Optimal completions of a frame.

P. G. Massey, M. A. Ruiz and D. Stojanoff
Depto. de Matemática, FCE-UNLP, La Plata, Argentina and IAM-CONICET *

Abstract

Given a finite sequence of vectors $\mathcal{F}_0$ in $\mathbb{C}^d$ we describe the spectral and geometrical structure of optimal completions of $\mathcal{F}_0$ obtained by adding a finite sequence of vectors with prescribed norms, where optimality is measured with respect to a general convex potential. In particular, our analysis includes the so-called Mean Square Error (MSE) and the Benedetto-Fickus' frame potential. On a first step, we reduce the problem of finding the optimal completions to the computation of the minimum of a convex function in a convex compact polytope in $\mathbb{R}^d$. As a second step, we show that there exists a finite set (that can be explicitly computed in terms of a finite step algorithm that depends on $\mathcal{F}_0$ and the sequence of prescribed norms) such that the optimal frame completions with respect to a given convex potential can be described in terms of a distinguished element of this set. As a byproduct we characterize the cases of equality in Lidskii's inequality from matrix theory.

AMS subject classification: 42C15, 15A60.
Keywords: frames, frame completions, majorization, Lidskii's inequality, Schur-Horn theorem.

Contents

*e-mail addresses: massey@mate.unlp.edu.ar , mruiz@mate.unlp.edu.ar , demetrio@mate.unlp.edu.ar

1 Introduction

A finite sequence of vectors $\mathcal{F} = \{f_i\}_{i\in\mathbb{I}_n}$ in $\mathbb{C}^d$ is a frame for $\mathbb{C}^d$ if the sequence spans $\mathbb{C}^d$. It is well known that finite frames provide redundant linear encoding-decoding schemes, that have proved useful in real life applications. Conversely, several research problems in this field have arise in the attempt to apply this theory in different contexts.

For example, the (linear) redundancy provided by finite frames translates into robustness properties of the transmission scheme that they induce, which make frames a useful device for transmission of signals through noisy channels; this last fact has posed several problems dealing with the determination of what is known in the literature as optimal frames for erasures (see [4, 5, 6, 15, 23, 28, 27]).

On the other hand, the so-called tight frames allow for redundant linear representations of vectors (signals) that are formally analogous to the linear representations given by orthonormal basis; this feature makes tight frames a distinguished class of frames that is of interest for applications. Conversely, in several applications we would like to consider tight frames that have some other prescribed properties leading to what is known in the literature as frame design problems [1, 7, 10, 13, 17, 18, 19, 26]. It is worth pointing out that in some cases it is not possible to find a frame fulfilling the previous demands; in [2] Benedetto and Fickus found an alternative approach to these situations by introducing a functional, called the frame potential, and showing that minimizers of the frame potential (within a convenient set of frames) are the natural substitutes of tight frames with prescribed parameters (see also [14, 21, 24, 30] and [11, 31, 32] for related problems in the context of fusion frames).

Recently, the following frame completion problem, related with the frame design problems mentioned above, was posed in [20]: given an initial sequence $\mathcal{F}_0$ in $\mathbb{C}^d$ and a sequence of positive numbers $\mathbf{a}$ then compute the sequences $\mathcal{G}$ in $\mathbb{C}^d$ whose elements have norms given by the sequence $\mathbf{a}$ and such that the completed sequence $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G})$ minimizes the so-called mean square error (MSE) of $\mathcal{F}$, which is a (convex) functional (see also [8, 19, 29] for completion problems for frames). The initial sequence of vectors can be considered as a checking device for the measurement, and therefore we search for a complementary set of measurements (given by vectors with prescribed norms) in such a way that the complete set of measurements is optimal with respect to the MSE. Notice there are other possible (convex) functionals that we could choose to minimize such as, for example, the frame potential. Therefore, a natural extension of the previous problem is: given a (convex) functional defined on the set of frames, compute the frame completions with prescribed norms that minimize this functional.

A first step towards the solution of this general version of the completion problem was made in [33]. There we showed that under certain hypothesis (feasible cases, see Section 2.4), optimal frame completions with prescribed norms do not depend on the particular choice of convex functional, as long as we consider *convex potentials*, that contain the MSE and the frame potential (see Section 2.2). On the other hand, it is easy to show examples in which the previous result does not apply (non-feasible cases); in these cases the optimal frame completions with prescribed norms are not known even for the MSE nor the frame potential.

In this paper we consider the frame completion problem of an initial sequence $\mathcal{F}_0$ in $\mathbb{C}^d$, for general sequences $\mathbf{a}$ of prescribed norms and for a fixed convex potential P_f - where f is a strictly convex function - in the non-feasible cases (see Section 2.4 for motivations and a detailed description of our main problem). In order to deal with the general problem we introduce and develop a class of pairs of positive matrices (called optimal matchings matrices, see the Appendix) that allow to reduce the problem to the computation of minimizers of a scalar convex function F (associated to f) in a compact convex domain in $\mathbb{R}^d$ (the same set for every map f). This constitutes a reduction of the optimization problem, that in turn can be attacked with several numerical tools in concrete examples. In fact, the convex domain has a natural and explicit description in terms of majorization, which is an algorithmic notion.

We also study the spectral and geometrical structure of local minimizers of P_f in the set of frame completions with prescribed norms, in terms of a geometrical approach to a perturbation problem. These last results allow to a second reduction of the problem: there is a finite set $E(\mathcal{F}_0\,,\,\mathbf{a})$ in $\mathbb{R}^d$ - that depends only on the initial family $\mathcal{F}_0$ and the finite sequence $\mathbf{a}$ of positive numbers - such that for any fixed convex potential P_f there exists a unique vector $\mu = \mu_f \in E(\mathcal{F}_0\,,\,\mathbf{a})$ (computable by a minimization on the finite set $E(\mathcal{F}_0\,,\,\mathbf{a})$ in terms of F) such that all optimal frame completions for P_f with prescribed norms can be computed in terms of μ.

In both methods, we describe the optimal vector of eigenvalues for the frame operator of the completing sequences. With this data, the optimal completions (which satisfy the norm restrictions) can be effectively computed by using a well known algorithm developed in [17] that implements the Schur-Horn theorem.

In all examples that we have computed numerically, we have found that the optimal spectrum of the completing sequences does not depend on the particular choice of convex potential P_f considered. Although at the present we have not been able to prove this fact, we state it as a conjecture. We have also observed two other common features of optimal solutions - that are also stated as conjectures - that allow to implement an efficient (and considerably faster) algorithm that computes an smaller set than $E(\mathcal{F}_0\,,\,\mathbf{a})$ that also enables to compute the optimal frame completions with prescribed norms with respect to a general convex potential P_f.

The paper is organized as follows. In Section 2 we state several facts and notions about frame theory in finite dimension and majorization, which is a notion from matrix theory; in this section we describe in detail the main problem of the present paper and some previous related results. In section 3 we reduce the problem of computing optimal frame completions with prescribed norms to a set of completions whose frame operators are optimal matchings of the frame operator of the initial set of vectors $\mathcal{F}_0$, in the sense described in the Section 6 (Appendix). Based on the results of the Appendix we obtain a first reduction of the problem and show that the optimal frame completions with prescribed norms for the convex potential P_f can be described in terms of the minimizers of an associated function F in a compact convex polytope. We also show that the spectral structure of optimal completions is unique and has some other features. In Section 4 we introduce two different topologies in the set of completions and consider the geometrical structure of local minimizers with respect to these topologies; in order to do this we apply tools from differential geometry that allow to solve a local perturbation problem for frames with prescribed norms. Using these results we show in Section 5 that optimal completions $\mathcal{F} = (\mathcal{F}_0, \mathcal{G})$ have the property that the vectors of the completing sequence $\mathcal{G}$ are eigenvectors of the frame operator $S_\mathcal{F}$ of the complete sequence $\mathcal{F}$. Based on this last fact we develop an algorithm (that can be effectively implemented) to compute optimal completions numerically. The analysis of the computed examples reflects some commons features of the numerical solutions. Based on these facts we state some other conjectures related with the spectral structure of optimal completions. Finally, in Section 6 we introduce pairs of positive matrices, that we call optimal matchings, and describe the structure of these pairs; this corresponds to the study of the case of equality in Lidskii's inequality from matrix theory.

2 Preliminaries

In this section we describe the basic notions that we shall consider throughout the paper. We first establish the general notations and then we recall the basic facts from frame theory that are related with our main results. Then, we describe submajorization which is a notion from matrix analysis, that will play a major role in this note. Finally, we recall the solution of the frame design problem in terms of majorization and give a detailed description of the optimal frame completion problem, which is the main topic of this paper.

2.1 General notations.

Given $m \in \mathbb{N}$ we denote by $\mathbb{I}_m = \{1, \dots, m\} \subseteq \mathbb{N}$ and $\mathbb{1} = \mathbb{1}_m \in \mathbb{R}^m$ denotes the vector with all its entries equal to 1. For a vector $x \in \mathbb{R}^m$ we denote by $x^\downarrow$ (resp. $x^\uparrow$) the rearrangement of x in decreasing (resp. increasing) order, and $(\mathbb{R}^m)^\downarrow = \{x \in \mathbb{R}^m : x = x^\downarrow\}$ the set of downwards ordered vectors.

Given $\mathcal{H} \cong \mathbb{C}^d$ and $\mathcal{K} \cong \mathbb{C}^n$, we denote by $L(\mathcal{H}, \mathcal{K})$ the space of linear operators $T : \mathcal{H} \to \mathcal{K}$. Given an operator $T \in L(\mathcal{H}, \mathcal{K})$, $R(T) \subseteq \mathcal{K}$ denotes the image of T, $\ker T \subseteq \mathcal{H}$ the null space of T and $T^* \in L(\mathcal{K}, \mathcal{H})$ the adjoint of T. If $\mathcal{K} = \mathcal{H}$ we denote by $L(\mathcal{H}) = L(\mathcal{H}, \mathcal{H})$, by $\mathcal{G}l(\mathcal{H})$ the group of all invertible operators in $L(\mathcal{H})$, by $L(\mathcal{H})^+$ the cone of positive operators and by $\mathcal{G}l(\mathcal{H})^+ = \mathcal{G}l(\mathcal{H}) \cap L(\mathcal{H})^+$. If $T \in L(\mathcal{H})$, we denote by $\sigma(T)$ the spectrum of T, by $\operatorname{rk} T = \dim R(T)$ the rank of T, and by $\operatorname{tr} T$ the trace of T.

By fixing orthonormal basis's (ONB's) of the Hilbert spaces involved, we shall identify operators with matrices, using the following notations: by $\mathcal{M}_{n,d}(\mathbb{C}) \cong L(\mathbb{C}^d, \mathbb{C}^n)$ we denote the space of complex $n \times d$ matrices. If $n = d$ we write $\mathcal{M}_d(\mathbb{C}) = \mathcal{M}_{d,d}(\mathbb{C})$; $\mathcal{H}(d)$ is the $\mathbb{R}$-subspace of selfadjoint matrices, $\mathcal{G}l(d)$ the group of all invertible elements of $\mathcal{M}_d(\mathbb{C})$, $\mathcal{U}(d)$ the group of unitary matrices, $\mathcal{M}_d(\mathbb{C})^+$ the set of positive semidefinite matrices, and $\mathcal{G}l(d)^+ = \mathcal{M}_d(\mathbb{C})^+ \cap \mathcal{G}l(d)$.

If $W \subseteq \mathcal{H}$ is a subspace we denote by $P_W \in L(\mathcal{H})^+$ the orthogonal projection onto W, i.e. $R(P_W) = W$ and $\ker P_W = W^\perp$. Given $x, y \in \mathcal{H}$ we denote by $x \otimes y \in L(\mathcal{H})$ the rank one operator given by $x \otimes y\,(z) = \langle z, y\rangle\, x$ for every $z \in \mathcal{H}$. Note that if $\|x\| = 1$ then $x \otimes x = P_{\operatorname{span}\{x\}}$.

Given $S \in \mathcal{M}_d(\mathbb{C})^+$, we write $\lambda(S) \in (\mathbb{R}^d_{\geq 0})^\downarrow$ the vector of eigenvalues of S - counting multiplicities - arranged in decreasing order. If $\lambda(S) = \lambda = (\lambda_1, \dots, \lambda_d) \in (\mathbb{R}^d_{\geq 0})^\downarrow$, a system $\mathcal{B} = \{h_i\}_{i\in\mathbb{I}_d} \subseteq \mathbb{C}^d$ is a "ONB of eigenvectors for S, λ" if it is an orthonormal basis for $\mathbb{C}^d$ such that $S\,h_i = \lambda_i\, h_i$ for every $i \in \mathbb{I}_d$. In other words, an orthonormal basis

$$\mathcal{B} = \{h_i\}_{i\in\mathbb{I}_d} \quad \text{is a "ONB of eigenvectors for } S\,,\ \lambda\text{"} \iff S = \sum_{i\in\mathbb{I}_d} \lambda_i \cdot h_i \otimes h_i \ . \tag{1}$$

For vectors in $\mathbb{C}^d$ we shall use the euclidean norm. On the other hand, for $T \in \mathcal{M}_{n,d}(\mathbb{C})$ we shall use the spectral norm, denoted $\|T\|$, given by $\|T\| = \max\limits_{\|x\|=1} \|Tx\|$.

2.2 Basic framework of finite frames

In what follows we consider (n, d)-frames. See [2, 9, 16, 22, 30] for detailed expositions of several aspects of this notion.

Let $d, n \in \mathbb{N}$, with $d \leq n$. Fix a Hilbert space $\mathcal{H} \cong \mathbb{C}^d$. A family $\mathcal{F} = \{f_i\}_{i\in\mathbb{I}_n} \in \mathcal{H}^n$ is an (n, d)-frame for $\mathcal{H}$ if there exist constants $A, B > 0$ such that

$$A\|x\|^2 \leq \sum_{i=1}^n |\langle x, f_i\rangle|^2 \leq B\|x\|^2 \quad \text{for every} \quad x \in \mathcal{H} \ . \tag{2}$$

The **frame bounds**, denoted by $A_\mathcal{F}, B_\mathcal{F}$ are the optimal constants in (2). If $A_\mathcal{F} = B_\mathcal{F}$ we call $\mathcal{F}$ a tight frame. Since $\dim \mathcal{H} < \infty$, a family $\mathcal{F} = \{f_i\}_{i\in\mathbb{I}_n}$ is an (n, d)-frame if and only if $\operatorname{span}\{f_i : i \in \mathbb{I}_n\} = \mathcal{H}$. We shall denote by $\mathbf{F} = \mathbf{F}(n, d)$ the set of all (n, d)-frames for $\mathcal{H}$.

Given $\mathcal{F} = \{f_i\}_{i\in\mathbb{I}_n} \in \mathcal{H}^n$, the operator $T_\mathcal{F} \in L(\mathcal{H}, \mathbb{C}^n)$ defined by

$$T_\mathcal{F}\, x = \big(\langle x, f_i\rangle\big)_{i\in\mathbb{I}_n} \ , \quad \text{for every} \quad x \in \mathcal{H} \tag{3}$$

is the **analysis** operator of $\mathcal{F}$. Its adjoint $T_\mathcal{F}^*$ is called the **synthesis** operator:

$$T_\mathcal{F}^* \in L(\mathbb{C}^n, \mathcal{H}) \quad \text{given by} \quad T_\mathcal{F}^*\, v = \sum_{i\in\mathbb{I}_m} v_i\, f_i \quad \text{for every} \quad v = (v_1, \dots, v_n) \in \mathbb{C}^n \ .$$

Finally, we define the **frame operator** of $\mathcal{F}$ as

$$S_{\mathcal{F}} = T_{\mathcal{F}}^* \, T_{\mathcal{F}} = \sum_{i\in\mathbb{I}_n} f_i \otimes f_i \in L(\mathcal{H})^+ \ .$$

Notice that, if $\mathcal{F} = \{f_i\}_{i\in\mathbb{I}_n} \in \mathcal{H}^n$ then $\langle S_{\mathcal{F}}\, x \, , \, x\rangle = \sum_{i\in\mathbb{I}_n} \big| \langle x \, , \, f_i\rangle \big|^2$ for every $x \in \mathcal{H}$. Hence, $\mathcal{F} \in \mathbf{F}(n \, , \, d)$ if and only if $S_{\mathcal{F}} \in \mathcal{G}l(\mathcal{H})^+$ and in this case $A_{\mathcal{F}} \, \|x\|^2 \le \langle S_{\mathcal{F}}\, x \, , \, x\rangle \le B_{\mathcal{F}} \, \|x\|^2$ for every $x \in \mathcal{H}$. In particular, $A_{\mathcal{F}} = \lambda_{\min}(S_{\mathcal{F}}) = \|S_{\mathcal{F}}^{-1}\|^{-1}$ and $\lambda_{\max}(S_{\mathcal{F}}) = \|S_{\mathcal{F}}\| = B_{\mathcal{F}}$. Moreover, $\mathcal{F}$ is tight if and only if $S_{\mathcal{F}} = \frac{\tau}{d} \, I_{\mathcal{H}}$, where $\tau = \operatorname{tr} S_{\mathcal{F}} = \sum_{i\in\mathbb{I}_n} \|f_i\|^2$.

