *Optim. Lett.*, 16, 1119–1128 (2022)
- [5] Gowda, M. S., Jeong, J.: Commutation principles in Euclidean Jordan algebras and normal decomposition systems. *SIAM J. Optim.*, 27, 1390–1402 (2017)
- [6] Gowda, M. S.: Optimizing certain combinations of spectral and linear/distance functions over spectral sets. *arXiv:1902.06640v2* (2019)
- [7] C. Greif, J. M. Varah: Minimizing the condition number for small rank modifications. *SIAM J. Matrix Anal. Appl.*, 29, 82–97 (2006)
- [8] Jeong, J., Jung, Y.M., Lim, Y.W.: Weak majorization, doubly substochastic maps, and some related inequalities in Euclidean Jordan algebras. *Linear Algebra Appl.*, 597, 133–154 (2020)
- [9] Jeong, J., Sossa, D.: Commutation principles for non-smooth variational problems on Euclidean Jordan algebras. *SIAM J. Optim.*, 35.3, 1551–1569 (2025).
- [10] Knyazev, A. K., Argentati, M. E.: Rayleigh-Ritz majorization error bounds with applications to FEM. *SIAM J. Matrix Anal. Appl.* 31.3, 1521–1537 (2009)
- [11] Lewis, A. S.: Lidskii’s theorem via nonsmooth analysis. *SIAM J. Matrix Anal. Appl.*, 21, 379–381 (1999)
- [12] Lourenço, B. F., Takeda, A.: Generalized subdifferentials of spectral functions over Euclidean Jordan algebras. *SIAM J. Optim.*, 30, 3387–3414 (2020)
- [13] Marshall, A. W., Olkin, I., Arnold, B. C.: *Inequalities: Theory of Majorization and Its Applications*. Springer, New York (2010).
- [14] Massey, P. G., Rios, N. B., Stojanoff, D.: Local Lidskii’s theorems for unitarily invariant norms. *Linear Algebra Appl.*, 557, 34–61 (2018)
- [15] Massey, P. G., Stojanoff, D., Zárate, S.: The spectral spread of Hermitian matrices. *Linear Algebra Appl.* 616, 19–44 (2021)
- [16] Massey, P. G., Stojanoff, D., Zárate, S.: Absolute variation of Ritz values, principal angles, and spectral spread. *SIAM J. Matrix Anal. Appl.* 42.4, 1506–1527 (2021)
- [17] Ramírez, H., Seeger, A., Sossa, D.: Commutation principle for variational problems on Euclidean Jordan algebras. *SIAM J. Optim.*, 23, 687–694 (2013)

- [18] Seeger, A.: Condition number minimization in Euclidean Jordan algebras. *SIAM J. Optim.*, 32, 635–658 (2022)
- [19] Sossa, D.: A Fiedler-type determinantal inequality in Euclidean Jordan algebras. *Linear Algebra Appl.*, 667, 151–164 (2023)
- [20] Stępniański, C.: An effective characterization of Schur-convex functions with applications. *J. Convex Anal.*, 14, 103–108 (2007)