The frame operator plays an important role in the reconstruction of a vector x using its frame coefficients $\{\langle x \, , \, f_i\rangle\}_{i\in\mathbb{I}_n}$. This leads to the definition of the canonical dual frame associated to $\mathcal{F}$: for every $\mathcal{F} = \{f_i\}_{i\in\mathbb{I}_n} \in \mathbf{F}(n \, , \, d)$, the **canonical dual** frame associated to $\mathcal{F}$ is the sequence $\mathcal{F}^\# \in \mathbf{F}$ defined by

$$\mathcal{F}^\# \stackrel{\text{def}}{=} S_{\mathcal{F}}^{-1} \cdot \mathcal{F} = \{S_{\mathcal{F}}^{-1} \, f_i\}_{i\in\mathbb{I}_m} \in \mathbf{F}(n \, , \, d) \ .$$

Therefore, we obtain the reconstruction formulas

$$x = \sum_{i\in\mathbb{I}_n} \langle x \, , \, f_i\rangle \, S_{\mathcal{F}}^{-1} \, f_i = \sum_{i\in\mathbb{I}_n} \langle x \, , \, S_{\mathcal{F}}^{-1} \, f_i\rangle \, f_i \quad \text{for every} \quad x \in \mathcal{H} \ . \tag{4}$$

Observe that the canonical dual $\mathcal{F}^\#$ satisfies that given $x \in \mathcal{H}$, then

$$T_{\mathcal{F}^\#} \, x = \big(\langle x \, , \, S_{\mathcal{F}}^{-1} \, f_i\rangle \big)_{i\in\mathbb{I}_n} = \big(\langle S_{\mathcal{F}}^{-1} \, x \, , \, f_i\rangle \big)_{i\in\mathbb{I}_n} \quad \text{for} \quad x \in \mathcal{H} \implies T_{\mathcal{F}^\#} = T_{\mathcal{F}} \, S_{\mathcal{F}}^{-1} \ . \tag{5}$$

Hence $T_{\mathcal{F}^\#}^* \, T_{\mathcal{F}} = I_{\mathcal{H}}$ and $S_{\mathcal{F}^\#} = S_{\mathcal{F}}^{-1} \, T_{\mathcal{F}}^* \, T_{\mathcal{F}} \, S_{\mathcal{F}}^{-1} = S_{\mathcal{F}}^{-1}$.

In their seminal work [2], Benedetto and Fickus introduced a functional defined (on unit norm frames), the so-called frame potential, given by

$$\operatorname{FP}\left(\{f_i\}_{i\in\mathbb{I}_n}\right) = \sum_{i,\, j\, \in\mathbb{I}_n} |\langle f_i \, , \, f_j\rangle|^2 \ .$$

One of their major results shows that tight unit norm frames - which form an important class of frames because of their simple reconstruction formulas - can be characterized as (local) minimizers of this functional among unit norm frames. Since then, there has been interest in (local) minimizers of the frame potential within certain classes of frames, since such minimizers can be considered as natural substitutes of tight frames (see [14, 30, 31]). Notice that, given $\mathcal{F} = \{f_i\}_{i\in\mathbb{I}_n} \in \mathcal{H}^n$ then $\operatorname{FP}(\mathcal{F}) = \operatorname{tr} S_{\mathcal{F}}^2 = \sum_{i\in\mathbb{I}_d} \lambda_i(S_{\mathcal{F}})^2$. These remarks have motivated the definition of general convex potentials as follows:

Definition 2.1. Let $f : [0 \, , \, \infty) \to [0 \, , \, \infty)$ be a convex function. Following [30] we consider the (generalized) convex potential associated to f, denoted P_f, given by

$$P_f(\mathcal{F}) = \operatorname{tr} \, f(S_{\mathcal{F}}) \quad \text{for} \quad \mathcal{F} = \{f_i\}_{i\in\mathbb{I}_n} \in \mathcal{H}^n \ . \qquad \triangle$$

Of course, one of the most important convex potential is the Benedetto-Fickus' (BF) frame potential. As shown in [30, Sec. 4] these convex functionals (which are related with the so-called entropic measures of frames) share many properties with the BF-frame potential. Indeed, under certain restrictions both the spectral and geometric structures of minimizers of these potentials coincide (see [30]).

Remark 2.2. The results that we shall develop in this work apply in the case of convex potentials P_f for a strictly convex function $f : [0, \infty) \to \mathbb{R}$. Notice that this formulation does not formally include the Mean Square Error (MSE), which is the convex potential associated with the strictly convex function $f : (0, \infty) \to (0, \infty)$ given by $f(x) = x^{-1}$, since f is not defined in 0 in this case. In order to include the MSE within our results we proceed as follows: we define $\tilde{f} : [0, \infty) \to (0, \infty]$ given by $\tilde{f}(x) = x^{-1}$ for $x > 0$ and $\tilde{f}(0) = \infty$. Assuming that $x < \infty$ and $x + \infty = x \cdot \infty = \infty$ for every $x \in (0 \, , \, \infty)$, it turns out that the new map $\tilde{f}$ is a (extended) strictly convex function and all the results obtained in this paper apply to the convex potential induced by $\tilde{f}$. $\triangle$

2.3 Submajorization

Next we briefly describe submajorization, a notion from matrix analysis theory that will be used throughout the paper. For a detailed exposition of submajorization see [3].

Given $x,\, y \in \mathbb{R}^d$ we say that x is **submajorized** by y, and write $x \prec_w y$, if

$$\sum_{i=1}^{k} x_i^{\downarrow} \le \sum_{i=1}^{k} y_i^{\downarrow} \quad \text{for every} \quad k \in \mathbb{I}_d \ .$$

If $x \prec_w y$ and $\operatorname{tr} x = \sum_{i=1}^{d} x_i = \sum_{i=1}^{d} y_i = \operatorname{tr} y$, then we say that x is **majorized** by y, and write $x \prec y$. If the two vectors x and y have different size, we write $x \prec y$ if the extended vectors (completing with zeros to have the same size) satisfy the previous relationship.

On the other hand we write $x \leqslant y$ if $x_i \le y_i$ for every $i \in \mathbb{I}_d$. It is a standard exercise to show that $x \leqslant y \implies x^{\downarrow} \leqslant y^{\downarrow} \implies x \prec_w y$. Majorization is usually considered because of its relation with tracial inequalities for convex functions. Indeed, given $x,\, y \in \mathbb{R}^d$ and $f : I \to \mathbb{R}$ a convex function defined on an interval $I \subseteq \mathbb{R}$ such that $x,\, y \in I^d$, then (see for example [3]):

1. If one assumes that $x \prec y$, then $\operatorname{tr} f(x) \stackrel{\text{def}}{=} \sum_{i=1}^{d} f(x_i) \le \sum_{i=1}^{d} f(y_i) = \operatorname{tr} f(y)$.
2. If only $x \prec_w y$, but the map f is also increasing, then still $\operatorname{tr} f(x) \le \operatorname{tr} f(y)$.
3. If $x \prec_w y$ and f is a strictly convex function such that $\operatorname{tr} f(x) = \operatorname{tr} f(y)$ then there exists a permutation σ of $\mathbb{I}_d$ such that $y_i = x_{\sigma(i)}$ for $i \in \mathbb{I}_d$.

Remark 2.3. Majorization between vectors in $\mathbb{R}^d$ is intimately related with the class of doubly stochastic $d \times d$ matrices, denoted by $\mathrm{DS}(d)$. Recall that a $d \times d$ matrix $D \in \mathrm{DS}(d)$ if it has non-negative entries and each row sum and column sum equals 1.

It is well known (see [3]) that given $x\,,\, y \in \mathbb{R}^d$ then $x \prec y$ if and only if there exists $D \in \mathrm{DS}(d)$ such that $D\,y = x$. As a consequence of this fact we see that if $x_1\,,\, y_1 \in \mathbb{R}^r$ and $x_2\,,\, y_2 \in \mathbb{R}^s$ are such that $x_i \prec y_i\,,\ i = 1\,,\,2$, then $x = (x_1\,,\, x_2) \prec y = (y_1\,,\, y_2)$ in $\mathbb{R}^{r+s}$.

Indeed, if D_1 and D_2 are the doubly stochastic matrices corresponding the previous majorization relations then $D = D_1 \oplus D_2 \in \mathrm{DS}(r+s)$ is such that $D\,y = x$. △

Submajorization can be extended to the context of self-adjoint matrices as follows: given $S_1\,,\, S_2 \in \mathcal{H}(d)$ we say that S_1 is **submajorized** by S_2 , denoted $S_1 \prec_w S_2$, if $\lambda(S_1) \prec_w \lambda(S_2)$. If $S_1 \prec_w S_2$ and $\operatorname{tr} S_1 = \operatorname{tr} S_2$ we say that S_1 is **majorized** by S_2 and write $S_1 \prec S_2$. Thus, $S_1 \prec S_2$ if and only if $\lambda(S_1) \prec \lambda(S_2)$. Notice that (sub)majorization is a spectral relation between self-adjoint operators.

We end this section by recalling the following result, known as Lidskii's inequality (see [3, III.4]).

Theorem 2.4 (Lidskii's inequality)**.** *Let* $A,\, B \in \mathcal{H}(d)$. *Then* $\lambda(A) + \lambda^{\uparrow}(B) \prec \lambda(A+B)$. □

Lidskii's inequality plays an important role in our study of optimal frame completion problems. Moreover, the case of equality in Lidskii's inequality, i.e. when $(\lambda(A) + \lambda^{\uparrow}(B))^{\downarrow} = \lambda(A+B)$ for $A,\, B \in \mathcal{H}(d)$, plays a central role in this paper. We completely characterize such pair of matrices - that we call optimal matching matrices - in the Appendix.

2.4 Frames and optimal completions with prescribed parameters

In several applied situations it is desired to construct a sequence $\mathcal{F}$ in such a way that the frame operator of $\mathcal{F}$ is given by some $S \in \mathcal{M}_d(\mathbb{C})^+$ and the squared norms of the frame elements are prescribed by a sequence of positive numbers $\mathbf{a} = (\alpha_i)_{i\in\mathbb{I}_n} \in \mathbb{R}_{>0}^n$. That is, given a fixed $S \in \mathcal{M}_d(\mathbb{C})^+$ and $\mathbf{a} \in \mathbb{R}_{>0}^n$, we analyze the existence (and construction) of a sequence $\mathcal{F} = \{f_i\}_{i\in\mathbb{I}_n}$ such that $S_\mathcal{F} = S$ and $\|f_i\|^2 = \alpha_i$, for $i \in \mathbb{I}_n$. This is known as the classical frame design problem. It has been treated by several research groups (see for example [1, 7, 10, 13, 17, 18, 19, 26]). In what follows we recall a solution of the classical frame design problem in the finite dimensional setting, in the way that it is convenient for our analysis.

Proposition 2.5 ([1, 29])**.** Let $B \in \mathcal{M}_d(\mathbb{C})^+$ with $\lambda(B) \in \mathbb{R}_+^{d\,\downarrow}$ and let $\mathbf{a} = (\alpha_i)_{i\in\mathbb{I}_k} \in \mathbb{R}_{>0}^k$. Then there exists a sequence $\mathcal{G} = \{g_i\}_{i\in\mathbb{I}_k} \in \mathcal{H}^k$ with frame operator $S_\mathcal{G} = B$ and such that $\|g_i\|^2 = \alpha_i$ for every $i \in \mathbb{I}_k$ if and only if $\mathbf{a} \prec \lambda(B)$ (completing with zeros if $k \neq d$). □

Recently, researchers have made a step forward in the classical frame design problem and have asked about the structure of **optimal** frames with prescribed parameters. For example, consider the following problem posed in [20]: let $\mathcal{H} \cong \mathbb{C}^d$ and let $\mathcal{F}_0 = \{f_i\}_{i\in\mathbb{I}_{n_o}} \in \mathcal{H}^{n_o}$ be a fixed (finite) sequence of vectors. Consider a sequence $\mathbf{a} = (\alpha_i)_{i\in\mathbb{I}_k} \in \mathbb{R}_{>0}^k$ such that $\operatorname{rk} S_{\mathcal{F}_0} \geq d-k$ and denote by $n = n_o + k$. Then, with this fixed data, the problem is to construct a sequence

$$\mathcal{G} = \{f_i\}_{i=n_o+1}^n \in \mathcal{H}^k \quad \text{with} \quad \|f_{n_o+i}\|^2 = \alpha_i \quad \text{for} \quad 1 \leq i \leq k \ ,$$

such that the resulting completed sequence is a frame $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) = \{f_i\}_{i\in\mathbb{I}_n} \in \mathbf{F}(n\,,\,d)$ whose MSE $\operatorname{tr} S_\mathcal{F}^{-1}$ is minimal among all possible such completions.

Note that there are other possible ways to measure robustness (optimality) of the completed frame $\mathcal{F}$ as above. For example, we can consider optimal (minimizing) completions, with prescribed norms, for the Benedetto-Fickus' potential. In this case we search for a frame $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) = \{f_i\}_{i\in\mathbb{I}_n} \in \mathbf{F}(n\,,\,d)$, with $\|f_{n_o+i}\|^2 = \alpha_i$ for $1 \leq i \leq k$, and such that its frame potential $\operatorname{FP}(\mathcal{F}) = \operatorname{tr} S_\mathcal{F}^2$ is minimal among all possible such completions. Indeed, this problem has been considered before in the particular case in which $\mathcal{F}_0 = \emptyset$ in [2, 14, 21, 24, 30].

In this paper we shall consider the problems of optimal completion with prescribed norms, where optimality is measured with respect to general convex potentials (see Definition 2.1). In order to describe our main problem we first fix the notation that we shall use throughout the paper.

Definition 2.6. Let $\mathcal{F}_0 = \{f_i\}_{i\in\mathbb{I}_{n_o}} \in \mathcal{H}^{n_o}$ and $\mathbf{a} = (\alpha_i)_{i\in\mathbb{I}_k} \in \mathbb{R}_{>0}^k$ such that $d - \operatorname{rk} S_{\mathcal{F}_0} \leq k$. Define $n = n_o + k$. Then

1. In what follows we say that $(\mathcal{F}_0\,,\,\mathbf{a})$ are initial data for the completion problem (CP).
2. For these data we consider the sets

$$\mathcal{C}_\mathbf{a}(\mathcal{F}_0) = \left\{\, \{f_i\}_{i\in\mathbb{I}_n} \in \mathcal{H}^n : \{f_i\}_{i\in\mathbb{I}_{n_o}} = \mathcal{F}_0 \quad \text{and} \quad \|f_{n_o+i}\|^2 = \alpha_i \ \text{ for } \ i \in \mathbb{I}_k \right\} ,$$

$$\text{and} \quad \mathcal{SC}_\mathbf{a}(\mathcal{F}_0) = \{S_\mathcal{F} :\ \mathcal{F} \in \mathcal{C}_\mathbf{a}(\mathcal{F}_0)\} \subseteq \mathcal{M}_d(\mathbb{C})^+ \ .$$

When the initial data $(\mathcal{F}_0\,,\,\mathbf{a})$ are fixed, we shall use throughout the paper the notations

$$S_0 = S_{\mathcal{F}_0} \quad , \quad \lambda = \lambda(S_0) \qquad \text{and} \qquad n = n_o + k \ . \qquad \triangle$$

Problem: (Optimal completions with prescribed norms with respect to P_f) Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP and let $f : [0,\infty) \to \mathbb{R}$ be a strictly convex function. Construct all possible $\mathcal{F} \in \mathcal{C}_\mathbf{a}(\mathcal{F}_0)$ that are the minimizers of P_f in $\mathcal{C}_\mathbf{a}(\mathcal{F}_0)$. △

Our analysis of the completed frame $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G})$ will depend on $\mathcal{F}$ through $S_\mathcal{F}$. Hence, the following description of $\mathcal{SC}_\mathbf{a}(\mathcal{F}_0)$ plays a central role in our approach.

Proposition 2.7. *Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. Then*

$$\mathcal{SC}_{\mathbf{a}}(\mathcal{F}_0)=\big\{S\in\mathcal{M}_d(\mathbb{C})^+\,:\,S\geq S_{\mathcal{F}_0}\quad \textit{and}\quad \mathbf{a}\prec\lambda(S-S_{\mathcal{F}_0})\big\}\ .$$

Proof. Observe that if $\mathcal{F}=(\mathcal{F}_0\,,\,\mathcal{G})\in\mathcal{H}^n$ then $S_{\mathcal{F}}=S_{\mathcal{F}_0}+S_{\mathcal{G}}$. Denote by $S_0=S_{\mathcal{F}_0}$ and $B=S-S_0$, for $S\in\mathcal{M}_d(\mathbb{C})^+$. Applying Proposition 2.5 to the matrix B (which must be nonnegative if $S\in\mathcal{SC}_{\mathbf{a}}(\mathcal{F}_0)$), we get the equality of the sets. □

Remark 2.8 (Optimal completion problem with prescribed norms: the feasible case)**.** Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP . Denote by $S_0=S_{\mathcal{F}_0}$, $\lambda=\lambda(S_0)$ and $t=\operatorname{tr}\lambda+\operatorname{tr}\mathbf{a}$. In [33] we introduced the following set

$$U_t(S_0\,,\,m)=\{S_0+B:\ B\in\mathcal{M}_d(\mathbb{C})^+\,,\ \operatorname{rk}B\leq d-m\ ,\ \operatorname{tr}(S_0+B)\ =\ t\ \}\ ,$$

where $m=d-k$. In [33, Theorem 3.12] it is shown that there exist $\prec$-minimizers in $U_t(S_0\,,\,m)$. Indeed, there exists $\nu=\nu(\lambda\,,\,m)\in(\mathbb{R}^d_{\geq 0})^{\downarrow}$ - that can be effectively computed by simple algorithms - such that $S\in U_t(S_0\,,\,m)$ is a $\prec$-minimizer if and only if $\lambda(S)=\nu$.

We say that the completion problem for $(\mathcal{F}_0,\mathbf{a})$ is **feasible** if $\mu\overset{\text{def}}{=}\nu-\lambda$ satisfies that $\mathbf{a}\prec\mu$, where $\nu=\nu(\lambda,m)$ is as above. In this case for any S which is a $\prec$-minimizer in $U_t(S_0\,,\,m)$ it holds that $\lambda(S-S_0)=\mu^{\downarrow}$ and hence, by Proposition 2.7, we conclude that $S\in\mathcal{SC}_{\mathbf{a}}(\mathcal{F}_0)$. Moreover, Proposition 2.7 also shows that $\mathcal{SC}_{\mathbf{a}}(\mathcal{F}_0)\subseteq U_t(S_0\,,\,m)$ and therefore S is a $\prec$-minimizer in $\mathcal{SC}_{\mathbf{a}}(\mathcal{F}_0)$. In this case, as a consequence of the results in Section 2.3, any completion $\mathcal{F}\in\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ such that $S_{\mathcal{F}}=S$ is a minimizer of P_f for any convex function $f:[0,\infty)\to\mathbb{R}$. That is, in the feasible case we have structural solutions of the completion problem, in the sense that these solutions do not depend on the particular choice of convex potential considered.

Nevertheless, it is easy to construct examples in which the completion problem for $(\mathcal{F}_0\,,\,\mathbf{a})$ is not feasible. For example, consider the frame $\mathcal{F}_0\in\mathbf{F}(7\,,\,5)$ whose synthesis operator is

$$T^*_{\mathcal{F}_0}=\begin{bmatrix} 0.9202 & -0.7476 & -0.4674 & 0.9164 & 0.1621 & 0.3172 & -0.5815\\ 0.4556 & 0.0164 & 0.0636 & 1.0372 & -1.6172 & 0.3688 & 0.2559\\ -0.0885 & -0.3495 & -0.9103 & 0.3672 & -0.6706 & -0.9252 & 0.6281\\ 0.1380 & -0.4672 & -0.6228 & -0.1660 & 0.9419 & 1.0760 & 1.1687\\ 0.7082 & 0.2412 & -0.1579 & -1.8922 & -0.4026 & 0.1040 & 1.6648 \end{bmatrix}\ . \tag{6}$$

In this case $\lambda=\lambda(S_{\mathcal{F}_0})=(9\,,5\,,4\,,2\,,1)$ and $t_0=\operatorname{tr}S_{\mathcal{F}_0}=21$. Fix the data $n=9$ (hence $k=2$), $\mathbf{a}=(3.5\,,\,2)$ and notice that then $t=t_0+\operatorname{tr}\mathbf{a}=26.5$ and $m=d-k=3$. Then, according to the results in [33] we know that the optimal spectrum for $U_t(S_0\,,\,m)$ is $\nu_{\lambda\,,\,m}(26.5)=(9\,,5\,,4.25\,,4.25\,,4)$. Therefore, we have that $\nu-\lambda=\mu=(2.25\,,\,3.25)$ so that $\mathbf{a}\not\prec\mu$, that is the completion problem $(\mathcal{F}_0,\mathbf{a})$ is not feasible.

The structure of the optimal completions with these norms was not known, even for the MSE. In what follows we shall give a complete description of the optimal frame completions - with respect to an arbitrary convex potential - for this initial data (see Example 5.10). △

3 The spectrum of the minimizers of P_f on $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$

Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. Let $\mu\in\mathbb{R}^d_{\geq 0}$ be such that $\mathbf{a}\prec\mu$. We consider the set

$$\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)\ \overset{\text{def}}{=}\ \{\mathcal{F}=(\mathcal{F}_0\,,\,\mathcal{G})\in\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0):\ \lambda(S_1)=\mu^{\downarrow}\}\subseteq\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)\ .$$

Notice that if $\mathcal{F}=(\mathcal{F}_0,\mathcal{G})$ then $S_{\mathcal{F}}=S_{\mathcal{F}_0}+S_{\mathcal{G}}$. By Proposition 2.7 we get the following partition:

$$\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)=\bigsqcup_{\mu\in\Gamma_d(\mathbf{a})}\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)\qquad\text{where}\qquad \Gamma_d(\mathbf{a})\ \overset{\text{def}}{=}\ \{\mu\in(\mathbb{R}^d_{\geq 0})^{\uparrow}:\ \mathbf{a}\prec\mu\}\ . \tag{7}$$

Theorem 3.1. *Consider the previous notations and fix* $\mu = \mu^\uparrow \in \Gamma_d(\mathbf{a})$. *Then,*

1. *The set* $\Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)) \stackrel{def}{=} \{\lambda(S_\mathcal{F}):\ \mathcal{F}\in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)\}$ *is convex.*

2. *The vector* $\nu = (\lambda(S_{\mathcal{F}_0}) + \mu)^\downarrow$ *is a* $\prec$*-minimizer in* $\Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu))$.

3. *If* $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)$ *is such that* $\lambda(S_\mathcal{F}) = \nu$ *then* $S_{\mathcal{F}_0}$ *and* $S_\mathcal{G}$ *commute.*

Proof. 1. First notice that the set of all frame operators $S_\mathcal{G} \in \mathcal{M}_d(\mathbb{C})^+$ such that $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)$ is closed under unitary equivalence. Indeed, if $U \in \mathcal{U}(n)$, then $U\,S_\mathcal{G}\,U^*$ is the frame operator of the sequence $U\cdot\mathcal{G} = \{U f_i\}_{i=n_o+1}^n$. Denote by $\lambda = \lambda(S_{\mathcal{F}_0})$. Therefore, it is straightforward to check that

$$\Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)) = \{\lambda(C):\ C = A+B,\ A,\,B\in\mathcal{H}(n),\ \lambda(A) = \lambda \quad \text{and} \quad \lambda(B) = \mu\}\ .$$

By Klyachko's theory on the sum of hermitian matrices with a given spectra [25] , we conclude that the set $\Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu))$ is convex.

2. Since the set of all frame operators $S_\mathcal{G} \in \mathcal{M}_d(\mathbb{C})^+$ such that $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)$ is closed under unitary equivalence it is clear that $\nu \in \Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu))$. On the other hand, given $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)$, then Lidskii's inequality (see Theorem 2.4) states that the vector $\nu \prec \lambda(S_{\mathcal{F}_0} + S_\mathcal{G}) = \lambda(S_\mathcal{F})$. This establishes that ν is a $\prec$-minimizer in $\Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu))$.

3. This is a restatement of Theorem 6.4. $\square$

Remark 3.2. Consider the previous notations and fix $\mu = \mu^\uparrow \in \Gamma_d(\mathbf{a})$. Let $f : [0\,,\,\infty) \to [0\,,\,\infty)$ be a strictly convex function and let P_f be the convex potential induced by f. By the results described in Section 2.3 and Theorem 3.1 we see that, if $\lambda = \lambda(S_{\mathcal{F}_0})$ then

$$\mathcal{F} \in \operatorname{argmin}\{P_f(\mathcal{G}) : \mathcal{G} \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)\} \iff \lambda(S_\mathcal{F}) = (\lambda+\mu)^\downarrow = (\,\lambda^\downarrow + \mu^\uparrow\,)^\downarrow\ . \tag{8}$$

That is, if we consider the partition of $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ described in Eq. (7), then in each slice $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)$ the minimizers of the potential P_f are characterized by the spectral condition (8).

This shows that in order to search for global minimizers of P_f on $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ we can restrict our attention to the set

$$\mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0) \stackrel{\rm def}{=} \big\{\,\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0):\ \lambda(S_\mathcal{F}) = \big(\,\lambda(S_{\mathcal{F}_0}) + \lambda^\uparrow(S_\mathcal{G})\,\big)^\downarrow\,\big\}\ . \tag{9}$$

Indeed, Eqs. (7) and (8) show that if $\mathcal{F}$ is a minimizer of P_f in $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ then $\mathcal{F} \in \mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)$, i.e.

$$\operatorname{argmin}\{P_f(\mathcal{F}) : \mathcal{F} \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)\} \ = \ \operatorname{argmin}\{P_f(\mathcal{F}) : \mathcal{F} \in \mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)\}\ . \tag{10}$$

Since the potential $P_f(\mathcal{F})$ depends on $\mathcal{F}$ through the eigenvalues of $S_\mathcal{F}$ we introduce the set

$$\mathcal{S}(\mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)) \stackrel{\rm def}{=} \{S_\mathcal{F}:\ \mathcal{F} \in \mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)\} \subseteq \mathcal{M}_d(\mathbb{C})^+\ . \tag{11}$$

Finally, for any $\lambda \in \mathbb{R}_{\geq 0}^d$, in what follows we shall also consider the set

$$\Lambda_{\mathbf{a}}^{\rm op}(\lambda) \stackrel{\rm def}{=} \{\lambda^\downarrow + \mu : \mu \in \Gamma_d(\mathbf{a})\} = \{\lambda^\downarrow + \mu^\uparrow :\ \mu \in \mathbb{R}_{\geq 0}^d \quad \text{and} \quad \mathbf{a} \prec \mu\}\ . \qquad \triangle$$

Theorem 3.3. *Let* $(\mathcal{F}_0\,,\,\mathbf{a})$ *be initial data for the CP. Denote by* $\lambda = \lambda(S_{\mathcal{F}_0})$. *Then*

1. *The set* $\Lambda_{\mathbf{a}}^{\rm op}(\lambda)$ *is compact and convex.*

2. *The spectral picture* $\{\lambda(S_\mathcal{F}):\ \mathcal{F} \in \mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)\} = \{\nu^\downarrow:\ \nu \in \Lambda_{\mathbf{a}}^{\rm op}(\lambda)\}$.

3. *If $\mathcal{F} = (\mathcal{F}_0\,, \mathcal{G}) \in \mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)$, with $\lambda^\uparrow(S_\mathcal{G}) = \mu$, then there exists $\{v_i : i \in \mathbb{I}_d\}$ an ONB of eigenvectors for $S_{\mathcal{F}_0}\,, \lambda$ such that*

$$S_\mathcal{G} = \sum_{i\in\mathbb{I}_d} \mu_i \cdot v_i \otimes v_i \quad \text{ and } \quad S_\mathcal{F} = S_{\mathcal{F}_0} + S_\mathcal{G} = \sum_{i\in\mathbb{I}_d} (\lambda_i + \mu_i)\, v_i \otimes v_i \ . \tag{12}$$

Proof. 1. If $\nu\,, \gamma \in \Lambda_{\mathbf{a}}^{\rm op}(\lambda)$ then there exist $\mu\,, \rho \in \Gamma_d(\mathbf{a})$ such that $\nu = \lambda^\downarrow + \mu$, $\gamma = \lambda^\downarrow + \rho$. Note that $\Gamma_d(\mathbf{a})$ is convex. Hence, if $t \in (0,1)$, then $\mu_t \stackrel{\rm def}{=} t\,\mu + (1-t)\,\rho \in \Gamma_d(\mathbf{a})$ and

$$t\,\nu + (1-t)\,\gamma = \lambda^\downarrow + t\,\mu + (1-t)\,\rho = \lambda^\downarrow + \mu_t \in \Lambda_{\mathbf{a}}^{\rm op}(\lambda) \ .$$

Item 2. is an immediate consequence of the definitions of $\Lambda_{\mathbf{a}}^{\rm op}(\lambda)$ and $\mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)$.

3. Let $\mathcal{F} = (\mathcal{F}_0\,, \mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)^{\rm op}$. Then the frame operator $S_\mathcal{G}$ is an optimal matching matrix for $S_{\mathcal{F}_0}$ in the sense of Eq. 35 (see the Appendix). Hence, the existence of an ONB $\{v_i : \ i \in \mathbb{I}_d\}$ for $S_0\,, \lambda$ satisfying Eq. (12) follows from Theorem 6.8. □

Let $(\mathcal{F}_0\,, \mathbf{a})$ be initial data for the CP. Recall that $\Gamma_d(\mathbf{a}) = \{\mu \in (\mathbb{R}_{\geq 0}^d)^\uparrow : \ \mathbf{a} \prec \mu\}$. In what follows we use the following notation: if $f : [0,\infty) \to \mathbb{R}$ is a function we consider $F : \mathbb{R}_{\geq 0}^d \to \mathbb{R}$ given by $F(\gamma) = \sum_{i\in\mathbb{I}_d} f(\gamma_i)$, for $\gamma \in \mathbb{R}_{\geq 0}^d$.

Theorem 3.4. *Let $(\mathcal{F}_0\,, \mathbf{a})$ be initial data for the CP and let $f : [0\,, \infty) \to [0\,, \infty)$ be a strictly convex function. Then there exists a vector $\mu(\lambda\,, \mathbf{a}\,, f) = \mu = \mu^\uparrow \in \Gamma_d(\mathbf{a})$ such that:*

1. *$\mathcal{F} = (\mathcal{F}_0\,, \mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ is a global minimizer of $P_f \iff \mathcal{F} \in \mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)$ and $\lambda^\uparrow(S_\mathcal{G}) = \mu$.*
2. *If we let $\lambda = \lambda(S_{\mathcal{F}_0})$ then μ is uniquely determined by the conditions*

$$\mu \in \Gamma_d(\mathbf{a}) \quad \text{ and } \quad F(\lambda + \mu) = \min_{\gamma\in\Gamma_d(\mathbf{a})} F(\lambda + \gamma) = \min_{\nu\in\Lambda_{\mathbf{a}}^{\rm op}(\lambda)} F(\nu) \ . \tag{13}$$

3. *Moreover, μ also satisfies that*

$$0 < \mu_i = \mu_{i+1} \implies \lambda_i = \lambda_{i+1} \quad \text{ for every } \quad i \in \mathbb{I}_{d-1} \ . \tag{14}$$

Proof. Notice that the map $F : \mathbb{R}_{\geq 0}^d \to [0\,, \infty)$ is also strictly convex, and it is invariant under permutations of the variables. Moreover,

$$P_f(\mathcal{F}) = \operatorname{tr} f(S_\mathcal{F}) = F(\lambda(S_\mathcal{F})) \quad \text{ for every } \quad \mathcal{F} \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0) \ . \tag{15}$$

Since $\Lambda_{\mathbf{a}}^{\rm op}(\lambda)$ is compact and convex and F is strictly convex then every local minimizer of F on $\Lambda_{\mathbf{a}}^{\rm op}(\lambda)$ coincide with a unique global minimizer denoted by $\nu = \nu(\mathbf{a}\,, \lambda\,, f) \in \Lambda_{\mathbf{a}}^{\rm op}(\lambda)$. Define $\mu = \nu - \lambda$ and notice that, by construction of the set $\Lambda_{\mathbf{a}}^{\rm op}(\lambda)$, $\mu = \mu^\uparrow$ and $\mathbf{a} \prec \mu$.

Recall that given $\mathcal{F} = (\mathcal{F}_0\,, \mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ then a necessary condition for $\mathcal{F}$ to be a global minimizer of P_f on $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ is that $\mathcal{F} \in \mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)$ (see Remark 3.2). Hence, by item 2 in Theorem 3.3, the fact that F is permutation invariant and Eq. (15) we conclude that $\mathcal{F} \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ is a global minimizer of P_f on $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ if and only if

$$\mathcal{F} = (\mathcal{F}_0\,, \mathcal{G}) \in \mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0) \quad \text{ and } \quad \lambda(S_\mathcal{F}) = \big(\,\lambda + \lambda^\uparrow(S_\mathcal{G})\,\big)^\downarrow = \nu^\downarrow \ .$$

Denote by $\rho = \lambda^\uparrow(S_\mathcal{G})$. Then $\mathbf{a} \prec \rho = \rho^\uparrow$ and hence $\lambda + \rho \in \Lambda_{\mathbf{a}}^{\rm op}(\lambda)$ is a minimizer of $F \iff \lambda + \rho = \nu \iff \rho = \mu$.

Assume now that $0 < \mu_i = \mu_{i+1}$ but $\lambda_i > \lambda_{i+1}$ for some $i \in \mathbb{I}_{d-1}$. We denote by ρ the vector obtained from μ be replacing the i-th and $(i+1)$-th entries of μ by

$$\rho_i = \mu_i - \varepsilon \ \text{ and } \ \rho_{i+1} = \mu_{i+1} + \varepsilon \ , \ \text{ where } \ 0 < \varepsilon < \min\{\tfrac{\lambda_i - \lambda_{i+1}}{2} \ , \ \mu_i\} \ .$$

Although it is possible that $\rho \neq \rho^\uparrow$, the facts that $(\mu_i \, , \, \mu_{i+1}) \prec (\rho_i \, , \, \rho_{i+1})$ and $\mu_j = \rho_j$ for every $j \in (\mathbb{I}_d \setminus \{i \, , \, i+1\})$ imply, by Remark 2.3, that $\mu \prec \rho$ and hence $\mathbf{a} \prec \mu \prec \rho$. Using Proposition 2.7 and fixing an ONB for $S_{\mathcal{F}_0} \, , \, \lambda$, we deduce that there exists $\mathcal{F}' = (\mathcal{F}_0 \, , \, \mathcal{G}') \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ such that $\lambda(S_{\mathcal{G}'}) = \rho^\downarrow$ and $\lambda(S_{\mathcal{F}'}) = (\lambda + \rho)^\downarrow$. Recall that $\nu = \lambda + \mu$. Note that

$$\nu_i = \lambda_i + \mu_i > \lambda_i + \rho_i > \lambda_{i+1} + \rho_{i+1} > \lambda_{i+1} + \mu_{i+1} = \nu_{i+1} \ ,$$

while $\nu_j = \lambda_j + \mu_j = \lambda_j + \rho_j$ for every $j \in (\mathbb{I}_d \setminus \{i \, , \, i+1\})$. Then, by Remark 2.3, we conclude that $\lambda + \rho \prec \nu$ and $(\lambda + \rho)^\downarrow \neq \nu^\downarrow$. Hence, if f is strictly convex the previous facts imply that $P_f(\mathcal{F}') < F(\nu)$, which contradicts the minimality of ν of the first part of this proof. □

Corollary 3.5. *Let $\mathcal{F} = (\mathcal{F}_0 \, , \, \mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ be a global minimizer of P_f . Then*

1. *If $z \in \sigma(S_{\mathcal{G}}) \setminus \{0\}$ then there exists $w \in \sigma(S_0)$ such that* $\ker(S_{\mathcal{G}} - z) \subseteq \ker(S_0 - w)$.
2. *In particular, if P denotes a sub-projection of the spectral projection $P(z)$ of $S_{\mathcal{G}}$ onto its eigenspace* $\ker(S_{\mathcal{G}} - z)$*, then P and S_0 commute.*

Proof. By Remark 3.2, the P_f-minimality of $\mathcal{F} = (\mathcal{F}_0 \, , \, \mathcal{G})$ in $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ implies that $\mathcal{F} \in \mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)$. Then, by Theorem 3.3, there exists $\{v_i : i \in \mathbb{I}_d\}$ an ONB of eigenvectors for $S_0 \, , \, \lambda = \lambda(S_0)$ such that Eq. (12) holds. Denote by $S_1 = S_{\mathcal{G}} \, , \, \mu = \lambda^\uparrow(S_1)$ and fix $z \in \sigma(S_1) \setminus \{0\}$. Consider the indexes

$$m(z) = \min\{i \in \mathbb{I}_d : \mu_i = z\} \quad \text{ and } \quad M(z) = \max\{i \in \mathbb{I}_d : \mu_i = z\} \ .$$

By Eq. (14) in Theorem 3.4 we know that there exists $w \in \sigma(S_0)$ such that $\lambda_i = w$ for every $m(z) \leq i \leq M(z)$. Then, we can use Eq. (12) and deduce that

$$\ker(S_{\mathcal{G}} - z) = \text{span}\{v_i : \ m(z) \leq i \leq M(z)\} \subseteq \ker(S_0 - w) \ .$$

Therefore, any projection P as in item 2 must satisfy that $P \cdot S_0 = S_0 \cdot P = wP$. □

Remark 3.6 (First reduction of the optimal CP problem)**.** Let $(\mathcal{F}_0 \, , \, \mathbf{a})$ be initial data for the CP and let $f : [0, \infty) \to \mathbb{R}$ be a strictly convex function. Consider the compact convex set $\Lambda_{\mathbf{a}}^{\rm op}(\lambda) \subseteq \mathbb{R}_{\geq 0}^d$ and define the strictly convex function $F : \Lambda_{\mathbf{a}}^{\rm op}(\lambda) \to \mathbb{R}$. Therefore F is continuous and hence

$$\exists ! \ \ \text{argmin} \ \{F(x) : x \in \Lambda_{\mathbf{a}}^{\rm op}(\lambda)\} = \nu \tag{16}$$

Theorem 3.4 states that $\mu(\lambda \, , \, \mathbf{a} \, , \, f) = \nu - \lambda$. Thus, $\mathcal{F} = (\mathcal{F}_0 \, , \, \mathcal{G}) \in \mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)$ is an optimal completion with respect to P_f if and only if $\lambda(S_{\mathcal{G}}) = \nu - \lambda$. Thus, the minimization problem in Eq. (16) constitutes a reduction of the optimization problem, that in turn can be attacked with several numerical tools in concrete examples. Notice that $\Lambda_{\mathbf{a}}^{\rm op}(\lambda)$ has a natural and explicit description in terms of majorization, which is an algorithmic notion.

In the next sections we develop a different approach to the computation of minimizers of P_f in $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ (see Section 5.2). △

4 Local minimizers of $P_f(\cdot)$ on $\mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)$

In applied situations it is quite useful to understand the structure of local minimizers of objective functions. In our case, the study of local minimizers allows us to give a detailed description of the geometrical structure of global minimizers. We shall consider two different topologies on the set $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$. On the one hand, we consider the pseudo-metric d_S given by

$$d_S(\mathcal{F} \, , \, \mathcal{F}') = \|S_{\mathcal{F}} - S_{\mathcal{F}'}\| \ ,$$

where $\|\cdot\|$ denotes the spectral norm on $\mathcal{M}_d(\mathbb{C})$. On the other hand, we also consider the punctual metric d_P given by

$$d_P(\mathcal{F}\,,\,\mathcal{F}\,') = \|T_{\mathcal{F}} - T_{\mathcal{F}\,'}\|\ ,$$

where as before $\|\cdot\|$ denotes the spectral norm. It is clear that the topology induced by d_P is strictly stronger in the sense that: if $\mathcal{F}_n \xrightarrow[n]{d_P} \mathcal{F}$ then $\mathcal{F}_n \xrightarrow[n]{d_S} \mathcal{F}$, while the converse is false. Hence, d_S-local minimizers are also d_P-local minimizers.

Let $f : [0\,,\infty) \to [0\,,\infty)$ be a strictly convex function. Recall from Remark 3.2 that global minimizers of P_f on $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ actually lie in $\mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0)$. Therefore we shall focus our interest in the geometrical and spectral structure of local minimizers $\mathcal{F} \in \mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0)$ of P_f .

4.1 The d_S-local minimizers on $\mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0)$ are global minimizers

Remark 4.1. Let $\mathcal{B}_1 = \{u_j\}_{j\in\mathbb{I}_d}$ and $\mathcal{B}_2 = \{v_j\}_{j\in\mathbb{I}_d}$ be two ONB for $\mathbb{C}^d$. Then there exist continuous curves $w_j : [0,1] \to \mathbb{C}^d$ ($j \in \mathbb{I}_d$) such that $w_j(0) = v_j\,,\ w_j(1) = u_j$ and such that $\{w_j(t)\}_{i\in\mathbb{I}_d}$ is an ONB for $\mathbb{C}^d$ for every $t \in [0,1]$.

In fact, given the unitary matrix $U \in \mathcal{U}(d)$ such that $U\,v_j = u_j$ for every $j \in \mathbb{I}_d$, there exists a unique $X \in \mathcal{G}l(d)^+$ with $\|X\| \leq 2\,\pi$ such that $e^{i\,X} = U$. Hence the continuous curve $\gamma_U : [0,1] \to \mathcal{U}(d)$ given by $\gamma_U(t) = e^{i\,t\,X}$ joins $\gamma_U(0) = I$ with $\gamma_U(1) = U$. Thus, the continuous curves $w_j(t) = \gamma_U(t)\,v_j$ enjoy the mentioned properties.

Assume further that there exists $S \in \mathcal{M}_d(\mathbb{C})^+$ such that both $\mathcal{B}_1$ and $\mathcal{B}_2$ are ONB of eigenvectors for S , $\lambda = \lambda(S)$. Hence, by Eq. (1) we have that

$$S = \sum_{i\in\mathbb{I}_d} \lambda_i\ u_i \otimes u_i = \sum_{i\in\mathbb{I}_d} \lambda_i\ v_i \otimes v_i\ .$$

In this case, it is easy to see that the unitary $U \in \mathcal{U}(d)$ such that $U\,v_j = u_j$ for every $j \in \mathbb{I}_d$ should also satisfy that $S\,U = U\,S$ and that $\gamma_U(t)\,S = S\,\gamma_U(t)$ for every $t \in [0,1]$.

Then the continuous curves $w_i : [0,1] \to \mathbb{C}^d$ previously constructed also satisfy that the basis $\{w_i(t)\}_{i\in\mathbb{I}_d}$ is an ONB of eigenvectors for S, λ, for every $t \in [0,1]$. In other words, for every $t \in [0,1]$ we have the identity

$$S = \gamma_U(t)\,S\,\gamma_U(t)^* = \sum_{i\in\mathbb{I}_d} \lambda_i\ \gamma_U(t)\,v_i \otimes \gamma_U(t)\,v_i = \sum_{i\in\mathbb{I}_d} \lambda_i\ w_i(t) \otimes w_i(t)\ . \qquad \triangle$$

Theorem 4.2. *Let $(\mathcal{F}_0\,,\mathbf{a})$ be initial data for the CP and fix a strictly convex function $f : [0\,,\infty) \to [0\,,\infty)$. Then every d_S-local minimizer of P_f on $\mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0)$ must be a* **global** *minimizer.*

Proof. Let $\mathcal{F}\,' = (\mathcal{F}_0\,,\mathcal{G}\,')$ be a global minimizer of P_f in $\mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0)$, so that $\lambda^\uparrow(S_{\mathcal{G}\,'}) = \mu = \mu(\lambda\,,\mathbf{a}\,,f)$ the vector of Theorem 3.4. On the other hand take $\mathcal{F} = (\mathcal{F}_0\,,\mathcal{G}) \in \mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0)$ a d_S-local minimizer of P_f on $\mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0)$. We denote by $S_{\mathcal{F}_0} = S_0$, $\lambda(S_0) = \lambda$ and $\lambda^\uparrow(S_\mathcal{G}) = \rho$. Then $\mu\,,\rho \in \Gamma_d(\mathbf{a})$ and by Theorem 3.3 (applied to both $\mathcal{F}$ and $\mathcal{F}\,'$) there exist two ONB's $\{u_i :\ i \in \mathbb{I}_d\}$ and $\{v_i :\ i \in \mathbb{I}_d\}$ such that

$$\begin{aligned} S_0 &= \sum_{i\in\mathbb{I}_d} \lambda_i\ u_i \otimes u_i = \sum_{i\in\mathbb{I}_d} \lambda_i\ v_i \otimes v_i \quad , \quad S_{\mathcal{F}\,'} = \sum_{i\in\mathbb{I}_d} (\lambda_i + \mu_i)\ u_i \otimes u_i\ , \\ &\text{and} \quad S_{\mathcal{F}} = \sum_{i\in\mathbb{I}_d} (\lambda_i + \rho_i)\ v_i \otimes v_i\ . \end{aligned} \tag{17}$$

Therefore, by Remark 4.1, there exists a family of continuous curves $w_i : [0,1] \to \mathcal{H}$ such that $w_i(0) = v_i$ and $w_i(1) = u_i$ for every $i \in \mathbb{I}_d$ and such that $\{w_i(t) :\ i \in \mathbb{I}_d\}$ is an ONB for S_0 and λ

for every $t \in [0,1]$. Define the continuous curve $\mathfrak{s} : [0,1] \to \mathcal{M}_d(\mathbb{C})^+$ given by

$$\begin{aligned} \mathfrak{s}(t) &= \sum_{i\in\mathbb{I}_d} (\lambda_i + t \cdot \mu_i + (1-t)\cdot \rho_i)\ w_i(t) \otimes w_i(t) \\ &= S_0 + \sum_{i\in\mathbb{I}_d} (t\cdot \mu_i + (1-t)\cdot\rho_i)\ w_i(t)\otimes w_i(t) \quad \text{for every} \quad t\in[0,1] \ . \end{aligned}$$

It is clear that $\mathfrak{s}(0) = S_{\mathcal{F}}$ and $\mathfrak{s}(1) = S_{\mathcal{F}'}$. We claim that $\mathfrak{s}(t) \in \mathcal{S}(\mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0))$ for every $t \in [0,1]$. Indeed, notice that $(t\cdot\mu + (1-t)\cdot\rho)^\uparrow = t\cdot\mu + (1-t)\cdot\rho$ and therefore $\mathbf{a} \prec (t\cdot\mu + (1-t)\cdot\rho)$ for $t \in [0,1]$. Hence, there is a map $\mathfrak{s}_1 : [0,1] \to \mathcal{M}_d(\mathbb{C})^+$ such that for every $t \in [0,1]$

$$\mathfrak{s}(t) = S_0 + \mathfrak{s}_1(t) \quad , \quad \mathbf{a} \prec \lambda(\mathfrak{s}_1(t)) \quad \text{and} \quad \lambda(\mathfrak{s}(t)) = \big(\lambda(S_0) + \lambda^\uparrow(\mathfrak{s}_1(t)\,)\,\big)^\downarrow \ .$$

These last facts prove our claim. Notice that then $h(t) = P_f(\mathfrak{s}(t))\,,\ t \in [0,1]$ is a strictly convex function that has local minima at $t = 0$ and $t = 1$, i.e. h is constant. Then $P_f(\mathcal{F}) = h(0) = h(1) = P_f(\mathcal{F}')$, and $\mathcal{F}$ is another global minimizer. □

Remark 4.3 (On d_S-local minimizers in $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)$ and $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ and a lifting problem)**.** The previous result raises the question about the spectral structure of d_S-local minima of P_f on $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)$ or on $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$. Indeed, let $\mathbf{a} \prec \mu = \mu^\uparrow$ and consider $\mathcal{F}$ a d_S-local minimizer in $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)$. As shown in Theorem 3.1 the set $\Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu))$ is convex. Therefore $\lambda(t) = t\cdot\lambda(S_{\mathcal{F}}) + (1-t)\cdot(\lambda_0+\mu)^\downarrow \in \Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu))$ for $t\in[0,1]$ is a continuous curve.

Assume that we can lift the curve $\lambda(\cdot)$ to a curve in $\mathcal{S}(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu))$ i.e., assume that there exists a continuous curve

$$\mathfrak{s} : [0,1] \to \mathcal{S}(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)) \qquad \text{such that} \qquad \lambda(\mathfrak{s}(t)) = \lambda(t) \ . \tag{18}$$

Then, we could argue as in Theorem 4.2 above and conclude that $\lambda(S_{\mathcal{F}}) = (\lambda_0+\mu)^\downarrow$, which in turn would also imply that d_S-local minimizers of P_f on $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ are also global minimizers. Although we conjecture that the lifting problem of Eq. (18) has a solution, we are not able to show that such a solution exists at this time. △

4.2 A geometrical approach for d_P-local minimizers on $\mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0)$

In what follows we consider a geometrical approach to the study of d_P-local minimizers. Our results are based on a perturbation result for finite sequences of vectors, which follows from the work in [31]. In order to describe the general setting, we begin by considering some well known facts from differential geometry. In what follows we consider the unitary group of a complex and finite dimensional inner product space $\mathcal{R}$, denoted $\mathcal{U}(\mathcal{R})$, together with its natural differential geometric (Lie) structure.

Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. Fix $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) = \{f_i\}_{i=1}^n \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$, where $n = k + n_{\rm o}$, $\mathcal{R} = R(S_{\mathcal{G}}) = \operatorname{span}\{\mathcal{G}\} \subseteq \mathbb{C}^d$, and $\tau = \operatorname{tr}\,\mathbf{a} = \sum_{i=1}^k \alpha_i > 0$. Consider the real vector space

$$\mathcal{H}_d(\mathcal{R})^\tau = \{S \in \mathcal{H}(d) :\ R(S) \subseteq \mathcal{R}\,,\ \operatorname{tr}\, S = \tau\}\ , \tag{19}$$

the cone $L_d(\mathcal{R})^+_\tau = \mathcal{H}_d(\mathcal{R})^\tau \cap \mathcal{M}_d(\mathbb{C})^+$, and the affine manifold

$$S_{\mathcal{F}_0} + \mathcal{H}_d(\mathcal{R})^\tau = \{S_{\mathcal{F}_0} + S :\ S \in \mathcal{H}_d(\mathcal{R})^\tau\} \subseteq \mathcal{H}(d)\ .$$

We define the smooth (and so d_P-continuous) map

$$\Phi_{\mathcal{F}} : \mathcal{U}(\mathcal{R})^k \to \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0) \subseteq \mathcal{H}^n \quad \text{given by} \quad \Phi_{\mathcal{F}}(U_i)_{i=1}^k = \{f_i\}_{i=1}^{n_{\rm o}} \cup \{U_i f_{i+n_{\rm o}}\}_{i=1}^k\ . \tag{20}$$

Finally, we consider the smooth map $\Psi_{\mathcal{F}} : \mathcal{U}(\mathcal{R})^k \to S_{\mathcal{F}_0} + \mathcal{H}_d(\mathcal{R})^\tau$ given by

$$\Psi_{\mathcal{F}}(U_i)_{i=1}^k = S_{\mathcal{F}_0} + \sum_{i=n_\circ+1}^{n} U_i f_i \otimes U_i f_i = S_{\mathcal{F}'} \quad \text{where} \quad \mathcal{F}' = \Phi_{\mathcal{F}}(U_i)_{i=1}^k \ . \tag{21}$$

Let us denote by $I^k = (I, \dots, I) \in \mathcal{U}(\mathcal{R})^k$. It turns out that in several cases (indeed, in a generic case) the map $\Psi_{\mathcal{F}}$ is an open map (in $S_{\mathcal{F}_0} + \mathcal{H}_d(\mathcal{R})^\tau$) around $\Psi_{\mathcal{F}}(I^k) = S_{\mathcal{F}}$. In order to characterize this situation we introduce the following notion.

Definition 4.4. Given a sequence $\mathcal{G} = \{g_i\}_{i\in\mathbb{I}_n}$ in $\mathbb{C}^d$ we say that $\mathcal{G}$ is **irreducible** if it can not be partitioned into two mutually orthogonal subsequences. △

Remark 4.5. Let $(\mathcal{F}_0\,, \mathbf{a})$ be initial data for the CP. Fix $\mathcal{F} = (\mathcal{F}_0\,, \mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$. Denote by $n = k + n_\circ$ and $\mathcal{R} = R(S_{\mathcal{G}}) = \text{span}\{\mathcal{G}\} \subseteq \mathbb{C}^d$. Consider the map $\Psi_{\mathcal{F}} : \mathcal{U}(\mathcal{R})^k \to S_{\mathcal{F}_0} + \mathcal{H}_d(\mathcal{R})^\tau$ defined in Eq. (21).

In [31] we have characterized when the map $\Psi_{\mathcal{F}}$ is a submersion in terms of certain commutant. Indeed, let $L_d(\mathcal{R})$ denote the (non unital) $*$-subalgebra of $\mathcal{M}_d(\mathbb{C})$ that contains all $T \in \mathcal{M}_d(\mathbb{C})$ such that $T = P_{\mathcal{R}}\, T\, P_{\mathcal{R}}$, where $P_{\mathcal{R}}$ denotes the orthogonal projection onto $\mathcal{R}$. Then, an immediate application of [31, Theorem 4.2.1.] shows that $\Psi_{\mathcal{F}}$ is a submersion if and only if the local commutant

$$\mathcal{M}(\mathcal{G}) \stackrel{\text{def}}{=} \{f_i \otimes f_i : \ n_\circ + 1 \le i \le n\}' \cap L_d(\mathcal{R}) \tag{22}$$

is trivial, i.e. $\mathcal{M}(\mathcal{G}) = \mathbb{C} \cdot P_{\mathcal{R}}$. Equivalently, $\Psi_{\mathcal{F}}$ is a submersion iff any $A \in L_d(\mathcal{R})$ such that $A\, f_i = a_i\, f_i$ for some $a_i \in \mathbb{C}$, $n_\circ + 1 \le i \le n$, must be $A = a\, P_{\mathcal{R}}$ for some $a \in \mathbb{C}$. It is straightforward to show that this last condition on the family $\{f_i \otimes f_i : \ n_\circ + 1 \le i \le n\}$ is equivalent to the fact that the sequence $\mathcal{G}$ is irreducible, in the sense defined above. Thus, we have proved the following statement: △

Proposition 4.6. Let $(\mathcal{F}_0\,, \mathbf{a})$ be initial data for the CP. Fix $\mathcal{F} = (\mathcal{F}_0\,, \mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$. Denote by $n = k + n_\circ$ and $\mathcal{R} = R(S_{\mathcal{G}}) = \text{span}\{\mathcal{G}\} \subseteq \mathbb{C}^d$. Then the following statements are equivalent:

1. The map $\Psi_{\mathcal{F}}$ of Eq. (21) is a submersion at $I^k \in \mathcal{U}(\mathcal{R})^k$.
2. The sequence $\mathcal{G}$ is irreducible.

In this case, the image of $\Psi_{\mathcal{F}}$ contains an open neighborhood of $\Psi_{\mathcal{F}}(I^k) = S_{\mathcal{F}}$ in $S_{\mathcal{F}_0} + \mathcal{H}_d(\mathcal{R})^\tau$. Hence, $\Psi_{\mathcal{F}}$ admits a smooth local cross section ψ around $S_{\mathcal{F}}$ such that $\psi(S_{\mathcal{F}}) = I^k$. □

Next we state a convenient reformulation of Proposition 4.6, in terms of the distance d_P.

Corollary 4.7. *Let $(\mathcal{F}_0\,, \mathbf{a})$ be initial data for the CP. Consider the smooth map*

$$\mathcal{S} : \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0) \to S_{\mathcal{F}_0} + \mathcal{H}_d(\mathcal{R})^\tau \quad \textit{given by} \quad \mathcal{S}(\mathcal{F}) = S_{\mathcal{F}} = S_{\mathcal{F}_0} + S_{\mathcal{G}} \tag{23}$$

for every $\mathcal{F} = (\mathcal{F}_0\,, \mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$. If we assume that a point $\mathcal{F} = (\mathcal{F}_0\,, \mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ satisfies that the sequence $\mathcal{G}$ is **irreducible***, then*

1. *The image of $\mathcal{S}$ contains an open neighborhood of $S_{\mathcal{F}}$ in $S_{\mathcal{F}_0} + \mathcal{H}_d(\mathcal{R})^\tau$.*
2. *The map $\mathcal{S}$ has a d_P-continuous local cross section φ around $S_{\mathcal{F}}$ such that $\varphi(S_{\mathcal{F}}) = \mathcal{F}$.*

Proof. Just define the d_P-continuous local cross section $\varphi = \Phi_{\mathcal{F}} \circ \psi$, where ψ is the smooth local cross section for $\Psi_{\mathcal{F}}$ of Proposition 4.6 and $\Phi_{\mathcal{F}}$ is the map of Eq. (20). □

4.3 The d_P-local minimizers of P_f in $\mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)$ are frames for $\mathcal{H}$

Definition 4.8. Let $\mathcal{F} = \{f_i\}_{i\in\mathbb{I}_k} \subseteq \mathcal{H}^k$. A **partition of $\mathcal{F}$ into irreducible subsequences** is a family $\{\mathcal{F}_i\}_{i\in\mathbb{I}_p}$ given by a partition $\Pi = \{J_i\}_{i\in\mathbb{I}_p}$ of the index set $\mathbb{I}_k$ in such a way that each $\mathcal{F}_i = \{f_j\}_{j\in J_i}$ satisfies that:

- The subspaces $W_i = \operatorname{span}\{\mathcal{F}_i\}$ $(i \in \mathbb{I}_p)$ are mutually orthogonal.
- Each subfamily $\mathcal{F}_i$ $(i \in \mathbb{I}_p)$ is irreducible. △

Notice that any sequence $\mathcal{F} = \{f_i\}_{i\in\mathbb{I}_k} \subseteq \mathcal{H}^k$ has a unique such partition. To see this, consider the subspace $\mathcal{R} = \operatorname{span}\{\mathcal{F}\} \subseteq \mathbb{C}^d$ and the (non-unital) $*$-subalgebra $\mathcal{M}(\mathcal{F}) = \{f_i\otimes f_i : i \in \mathbb{I}_k\}' \cap L_d(\mathcal{R})$. If $\mathcal{F}$ is not irreducible, then $\mathcal{M}(\mathcal{F})$ contains a unique sequence of minimal orthogonal projections $\{Q_i\}_{i\in\mathbb{I}_p}$ such that $Q_i\,Q_j = 0$ for $i, j \in \mathbb{I}_p$ such that $i \neq j$ and $\sum_{i\in\mathbb{I}_p} Q_i = P_{\mathcal{R}}$. In this case, we have that

$$Q_i\, f_j = \varepsilon(i,j)\, f_j \quad \text{ for every } \; i \in \mathbb{I}_p \text{ and } \; j \in \mathbb{I}_k \ ,$$

where $\varepsilon(i,j) \in \{0,1\}$. Let $J_i = \{j \in \mathbb{I}_k : \ \varepsilon(i,j) = 1\}$ for $i \in \mathbb{I}_p$. Let $\Pi = \{J_i\}_{i\in\mathbb{I}_p}$. The fact that $\sum_{i\in\mathbb{I}_p} Q_i = P_{\mathcal{R}}$ implies that Π is a partition of $\mathbb{I}_p$. The fact that $\{Q_i\}_{i\in\mathbb{I}_p}$ is a family of mutually orthogonal projections imply that the subspaces $W_i = \operatorname{span}\{f_j : \ j \in J_i\} = R(Q_i)$ are mutually orthogonal, while the fact that each Q_i is a minimal projection in $\mathcal{M}(\mathcal{F})$ implies that each $\mathcal{F}_i = \{f_j\}_{j\in J_i}$ is irreducible. Then $\Pi = \{J_i\}_{i\in\mathbb{I}_p}$ has the desired properties.

Lemma 4.9. *Let $f : [0\,,\,\infty) \to [0\,,\,\infty)$ be a strictly convex function and let $\{a_i\}_{i\in\mathbb{I}_n} \in \mathbb{R}^n_{>0}$ for some $n \geq d$. If $\mathcal{F} = \{f_i\}_{i\in\mathbb{I}_n}$ is a d_P-local minimizer of P_f in the set*

$$\mathcal{B}(\mathbf{a}) = \{\mathcal{G} = \{g_i\}_{i\in\mathbb{I}_n} \in \mathcal{H}^n : \ \|g_i\|^2 = a_i\,,\ i \in \mathbb{I}_n\} \ ,$$

then $\mathcal{F}$ is a frame for $\mathcal{H}$.

Proof. Let $\Pi = \{J_i\}_{i\in\mathbb{I}_p}$ be a partition of $\mathbb{I}_n$ such that, if $\mathcal{F}_i = \{f_j\}_{j\in J_i}$ for $i \in \mathbb{I}_p$, then $\{\mathcal{F}_i\}_{i\in\mathbb{I}_p}$ is a partition of $\mathcal{F}$ into irreducible subsequences. Recall that in this case the subspaces $W_i \overset{\text{def}}{=} \operatorname{span}\{\mathcal{F}_i\}$ $(i \in \mathbb{I}_p)$ are mutually orthogonal. Hence, it is easy to see that each subfamily $\mathcal{F}_i$ is a d_P-local minimizer of P_f in the set

$$\{\{g_j\}_{j\in J_i} : \ g_j \in W_i\,,\ \|g_i\| = \|f_i\|\,,\ j \in J_i\}\,.$$

By [30, Corollary 3] and the properties of Π, each $\mathcal{F}_i$ is a c_i-tight frame for W_i, for some $c_i > 0$, $i \in \mathbb{I}_p$. Therefore

$$S_{\mathcal{F}} = \sum_{i\in\mathbb{I}_p} S_{\mathcal{F}_i} = \sum_{i\in\mathbb{I}_p} c_i\, P_{W_i} \ .$$

Notice that, in particular, $S_{\mathcal{F}}\, f_j = c_i\, f_j$ for every $j \in J_i$.

Assume now that $\mathcal{F}$ is not a frame for $\mathcal{H}$. Then, there exists $i \in \mathbb{I}_p$ and $q, s \in J_i$ such that $\langle f_q, f_s\rangle \neq 0$, because otherwise $\mathcal{F}$ would be a sequence of mutually orthogonal vectors and then, since $n \geq d$ then we would have $n = d$ and that $\operatorname{span}\mathcal{F} = \mathcal{H}$. In particular, for this choice of indexes we have that $a_s = \|f_s\|^2 < c_i$, because

$$c_i\, \|f_s\|^2 = \langle S_{\mathcal{F}}\, f_s\,,\, f_s\rangle \geq |\langle f_s\,,\, f_s\rangle|^2 + |\langle f_s\,,\, f_q\rangle|^2 = (\|f_s\|^2 + \frac{|\langle f_s\,,\, f_s\rangle|^2}{\|f_s\|^2})\, \|f_s\|^2 \ .$$

We are assuming that $\ker S_{\mathcal{F}} \neq \{0\}$. Hence there exists $g \in \ker S_{\mathcal{F}}$ with $\|g\| = \|f_s\|$. Let

$$f_s(t) = \cos(t)\cdot f_s + \sin(t)\cdot g \quad \text{ for every } \quad t \in [0,1] \ ,$$

so that $f_s(0) = f_s$ and $f_s(1) = g$. Notice that $\|f_s(t)\| = \|f_s\|$ for every $t \in [0,1]$. Let $\mathcal{F}(t)$ be the sequence obtained from $\mathcal{F}$ by replacing f_s by $f_s(t)$ and let $\mathfrak{s}(t)$ denote the frame operator of $\mathcal{F}(t)$, for each $t \in [0,1]$. Then

$$\mathfrak{s}(t) = [S_{\mathcal{F}} - (f_s \otimes f_s)] + f_s(t) \otimes f_s(t) \quad \text{for every} \quad t \in [0,1] \ .$$

The inequality $a_s = \|f_s\|^2 < c_i$ implies that $S_{\mathcal{F}} - (f_s \otimes f_s) \in \mathcal{M}_d(\mathbb{C})^+$ and also that $R(S_{\mathcal{F}} - (f_s \otimes f_s)) = R(S_{\mathcal{F}})$. Indeed, $S_{\mathcal{F}} - (f_s \otimes f_s) = [a_s^{-1}(c_i - a_s)] \cdot f_s \otimes f_s + S'$ with $S' \in \mathcal{M}_d(\mathbb{C})^+$; in this case $\lambda(S')$ is obtained from $\lambda(S_{\mathcal{F}})$ by setting one of the occurrences of c_i in $\lambda(S)$ equal to 0, and $f_s \in \ker S'$. Thus,

$$\mathfrak{s}(t) = S' + [a_s^{-1}(c_i - a_s) \cdot f_s \otimes f_s + f_s(t) \otimes f_s(t)] \quad \text{with} \quad f_s \, , \, f_s(t) \in \ker S' \ , \tag{24}$$

for every $t \in [0,1]$. Using again the inequality $a_s = \|f_s\|^2 < c_i$, let us define

$$\lambda(t) = \lambda([a_s^{-1}(c_i - a_s)] \cdot f_s \otimes f_s + f_s(t) \otimes f_s(t)) = (\lambda_1(t) \, , \, \lambda_2(t) \, , \, 0 \, , \, \ldots \, , \, 0) \in (\mathbb{R}^d_{\geq 0})^{\downarrow} \ .$$

Then $\lambda(0) = (c_i \, , 0 \, , \, \ldots \, , \, 0)$, $\lambda(1) = (c_i - a_s \, , \, a_s \, , \, 0 \, , \, \ldots \, , \, 0)^{\downarrow}$ and $\lambda_2(t) > 0$ for $t > 0$. Then there exists $t_0 \in (0,1)$ such that for $0 < t < t_0$, $\lambda_2(t) < \varepsilon$ for $\varepsilon > 0$ such that $\varepsilon < \min_{1 \leq j \leq p} c_j$ and $\varepsilon < \lambda_1(t) = (c_i - \lambda_2(t))$. By the previous remarks, it follows that $\lambda(\mathfrak{s}(t))$ is obtained from $\lambda(S_{\mathcal{F}})$ by replacing one occurrence of c_i by $\lambda_1(t)$ and one occurrence of 0 by $\lambda_2(t)$. Therefore, if $r = \operatorname{rk} S_{\mathcal{F}}$ then $\lambda_j(\mathfrak{s}(t)) \leq \lambda_j(S_{\mathcal{F}})$ for $1 \leq j \leq r$ and $\operatorname{tr} S_{\mathcal{F}} = \sum_{j=1}^{r+1} \lambda_j(\mathfrak{s}(t)) = \operatorname{tr} \mathfrak{s}(t)$ imply that $\lambda(\mathfrak{s}(t)) \prec \lambda(S_{\mathcal{F}})$ for $0 < t < t_0$.

These facts show that $\mathcal{F}(t)$ converges with respect to the d_P-metric as $t \to 0^+$, while $P_f(\mathcal{F}(t)) < P_f(\mathcal{F})$ for $t \in (0, t_0)$. This contradicts the assumption that $\mathcal{F}$ is a d_P-local minimum of P_f and thus we should have that $R(S_{\mathcal{F}}) = \mathcal{H}$, i.e. $\mathcal{F}$ is a frame. □

Theorem 4.10. *Let $(\mathcal{F}_0 \, , \mathbf{a})$ be initial data for the CP. Let $\mathcal{F} = (\mathcal{F}_0 \, , \mathcal{G}) \in \mathcal{C}^{\text{op}}_{\mathbf{a}}(\mathcal{F}_0)$ be a d_P-local minimizer of P_f on $\mathcal{C}^{\text{op}}_{\mathbf{a}}(\mathcal{F}_0)$, for some strictly convex function $f : [0 \, , \infty) \to [0 \, , \infty)$. Then $\mathcal{F}$ is a frame, i.e. $S = S_{\mathcal{F}} \in \mathcal{G}l(d)^+$.*

Proof. Denote by $S_0 = S_{\mathcal{F}_0}$, $\lambda(S_0) = \lambda = \lambda^{\downarrow}$, $S_1 = S_{\mathcal{G}}$ and $\lambda(S_1) = \mu^{\downarrow}$ for some $\mathbf{a} \prec \mu = \mu^{\uparrow}$. Since $\mathcal{F} = (\mathcal{F}_0 \, , \mathcal{G}) \in \mathcal{C}^{\text{op}}_{\mathbf{a}}(\mathcal{F}_0)$, by Theorem 3.3 there exists an ONB $\{v_i : i \in \mathbb{I}_d\}$ of eigenvectors for $S_0 \, , \lambda$ such that $S = S_0 + S_1 = \sum_{i \in \mathbb{I}_d} (\lambda_i + \mu_i) \, v_i \otimes v_i$. If $S \notin \mathcal{G}l(d)^+$, let

$$r = \max\{i \in \mathbb{I}_d : \lambda_i \neq 0\} < \min\{j \in \mathbb{I}_d : \mu_j \neq 0\} - 1 \ . \tag{25}$$

Then $\mathcal{H}_r = \operatorname{span}\{v_i : i > r\} = \ker S_0$, and S_1 acts on $\mathcal{H}_r$. The minimality of $\mathcal{F}$ in $\mathcal{C}^{\text{op}}_{\mathbf{a}}(\mathcal{F}_0)$ implies that $\mathcal{G}$ is a d_P-local minimizer of P_f in the set $(n = k + n_{\circ})$

$$\mathcal{B}_k(\mathcal{H}_r) \overset{\text{def}}{=} \{\mathcal{G} = \{g_i\}_{i \in \mathbb{I}_k} \in \mathcal{H}_r^k : \ \|g_i\|^2 = \alpha_i \, , \ i \in \mathbb{I}_k\} \ ,$$

because $\lambda(S_0 + S_{\mathcal{G}}) = (\lambda \, , \, \lambda(S_{\mathcal{G}}))^{\downarrow} \implies P_f(\mathcal{F}_0 \, , \mathcal{G}) = P_f(\mathcal{F}_0) + P_f(\mathcal{G})$ for every $\mathcal{G} \in \mathcal{B}_k(\mathcal{H}_r)$. By Lemma 4.9, we deduce that $S_1 \in \mathcal{G}l(\mathcal{H}_r)^+$, contradicting Eq. (25). □

5 On the structure of global minimizers of P_f on $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$

Let $(\mathcal{F}_0 \, , \mathbf{a})$ be initial data for the CP and let $f : [0, \infty) \to \mathbb{R}$ be a strictly convex function. In this section we obtain a description of the geometrical structure of global minimizers of P_f on $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$. We accomplish this by studying the structure of d_P-local minimizers of P_f in terms of perturbation results for the classical frame design problem. This geometrical structure of global minimizers allow us to obtain an finite step algorithm that produces a finite set (that does not depend on f) which completely describes the optimal frame completions $\mathcal{F} \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ for P_f.

5.1 Partitions into irreducible subsequences

From now on we shall fix a strictly convex function $f : [0\,,\,\infty) \to [0\,,\,\infty)$.

The goal of this section is the following Theorem on the spectral and geometrical structure of global minimizers of $P_f(\cdot)$ on $\mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)$. The proof is divided into some lemmas that we state after the main result. Recall that $\Gamma_d(\mathbf{a}) = \{\mu \in (\mathbb{R}_{\geq 0}^d)^\uparrow : \ \mathbf{a} \prec \mu\}$.

Theorem 5.1. *Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. Denote by $\lambda = \lambda(S_{\mathcal{F}_0})$. Then*

1. *There exists a vector $\mu = \mu(\lambda\,,\,\mathbf{a}\,,\,f) \in \Gamma_d(\mathbf{a})$ such that*

$$\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0) \text{ is a global minimizer of } P_f \iff \mathcal{F} \in \mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0) \text{ and } \lambda^\uparrow(S_{\mathcal{G}}) = \mu\ .$$

Assume now that $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G})$ is a global minimizer of P_f on $\mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)$. Let $\{\mathcal{G}_i\}_{i\in\mathbb{I}_p}$ be a partition of $\mathcal{G}$ into irreducible subfamilies, where $\mathcal{G}_i = \{f_j\}_{j\in J_i}$ for a partition $\{J_i\}_{i\in\mathbb{I}_p}$ of the set of indexes $\{i : \ 1 \leq i \leq k\}$. Then for each $i \in \mathbb{I}_p$

2. *The frame operators $S_{\mathcal{G}_i}$ and $S_{\mathcal{F}_0}$ commute.*
3. *There exists $c_i \in \mathbb{R}_{>0}$ such that $S_{\mathcal{F}}\, f_j = c_i\, f_j \quad$ for every $\quad j \in J_i$.*

Proof. Item 1 was shown in Theorem 3.4.

2. Assume now that $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G})$ is a global minimizer of P_f on $\mathcal{C}_{\mathbf{a}}^{\rm op}(\mathcal{F}_0)$. Then

$$S_{\mathcal{G}} = \bigoplus_{i\in\mathbb{I}_p} S_{\mathcal{G}_i} \implies \sigma(S_1) = \bigcup_{i\in\mathbb{I}_p} \sigma(S_{\mathcal{G}_i})\ .$$

Let $P(\alpha)$ (resp. $P_i(\alpha)$) denote the spectral projection of $S_{\mathcal{G}}$ (resp. $S_{\mathcal{G}_i}$) associated with $\alpha \in \sigma(S_{\mathcal{G}})$ (or 0 in case $\alpha \notin \sigma(S_{\mathcal{G}_i})$). Then, for every $i \in \mathbb{I}_p$ we have that

$$S_{\mathcal{G}_i} = \sum_{\alpha\in\sigma(S_{\mathcal{F}})} \alpha\ P_i(\alpha) \quad \text{ with } \quad \sum_{i\in\mathbb{I}_p} P_i(\alpha) = P(\alpha)\ , \quad \alpha \in \sigma(S_{\mathcal{G}})\ .$$

Thus, each $P_i(\alpha)$ is a sub-projection of $P(\alpha)$ for $i \in \mathbb{I}_p$. If we consider $\alpha \in \sigma(S_{\mathcal{G}})$, $\alpha \neq 0$, then Corollary 3.5 shows that $P_i(\alpha)$ commutes with $S_{\mathcal{F}_0}$, for every $i \in \mathbb{I}_p$. This last fact implies that $S_{\mathcal{G}_i}$ commutes with $S_{\mathcal{F}_0}$, for every $i \in \mathbb{I}_p$.

3. It is a consequence of item 2 and the following Remark and Lemmas. □

Remark 5.2. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP and let $t = \operatorname{tr} \mathbf{a}$. Denote by $S_0 = S_{\mathcal{F}_0}$ and $\lambda = \lambda(S_0)$. Consider the set

$$U_t(S_0\,,\,m) = \{S_0 + B : \ B \in \mathcal{M}_d(\mathbb{C})^+\,,\ \operatorname{rk} B \leq d - m\ ,\ \operatorname{tr}(S_0 + B)\ =\ t\ \}\ ,$$

where $n = k + n_{\rm o}$ and $m = d - k$. It is shown in [33, Theorem 3.12] that there exist $\prec$-minimizers in $U_t(S_0\,,\,m)$. Indeed, there exists $\nu = \nu(\lambda\,,\,m) \in (\mathbb{R}_{\geq 0}^d)^\downarrow$ such that $S \in U_t(S_0\,,\,m)$ is a $\prec$-minimizer if and only if $\lambda(S) = \nu$. In this case, there exist $c > 0$ and $\{v_i : \ i \in \mathbb{I}_d\}$, an ONB for S_0 and λ such that

1. $S - S_0 = \sum_{i=1}^d \rho_i \cdot v_i \otimes v_i$, where $\rho = \rho(\lambda\,,\,m) = \lambda(S - S_0)^\uparrow$;
2. $\nu = (\lambda + \rho^\uparrow)^\downarrow$ and $\lambda_i(S_0) + \rho_i = c$ whenever $\rho_i \neq 0$.

As a consequence of these facts we get $Sf = c\,f$ for every $f \in R(S - S_0)$. Moreover, if $S' \in U_t(S_0\,,\,m)$ is another matrix such that $\lambda(S' - S_0)^\uparrow = \rho$ and $S' - S_0 = \sum_{i=1}^d \rho_i\, w_i \otimes w_i$, where $\{w_i : \ i \in \mathbb{I}_d\}$ is some ONB for S_0 and λ, then $\lambda(S') = \nu$ and S' is a $\prec$-minimizer in $U_t(S_0\,,\,m)$.

Assume now that $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ is such that S_0 and $S_\mathcal{G}$ commute. Denote by

$$\mathcal{R} = R(S_\mathcal{G}) \quad , \quad \mu = \lambda^\uparrow(S_\mathcal{G}) \quad , \quad k' = \operatorname{rk} S_\mathcal{G} \quad , \quad m' = d - k' = \max\{i \in \mathbb{I}_d : \mu_i = 0\}$$

and $\tau = \operatorname{tr}\, \mathbf{a}$. Note that $\mathcal{R}$ reduces $S_{\mathcal{F}_0}$. Write $S_\mathcal{R} = S_{\mathcal{F}_0}|_\mathcal{R} \in L(\mathcal{R})^+_\tau$. We get the identity

$$S_{\mathcal{F}_0} + L_d(\mathcal{R})^+_\tau = S_{\mathcal{F}_0}|_{\mathcal{R}^\perp} \oplus \big(S_\mathcal{R} + L(\mathcal{R})^+_\tau\big)\ , \tag{26}$$

where $L_d(\mathcal{R})^+_\tau$ is the sapce defined in Eq. (19). If we identify $\mathcal{R}$ with $\mathbb{C}^{k'}$ we have that

$$S_\mathcal{R} + L(\mathcal{R})^+_\tau = U_s(S_\mathcal{R}\,,\,0) \subseteq \mathcal{M}_{k'}(\mathbb{C})\ ,$$

where $s = \tau + \operatorname{tr}\, S_\mathcal{R}$. By the previous comments there exists $S_\tau \in S_\mathcal{R} + L(\mathcal{R})^+_\tau$ such that $\lambda(S_\tau) = \nu(\lambda(S_\mathcal{R})\,,\,0) \in \mathbb{R}^{k'}_{\geq 0}$, which is a $\prec$-minimizer in $U_s(S_\mathcal{R}\,,\,0) = S_\mathcal{R} + L(\mathcal{R})^+_\tau$. As a consequence of Eq. (26) and Remark 2.3, we conclude that

$$S_1 \overset{\text{def}}{=} S_{\mathcal{F}_0}|_{\mathcal{R}^\perp} \oplus S_\tau \in S_{\mathcal{F}_0} + L_d(\mathcal{R})^+_\tau \text{ is a } \prec\text{-minimizer in } S_{\mathcal{F}_0} + L_d(\mathcal{R})^+_\tau\ .$$

Notice that $\lambda(S_1) = \big(\,\lambda(S_{\mathcal{F}_0}|_{\mathcal{R}^\perp})\,,\,\lambda(S_\tau)\,\big)^\downarrow \in \mathbb{R}^d$. Moreover, by items 1 and 2 above, we see that in this case there exists an ONB (for $\mathcal{R}$) $\{v_i\}_{i\in\mathbb{I}_{k'}}$ for $S_\mathcal{R}$ and $\lambda(S_\mathcal{R}) \in \mathbb{R}^{k'}$ such that

$$S_\tau - S_\mathcal{R} = \sum_{i\in\mathbb{I}_{k'}} \rho_i\, v_i \otimes v_i\ , \quad \text{where} \quad \rho = \lambda(S_\tau - S_\mathcal{R})^\uparrow \in \mathbb{R}^{k'}\ , \tag{27}$$

and there exists $c \in \mathbb{R}_{>0}$ such that $\lambda_i(S_\mathcal{R}) + \rho_i = c$ whenever $\rho_i \neq 0$. Hence, in this case we obtain that

$$S_1 f = c\ f \qquad \text{for every} \qquad f \in R(S_\tau - S_\mathcal{R}) \subseteq \mathcal{R}\ . \tag{28}$$

△

Lemma 5.3. *Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. Fix a subspace $\mathcal{R} \subseteq \mathbb{C}^d$ which reduces $S_{\mathcal{F}_0}$. Let $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ be a d_P-local minimizer of P_f on the set*

$$\big\{\,\mathcal{F}' = (\mathcal{F}_0\,,\,\mathcal{G}') \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)\ :\ R(S_{\mathcal{G}'}) \subseteq \mathcal{R}\,\big\}\ .$$

Assume further that $S_0 = S_{\mathcal{F}_0}$ and $S_\mathcal{G}$ commute and that the sequence $\mathcal{G}$ is **irreducible**. *Then*

1. *The frame operator $S_\mathcal{F}$ is a $\prec$-minimizer in $S_{\mathcal{F}_0} + L_d(\mathcal{R})^+_\tau$.*

2. *The subspace $\mathcal{R}$ is contained in a eigenspace of $S_\mathcal{F}$.*

In particular, there exists $c \in \mathbb{R}_{>0}$ such that $S_\mathcal{F}\, f_i = c\, f_i$, for $n_\circ + 1 \leq i \leq n$.

Proof. Let $k' = \operatorname{rk} S_\mathcal{G}$ and $m' = d - k'$. Since by hypothesis S_0 and $S_\mathcal{G}$ commute, arguing as in Remark 5.2 we conclude that there exists $S_1 = S_{\mathcal{F}_0}|_{\mathcal{R}^\perp} \oplus S_\tau \in S_{\mathcal{F}_0} + L_d(\mathcal{R})^+_\tau$ such that S_1 is a $\prec$-minimizer in $S_{\mathcal{F}_0} + L_d(\mathcal{R})^+_\tau$. On the other hand, $S_\mathcal{G}$ and $S_\mathcal{F}$ also commute so that there exists an ONB of $\mathbb{C}^d$ of eigenvectors of $S_\mathcal{F}$ and $S_\mathcal{G}$, denoted $\{v_i :\ i \in \mathbb{I}_d\}$, such that $\{v_i\}_{i\in\mathbb{I}_{k'}}$ is an ONB for $S_\mathcal{R} \overset{\text{def}}{=} S_0|_\mathcal{R} \in L(\mathcal{R})^+_\tau$ and $\lambda(S_\mathcal{R})$. In other words

$$S_\mathcal{F} = \sum_{i\in\mathbb{I}_d} \alpha_i \cdot v_i \otimes v_i \quad , \quad S_\mathcal{R} = \sum_{i\in\mathbb{I}_{k'}} \lambda_i(S_\mathcal{R}) \cdot v_i \otimes v_i \quad \text{and} \quad S_\mathcal{G} = \sum_{i\in\mathbb{I}_{k'}} \beta_i \cdot v_i \otimes v_i\ ,$$

for some $(\alpha_i)_{i=1}^d \in \mathbb{R}_{\geq 0}^d$ and $(\beta_i)_{i=1}^{k'} \in \mathbb{R}_{\geq 0}^{k'}$. Let $\rho = \lambda(S_\tau - S_\mathcal{R})^\uparrow \in \mathbb{R}^{k'}$ be as in Eq. (27) and consider the continuous curve $\mathfrak{s} : [0,1] \to S_{\mathcal{F}_0} + L_d(\mathcal{R})_\tau^+$ given by

$$\mathfrak{s}(x) = S_{\mathcal{F}_0} + \sum_{i\in\mathbb{I}_{k'}} [x \cdot \beta_i + (1-x)\cdot \rho_i] \cdot v_i \otimes v_i \quad \text{for} \quad x \in [0,1] \ .$$

First, notice that $\mathfrak{s}(x)$ is a segment (so, in particular, a continuous curve) joining $\mathfrak{s}(0) = S_1 = S_{\mathcal{F}_0}|_{\mathcal{R}^\perp} \oplus S_\tau$ and $\mathfrak{s}(1) = S_\mathcal{F}$. Consider now the map $h : [0,1] \to \mathbb{R}$ given by

$$\begin{aligned} h(x) &= \operatorname{tr} f\big(\mathfrak{s}(x)\big) = \textstyle\sum_{i\in\mathbb{I}_d} f(\lambda_i(\mathfrak{s}(x))) \\ &= \textstyle\sum_{i=k'+1}^d f(\alpha_i) + \sum_{i\in\mathbb{I}_{k'}} f\big(\lambda_i(S_\mathcal{R}) + x\cdot\beta_i + (1-x)\cdot\rho_i\big) \end{aligned}$$

for every $x \in [0,1]$. Since the sequence $\mathcal{G}$ is irreducible then Corollary 4.7, implies that the map $\mathcal{S} : \mathcal{C}_\mathbf{a}(\mathcal{F}_0) \to S_0 + L_d(\mathcal{R})_\tau^+$ defined in Eq. (23) has a d_P-continuous local cross section φ around $S_\mathcal{F}$ such that $\varphi(S_\mathcal{F}) = \mathcal{F}$. Then, the fact that $\mathcal{F}$ is a d_P-local minimizer of P_f implies that h has a local minimizer at $1 \in [0,1]$. But this h is a strictly convex function on $[0,1]$ that has a global minimum at $x = 0$, since $\mathfrak{s}(0)$ is a $\prec$-minimizer in $S_{\mathcal{F}_0} + L_d(\mathcal{R})_\tau^+$.

This implies that h is constant on $[0,1]$ and hence the segment $\lambda(\mathfrak{s}(x))$, $x \in [0,1]$, reduces to a point. Thus $\beta_i = \rho_i$ for every $i \in \mathbb{I}_{k'}$. Hence $S_\mathcal{G} = S_\tau - S_\mathcal{R}$ and $S_\mathcal{F} = S_{\mathcal{F}_0}|_{\mathcal{R}^\perp} \oplus S_\tau = S_1$. By Eq. (28) of Remark 5.2, there exists a $c \in \mathbb{R}_{\geq 0}$ such that $S_\mathcal{F}\, f_i = S_\tau\, f_i = c\, f_i$ for $n_\circ + 1 \leq i \leq n$ (since $f_i \in \mathcal{R} = R(S_\mathcal{G}) = R(S_\tau - S_\mathcal{R})$ for these indexes). This last fact proves item 2 of the statement. □

Lemma 5.4. *Let $(\mathcal{F}_0\,, \mathbf{a})$ be initial data for the CP. Let $\mathcal{F} = (\mathcal{F}_0\,, \mathcal{G}) \in \mathcal{C}_\mathbf{a}(\mathcal{F}_0)$ be a d_P-local minimizer of P_f on $\mathcal{C}_\mathbf{a}(\mathcal{F}_0)$. Let $\{\mathcal{G}_i\}_{i\in\mathbb{I}_p}$ be a partition of $\mathcal{G}$ into irreducible subfamilies, where $\mathcal{G}_i = \{f_j\}_{j\in J_i}$ for a partition $\{J_i\}_{i\in\mathbb{I}_p}$ of the set of indexes $\{i : \ n_\circ + 1 \leq i \leq n\}$. Assume that $S_{\mathcal{G}_i}$ and S_0 commute, for every $i \in \mathbb{I}_p$. Then there exist positive numbers*

$$c_1\,, \ldots, c_p \in \mathbb{R}_{>0} \quad \textit{such that} \quad S_\mathcal{F} f_j = c_i\, f_j \ , \quad j \in J_i \ , \ i \in \mathbb{I}_p \ .$$

Proof. Notice that, by construction, the ranges of the frame operators $S_{\mathcal{G}_i}$ and $S_{\mathcal{G}_j}$ are orthogonal whenever $i \neq j$. Fix $i \in \mathbb{I}_p$. The hypothesis allows us to apply Lemma 5.3 to the sequence $(\mathcal{F}_0\,, \mathcal{G}_i) \in \mathcal{C}_{\mathbf{a}_i}(\mathcal{F}_0)$, where $\mathbf{a}_i = (\|f_j\|^2)_{j\in J_i}$. In this case we conclude that there exists $c_i \in \mathbb{R}_{>0}$ such that $(S_{\mathcal{F}_0} + S_{\mathcal{G}_i})\, f_j = c_i\, f_j$, for every $j \in J_i$. Hence,

$$S_\mathcal{F}\, f_j = (S_{\mathcal{F}_0} + S_\mathcal{G})\, f_j = (S_{\mathcal{F}_0} + \bigoplus_{l\in\mathbb{I}_p} S_{\mathcal{G}_l})\, f_j = (S_{\mathcal{F}_0} + S_{\mathcal{G}_i})\, f_j = c_i\, f_j \ ,$$

for every $j \in J_i$. □

5.2 A finite step algorithm to compute global minimizers

In this section we obtain, as a consequence of Theorem 5.1, an algorithmic solution of the optimal frame completion problem with prescribed norms with respect to a general convex potential P_f. The key step is the introduction of the following finite set:

5.5. Let $(\mathcal{F}_0\,, \mathbf{a})$ be initial data for the CP. In order to find the minimizers for the CP with parameters $(\mathcal{F}_0\,, \mathbf{a})$ we construct a finite set $E(\mathcal{F}_0\,, \mathbf{a}) \subseteq (\mathbb{R}_{\geq 0}^d)^\uparrow$ as follows:

Set $r \in \mathbb{I}_d$. Consider a partition $\{K_i\}_{i\in\mathbb{I}_p}$ of the set $\{r, \ldots, d\}$ for some $1 \leq p \leq (d-r)+1$ and define the subsequences of $\lambda = \lambda(S_{\mathcal{F}_0})$ given by

$$\Lambda_i = \{\lambda_j\}_{j\in K_i} \in \mathbb{R}_{\geq 0}^{|K_i|} \ , \quad \text{for every} \quad i \in \mathbb{I}_p \ .$$

Consider also a partition $\{J_i\}_{i\in\mathbb{I}_p}$ of the set $\{1,\dots,k\}$ and define the subsequences of $\mathbf{a} = (\alpha_i)_{i=1}^k \in \mathbb{R}^k$ given by

$$\mathbf{a}_i = \{\alpha_j\}_{j\in J_i} \in \mathbb{R}_{>0}^{|J_i|} \ , \qquad \text{for every} \quad i \in \mathbb{I}_p \ .$$

For each $i \in \mathbb{I}_p$ define $c_i = |K_i|^{-1} \cdot (\text{tr } \Lambda_i + \text{tr } \mathbf{a}_i)$ and $\Gamma_i = \{c_i - \lambda_j\}_{j\in K_i}$. Let

$$\mu \in \mathbb{R}^d \quad \text{ be given by } \quad \mu_j = (\Gamma_i)_j = c_i - \lambda_j \quad \text{ if } \quad j \in K_i \ , \tag{29}$$

and $\mu_j = 0$ if $j < r$. We now check whether for every $i \in \mathbb{I}_p$ it holds that:

$$\Gamma_i \in \mathbb{R}_{\geq 0}^{|K_i|} \quad , \quad \mathbf{a}_i \prec \Gamma_i \quad \text{ and that } \quad \mu = \mu^\uparrow \in (\mathbb{R}_{\geq 0}^d)^\uparrow \ . \tag{30}$$

In this case we declare this μ as a member of $E(\mathcal{F}_0 \, , \, \mathbf{a})$. Otherwise we drop this μ. The set $E(\mathcal{F}_0 \, , \, \mathbf{a})$ is then obtained by this procedure, as we vary $1 \leq r \leq d$ and the partitions previously considered. Therefore, $E(\mathcal{F}_0 \, , \, \mathbf{a})$ is a finite set.

A straightforward computation using Proposition 2.5 and Eq. (30) shows that for every $\gamma \in E(\mathcal{F}_0 \, , \, \mathbf{a})$ there exists a completion $\mathcal{F}' = (\mathcal{F}_0 \, , \, \mathcal{G}') \in \mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)$ such that $\lambda^\uparrow(S_{\mathcal{G}'}) = \gamma$ and $\lambda(S_{\mathcal{F}'}) = (\lambda+\gamma)^\downarrow$. We remark that the set $E(\mathcal{F}_0 \, , \, \mathbf{a})$ can be explicitly computed in a finite step algorithm, in terms of $\lambda = \lambda(S_{\mathcal{F}_0})$ and $\mathbf{a}$ (see Section 5.3 below for details). △

Fix now a strictly convex function $f : [0 \, , \, \infty) \to [0 \, , \, \infty)$. Recall that we denote by $F : \mathbb{R}_{\geq 0}^d \to \mathbb{R}_{\geq 0}$ the map given by $F(\gamma) = \sum_{i\in\mathbb{I}_d} f(\gamma_i)$ for every $\gamma \in \mathbb{R}_{\geq 0}^d$.

Theorem 5.6. *Let $(\mathcal{F}_0 \, , \, \mathbf{a})$ be initial data for the CP and let $\lambda = \lambda(S_{\mathcal{F}_0})$. Then*

1. *The vector $\mu = \mu(\lambda \, , \, \mathbf{a} \, , \, f) \in (\mathbb{R}_{\geq 0}^d)^\uparrow$ of Theorem 5.1 satisfies that $\mu \in E(\mathcal{F}_0 \, , \, \mathbf{a})$.*
2. *Moreover, this vector μ is uniquely determined by the equation*

$$F(\lambda + \mu) = \min \, \{ F(\lambda+\gamma) : \gamma \in E(\mathcal{F}_0 \, , \, \mathbf{a}) \ \} \ . \tag{31}$$

That is, a completion $\mathcal{F} = (\mathcal{F}_0 \, , \, \mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ is a P_f global minimizer if and only if $\mathcal{F} \in \mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)$, $\mu = \lambda^\uparrow(S_{\mathcal{G}}) \in E(\mathcal{F}_0 \, , \, \mathbf{a})$ and it satisfies Eq. (31).

Proof. Denote by $\mu = \mu(\lambda \, , \, \mathbf{a} \, , \, f) \in (\mathbb{R}_{\geq 0}^d)^\uparrow$, the vector of Theorem 5.1. Let $\mathcal{F} = (\mathcal{F}_0 \, , \, \mathcal{G})$ be a global minimizer of P_f on $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$. In this case, by Theorem 3.1, $S_{\mathcal{F}_0}$ and $S_{\mathcal{G}}$ commute and $\lambda(S_{\mathcal{F}}) = (\lambda+\mu)^\downarrow$, where $\mu = \mu^\uparrow \in \mathbb{R}_{\geq 0}^d$ is such that $\lambda(S_{\mathcal{G}}) = \mu^\downarrow$.

Let $\{\mathcal{G}_i\}_{i\in\mathbb{I}_p}$ be a partition of $\mathcal{G}$ into irreducible subfamilies, corresponding to the partition $\{J_i\}_{i\in\mathbb{I}_p}$ of $\{n_\circ+1,\dots,n\}$, for some $1 \leq p \leq d$. Notice that in this case $S_{\mathcal{G}} = \oplus_{i\in\mathbb{I}_p} S_{\mathcal{G}_i}$. This last fact shows that there exists a partition $\{K_i\}_{i\in\mathbb{I}_p}$ such that $\lambda(S_{\mathcal{G}_i}) = (\Gamma_i \, , \, 0_i)$ where $\Gamma_i = \{\mu_j\}_{j\in K_i}$ and $0_i \in \mathbb{R}^{d-|K_i|}$ for every $i \in \mathbb{I}_p$. Then $\mu = (\oplus_{i\in\mathbb{I}_p}\Gamma_i)^\uparrow$.

Fix $i \in \mathbb{I}_p$. Theorem 5.1 implies that there exists $c_i > 0$ such that $S_{\mathcal{F}} f_j = c_i \, f_j$ for every $j \in J_i$ and $S_{\mathcal{G}_i}$ and that $S_{\mathcal{F}_0}$ commute. This fact implies that $S_{\mathcal{F}}|_{R_i} = c_i \, I_{R_i}$, where $R_i = R(S_{\mathcal{G}_i})$ and P_{R_i} denotes the identity operator on R_i . Therefore, we conclude that $c_i = \lambda_j + \mu_j$ for every $j \in K_i$. Hence $\Gamma_i = (c_i - \lambda_j)_{j\in K_i}$ and

$$c_i = |K_i|^{-1} \cdot \sum_{j\in K_i} (\lambda_j + \mu_j) = |K_i|^{-1} \cdot (\text{tr } \Lambda_i + \text{tr } \{\alpha_j\}_{j\in J_i}) \ ,$$

since $S_{\mathcal{G}_i} = \sum_{j\in J_i} f_j \otimes f_j$. This shows that $\text{tr } S_{\mathcal{G}_i} = \sum_{j\in J_i} \|f_j\|^2 = \sum_{j\in J_i} \alpha_j$. Moreover, the previous identity and Proposition 2.5 imply that $\mathbf{a}_i \prec \Gamma_i$, where $\mathbf{a}_i = \{\alpha_j\}_{j\in J_i}$. Hence, we conclude that the vector μ of Theorem 3.4 satisfies that $\mu \in E(\mathcal{F}_0 \, , \, \mathbf{a})$, as defined in 5.5.

As we mentioned before, for every $\gamma \in E(\mathcal{F}_0 \, , \, \mathbf{a})$ there exists a completion $\mathcal{F}' = (\mathcal{F}_0 \, , \, \mathcal{G}') \in \mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)$ such that $\lambda^\uparrow(S_{\mathcal{G}'}) = \gamma$ and $\lambda(S_{\mathcal{F}'}) = (\lambda+\gamma)^\downarrow$. Hence the vector μ satisfies Eq. (31). The converse implication now follows from item 1 and Theorem 3.4. □

Remark 5.7. Let $E(\mathcal{F}_0\,,\,\mathbf{a})\subseteq(\mathbb{R}^d_{\geq 0})^\uparrow$ be the finite set defined in 5.5 and assume that there exists $\mu\in E(\mathcal{F}_0\,,\,\mathbf{a})$ such that $\lambda+\mu$ is a $\prec$-minimizer for the set $\lambda+E(\mathcal{F}_0\,,\,\mathbf{a})$ i.e., such that

$$\lambda+\mu\prec\lambda+\gamma \quad \text{for every} \quad \gamma\in E(\mathcal{F}_0\,,\,\mathbf{a})\ . \tag{32}$$

Then, by Theorem 5.6 and the comments in Section 2.3 we see that μ coincides with $\mu(\lambda\,,\,\mathbf{a}\,,\,f)$, the vector of Theorem 5.1, for all strictly convex functions $f:[0\,,\,\infty)\to[0\,,\,\infty)$.

That is, given an arbitrary strictly convex functions $f:[0\,,\,\infty)\to[0\,,\,\infty)$ then a completion $\mathcal{F}=(\mathcal{F}_0\,,\,\mathcal{G})\in\mathcal{C}_\mathbf{a}(\mathcal{F}_0)$ is a global minimizer of P_f in $\mathcal{C}_\mathbf{a}(\mathcal{F}_0)$ if and only if $\lambda^\uparrow(S_\mathcal{G})=\mu$. Moreover, a similar argument shows that in this case

$$\lambda(S_{\mathcal{F}_0})+\mu\in\Lambda^{\rm op}_\mathbf{a}(\lambda(S_{\mathcal{F}_0})) \quad \text{is a } \prec\text{-minimizer in} \quad \Lambda^{\rm op}_\mathbf{a}(\lambda(S_{\mathcal{F}_0}))\ .$$

Therefore μ (resp. $\lambda(S_{\mathcal{F}_0})+\mu$) is an structural (spectral) solution to the problem of minimizing P_f , in the sense that the solution does not depend of the particular choice of the strictly convex function f.

Such structural solutions exist if we assume that the completion problem is feasible (see Remark 2.8). Numerical examples suggest that such a majorization minimizer always exists (see Section 5.3). These facts induce the following conjecture: △

Conjecture 5.8. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. Then there exists $\mu\in E(\mathcal{F}_0\,,\,\mathbf{a})$ such that $\lambda_{\mathcal{F}_0}+\mu$ satisfies the majorization minimality of Eq. (32). □

5.3 Algorithmic implementation: some examples and conjectures.

As it was described in the previous section, an algorithm can be developed in order to compute explicitly the set $E(\mathcal{F}_0\,,\,\mathbf{a})$ and the finite set of possible minimizers $\nu=\lambda+\mu$, $\mu\in E(\mathcal{F}_0\,,\,\mathbf{a})$ constructed from it. A proposed algorithm scheme is the following:

5.9. Given the initial data $\lambda\in(\mathbb{R}^d_{\geq 0})^\downarrow$ and $\mathbf{a}=(\alpha_i)_{i=1}^k$, we set $n=k+n_\circ$ as before.

Step 1. For each $r\in\mathbb{I}_r$ set $\lambda(r)=(\lambda_j)_{j=r}^d$. For such tail of λ, of length $l=d-r+1$, we consider the minimum $m=\{l,\ k\}$. Now, for each $p\in\mathbb{I}_m$,

- We compute all possible partitions of $\lambda(r)$ in p parts. We do the same with $\mathbf{a}$.
- Fixed a partition for $\lambda(r)$ and one of $\mathbf{a}$, we pair the sets of both partitions and compute for every pair the constant c and check majorization as it was described in Eq. (30).
- In case that the majorization conditions are satisfied for all pairs in these partitions for $\lambda(r)$ and $\mathbf{a}$, the vector μ is constructed as in Eq. (29).
- If $\mu=\mu^\uparrow$ then is μ stored in the set $E(\mathcal{F}_0\,,\,\mathbf{a})$.

Step 2. The set $N(\mathcal{F}_0\,,\,\mathbf{a})=\{\lambda+\mu:\mu\in E(\mathcal{F}_0\,,\,\mathbf{a})\}$ is constructed from that stored data.

Step 3. We search for the vector $\nu\in N(\mathcal{F}_0\,,\,\mathbf{a})$ of minimum euclidean norm.

Then this ν is a minimizer for the map $F(x)=\sum_{i\in\mathbb{I}_d}x_i^2$ associated to the frame potential on the set $\{\lambda(S_\mathcal{F}):\mathcal{F}\in\mathcal{C}_\mathbf{a}(\mathcal{F}_0)\}$. Moreover $\mu=\nu-\lambda$ is the vector of Theorem 3.4, which allows to construct (via the Schur-Horn algorithm) optimal completions in $\mathcal{C}^{\rm op}_\mathbf{a}(\mathcal{F}_0)$ with respect to the Benedetto-Fickus's frame potential. By Theorem 5.6, the globarl minimizers corresponding to a different potential in $\mathcal{C}^{\rm op}_\mathbf{a}(\mathcal{F}_0)$ can be computed similarly, i.e. by minimizing the corresponding convex function on the set $N(\mathcal{F}_0\,,\,\mathbf{a})$.

Step 4. Finally, we test if the vector ν obtained in Step 3 is a minimizer for majorization in $N(\mathcal{F}_0\,,\,\mathbf{a})$. In that case, the algorithm succeed in finding the minimizer for every convex potential P_f. △

In all examples in which we have applied the previous algorithm, the Step 4 confirmed that the minimizer for the frame potential in $N(\mathcal{F}_0\,,\,\mathbf{a})$ is actually the minimizer for majorization, which suggests a positive answer to the Conjecture 5.8 (see the comments in Remark 5.7).

Example 5.10. Consider the set of vectors $\mathcal{F}_0 \in \mathbf{F}(7\,,\,5)$ given in (6) and let $\mathbf{a} = \{3.5\,,\,2\}$ as it was pointed out in Subsection 2.4, with that initial data, the completion problem is not feasible. Nevertheless, if we apply the algorithm described above, the optimal spectrum μ and ν can be computed, since we can describe the set $N(\mathcal{F}_0\,,\,\mathbf{a})$.

Indeed in this case $N(\mathcal{F}_0\,,\,\mathbf{a}) = \{(9\,,5\,,4.5\,,4\,,4)\,,\ (9\,,6.5\,,5\,,4\,,2)\}$ so $\nu = (9\,,5\,,4.5\,,4\,,4)$ (where $\mu = (0\,,0\,,0\,,2\,,3.5)$) and an optimal completion is given by:

$$T^*_{\mathcal{F}_1} = \begin{bmatrix} 0.0441 & -1.3541 \\ 0.6901 & 0.5701 \\ -1.2093 & 0.0887 \\ -0.0569 & 0.8836 \\ 0.2371 & -0.7435 \end{bmatrix}. \tag{33}$$

In this case, the vector μ is constructed with the partitions $K_1 = \{2\}$, $K_2\{1\}$ of the two smaller eigenvalues in $\lambda = \lambda(S_{\mathcal{F}_0}) = (9\,,5\,,4\,,2\,,1)$ which are paired with $J_1 = \{2\}$ and $J_2 = \{3.5\}$ of $\mathbf{a}$, using the notation introduced in Section 5.2.

If we now set $\mathbf{a} = (2\,,\frac{1}{4}\,,\frac{1}{4}\,,\frac{1}{4})$, again the problem is not feasible (see [33]). In this case the algorithm yields a $N(\mathcal{F}_0\,,\,\mathbf{a})$ with 23 elements with a minimizer for majorization given by $\nu = (9\,,5\,,4\,,3\,,2.75)$. In this case, the partitions of λ are K_1 and K_2 of previous example, and $J_1 = \{\frac{1}{4}\,,\frac{1}{4}\,,\frac{1}{4}\}$ and $J_2 = \{2\}$ is the partition of $\mathbf{a}$. Finally, an optimal completion of $\mathcal{F}_0$ with prescribed norms is given by:

$$T^*_{\mathcal{F}_1} = \begin{bmatrix} 0.0156 & 0.0156 & 0.0156 & -1.0236 \\ 0.2440 & 0.2440 & 0.2440 & 0.4310 \\ -0.4275 & -0.4275 & -0.4275 & 0.0670 \\ -0.0201 & -0.0201 & -0.0201 & 0.6679 \\ 0.0838 & 0.0838 & 0.0838 & -0.5620 \end{bmatrix}. \tag{34}$$

Example 5.11. If $\mathbf{a} = (5.35\,,4.66\,,3.2\,,2.5\,,1.2\,,1\,,0.65)$ and let $\mathcal{F}_0$ be any family in $\mathbf{F}(n_\circ\,,\,6)$ such that $\lambda = \lambda(S_{\mathcal{F}_0}) = (5.75\,,5.4\,,4.25\,,4.25\,,3\,,2)$, (this is also a non-feasible example) then $N(\mathcal{F}_0\,,\,\mathbf{a})$ has 744 elements, and a minimizer is $\nu = (7.505\,,7.505\,,7.45\,,6.9167\,,6.9167\,,6.9167)$. In this example, the partitions for λ ($r_0 = 1$) and $\mathbf{a}$ involved in the computation of the optimal μ are $K_1 = \{5.75\,,5.4\,,4.25\}$, $K_2 = \{4.25\}$ and $K_3 = \{3\,,2\}$ and $J_1 = \{2.5\,,1.2\,,1\,,0.65\}$, $J_2 = \{3.2\}$ and $J_3 = \{5.35\,,4.66\}$ respectively. △

It is worth to note that the number of iterations done in Step 1 grows rapidly with d and k, and the size of $N(\mathcal{F}_0\,,\,\mathbf{a})$ also increases. As a consequence of these facts, the algorithm described in 5.9 is hard to implement for completion problems involving a large number of prescribed norms or for completion problems in $\mathbb{C}^d$ for large d. Nevertheless, in the previous examples (and several others considered for this work) it turned out (besides the fact that Conjecture 5.8 is verified in all examples) that the index-partition of λ and $\mathbf{a}$ in the $\prec$-minimizer consist of sets of **consecutive** elements, both for λ and $\mathbf{a}$. Moreover, in all examples the partitions are paired in such a way that the partitions with the greater elements of λ corresponds to those of $\mathbf{a}$ with the smaller entries (see the description of Λ_i and J_i in previous examples). Moreover, in all examples considered, the minimizer has the property that the sets of vectors corresponding to the partitions with the greater norms of $\mathbf{a}$ are linearly independent, with the exception of the last partition of $\mathbf{a}$. This structure is consistent with the solution for the classical completion problem with $\mathcal{F}_0 = \emptyset$ (see [2, 14, 30]).

This allows to develop a faster algorithm which tests a smaller set of partitions for λ and $\mathbf{a}$ which reduces considerably the time of computation and data storage. Thus, our numerical computations lead to the following Conjecture for the construction of the $\prec$-minimizer:

Conjecture 5.12. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP, and assume that $\mathbf{a}$ is arranged in decreasing order. Then, using the notations of 5.5, the minimizing vector $\mu \in E(\mathcal{F}_0\,,\,\mathbf{a})$ of Theorem 5.6 satisfies that:

1. It is constructed from consecutive partitions of λ and $\mathbf{a}$. In other words, that each set J_j and K_j in the partitions $\{J_j\}_{j\in\mathbb{I}_p}$ and $\{K_j\}_{j\in\mathbb{I}_p}$ given in 5.5 describing μ, consists of consecutive indexes.

2. The partitions of λ and $\mathbf{a}$ are paired in opposite order: the sets in the partition of λ with the larger elements are compared with those sets in the correspondent partition of $\mathbf{a}$ with the smaller elements. Moreover, the correspondent sets in both partitions have the same number of elements, except possibly the sets with the smallest and greatest entries of $\mathbf{a}$ and λ respectively. More explicitly, there exists $1 \le r_0 \le d$ such that $m = d - r_0 + 1 \le k$ and a sequence $r_0 \le r_1 < \ldots < r_p = d$ such that:

$$K_j = \{r_{j-1}+1\,,\,\ldots\,,\,r_j\}\ ,\quad J_j = \{d - r_j + 1\,,\,\ldots\,,\,d - r_{j-1}\}\ ,\qquad \text{for} \quad 2 \le j \le p\ ,$$

$$K_1 = \{r_0\,,\,\ldots\,,\,r_1\}\ ,\ J_1 = \{d - r_1 + 1\,,\,\ldots\,,\,k\}\ ,$$

and such that μ is constructed as in 5.5 in terms of $\{K_j\}_{j\in\mathbb{I}_p}$ and $\{J_j\}_{j\in\mathbb{I}_p}$. △

In the following example we verify that the algorithm implemented following the scheme in 5.9 and the simplified (and faster) version of this algorithm that assumes that Conjecture 5.12 holds, produce the same solution to the optimal completion problem with respect to the Benedetto-Fickus' frame potential.

Example 5.13. Given the initial data

$$\lambda = \lambda(S_{\mathcal{F}_0}) = (7\,,\,6\,,\,5.5\,,\,4\,,\,2.5\,,\,1\,,\,0.5\,,\,0.3) \quad \text{and} \quad \mathbf{a} = (5\,,\,4.5\,,\,1.2\,,\,1\,,\,0.8\,,\,0.5)\ ,$$

then applying the algorithm described in 5.9 we obtain that the optimal completion with prescribed norms $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G})$ has eigenvalues $\nu = (7\,,\,6\,,\,5.5\,,\,5.3\,,\,5\,,\,4\,,\,3.5\,,\,3.5)$. If we only check the partitions described in Conjecture 5.12, then we obtain the same optimal eigenvalues ν, with the partitions $J_1 = \{1.2\,,\,1\,,\,0.8\,,\,0.5\}$, $J_2 = \{4.5\}$, $J_3 = \{5\}$ and $K_1 = \{2.5\,,\,1\}$, $K_2 = \{0.5\}$, $K_3 = \{0.3\}$ for $\mathbf{a}$ and λ respectively ($r_0 = 5$). But there are only 5 cases constructed from this kind of partitions in a set $N(\mathcal{F}_0\,,\,\mathbf{a})$ with 322 elements. △

6 Appendix: Equality in Lidskii's inequality

Fix $S_0 \in \mathcal{M}_d(\mathbb{C})^+$. In this section we characterize those matrices

$$S_1 \in \mathcal{M}_d(\mathbb{C})^+ \quad \text{such that} \quad \lambda(S_0 + S_1) = \Big(\lambda^{\downarrow}(S_0) + \lambda^{\uparrow}(S_1)\Big)^{\downarrow}\ . \tag{35}$$

If $S_1 \in \mathcal{M}_d(\mathbb{C})^+$ satisfies Eq. (35) then we say that S_1 is an **optimal matching matrix** for S_0. Note that optimal matching matrices correspond to the cases of equality in Lidskii's inequality, as stated in Theorem 2.4.

Although we have defined this notion for positive matrices (since we interested in its application to frame operators) similar definitions and conclusions holds for general hermitian matrices (by translations by convenient multiples of the identity).

6.1 Optimal matching matrices commute

In this section we study the case of equality in Lidskii's inequality and show that if S_1 is an optimal matching for S_0 (i.e. S_1 is as in Eq. (35)) then $S_0\,S_1 = S_1\,S_0$.

We begin by revisiting some classical matrix analysis results. We shall give short proofs of them in order to handle these proofs for the equality cases in which we are interested here.

Lemma 6.1 (Weyl's inequalities)**.** *Let $A,\ B \in \mathcal{H}(d)$. Then,*

$$\lambda_j(A+B) \le \lambda_i(A) + \lambda_{j-i+1}(B) \quad \textit{for} \quad i \le j\,, \tag{36}$$

$$\lambda_j(A+B) \ge \lambda_i(A) + \lambda_{j-i+d}(B) \quad \textit{for} \quad i \ge j\,. \tag{37}$$

Moreover, if there exists $i \le j$ (resp. $i \ge j$) such that

$$\lambda_j(A+B) = \lambda_i(A) + \lambda_{j-i+1}(B) \tag{38}$$

(resp. $\lambda_j(A+B) = \lambda_i(A) + \lambda_{j-i+d}\,(B)$) then there exists a unit vector x such that

$$(A+B)\,x = \lambda_j(A+B)\,x\,,\ \ A\,x = \lambda_i(A)\,x\,,\ \ B\,x = \lambda_{j-i+1}(B)\,x\ ,$$

(resp. $(A+B)\,x = \lambda_j(A+B)\,x\,,\ A\,x = \lambda_i(A)\,x\,,\ B\,x = \lambda_{j-i+d}\,(B)\,x$).

Proof. We begin by proving (36). Let u_j , v_j and w_j denote the eigenvectors of A, B and $A+B$ respectively, corresponding to their eigenvalues arranged in decreasing order. Let $i \le j$ and consider the three subspaces spanned by the sets $\{w_1, \ldots, w_j\}$, $\{u_i, \ldots, u_n\}$ and $\{v_{j-i+1}, \ldots, v_n\}$. Since the dimensions of these subspaces are j, $n-i+1$ and $n-j+i$ respectively, we see that they have a non trivial intersection. If x is a unit vector in the intersection of these subspaces then

$$\lambda_j(A+B) \le \langle\,(A+B)\,x\,,\,x\rangle = \langle A\,x\,,\,x\rangle + \langle B\,x\,,\,x\rangle \le \lambda_i(A) + \lambda_{j-i+1}(B)\ .$$

If we further assume that equality (38) holds for these indexes then we deduce that

$$\langle\,(A+B)x\,,\,x\rangle = \lambda_j(A+B) \quad , \quad \langle A\,x\,,\,x\rangle = \lambda_i(A) \quad \text{and} \quad \langle B\,x\,,\,x\rangle = \lambda_{j-i+1}(B)\ .$$

Because x lies in the intersection of the previous subspaces, these last facts imply that $(A+B)\,x = \lambda_j(A+B)\,x$, $A\,x = \lambda_i(A)\,x$ and $\langle B\,x, x\rangle = \lambda_{j-i+1}(B)\,x$. The inequality (37) and the equality (38) for the case $i \ge j$ follow similarly. □

Corollary 6.2 (Weyl's monotonicity principle)**.** *Let $A \in \mathcal{H}(d)$ and $B \in \mathcal{M}_d(\mathbb{C})^+$. Then*

$$\lambda_j(A+B) \ge \lambda_j(A) \qquad \textit{for every} \quad j \in \mathbb{I}_d\ . \tag{39}$$

If there exists $J \subseteq \mathbb{I}_d$ such that $\lambda_j(A+B) = \lambda_j(A)$ for every $j \in J$, then there exists an orthonormal system $\{x_j\}_{j\in J}$ such that $A\,x_j = \lambda_j(A)\,x_j$ and $B\,x_j = 0$ for every $j \in J$.

Proof. Inequality (39) follows easily from Lemma 6.1 (with $i = j$). The second part follows by induction on the set $|J|$: Fix $j_0 \in J$. By Eq. (37) with $i = j = j_0$, there exists a unit vector x_{j_0} such that $A\,x_{j_0} = \lambda_{j_0}(A)\,x_{j_0}$ and $B\,x_{j_0} = \lambda_d(B)\,x_{j_0} = 0$.

This proves the case $|J| = 1$. If $|J| > 1$, consider the space $W = \{x_{j_0}\}^\perp \subseteq \mathbb{C}^d$ which reduces A, B and $A+B$. Let $I = \{j :\ j \in J\,,\ j < j_0\} \cup \{j-1 :\ j \in J\,,\ j > j_0\}$. The operators $A|_W \in L(W)^{\rm sa}$ and $B|_W \in L(W)^+$ satisfy that $\lambda_j(A|_W + B|_W) = \lambda_j(A|_W)$ for every $j \in I$, with $|I| = |J| - 1$. By the inductive hypothesis we can find an orthonormal system $\{x_j\}_{j\in I} \subseteq W$ which satisfy the desidered properties. □

Proposition 6.3. *Let A, $B \in \mathcal{H}(d)$. Then the equality*

$$\big(\lambda(A+B)-\lambda(A)\big)^{\downarrow}=\lambda(B) \implies \quad A \text{ and } B \text{ commute} .$$

Proof. We can assume that B is not a multiple of the identity. By hypothesis, there exists permutation $\sigma \in \mathbb{S}_d$ such that $\lambda_j(B)=\lambda_{\sigma(j)}(A+B)-\lambda_{\sigma(j)}(A)$ for every $j\in\mathbb{I}_d$. Therefore, there exists an increasing sequence $\{J_k\}_{k=1}^d$ of subsets of $\mathbb{I}_d$ such that $|J_k|=k$ and

$$\sum_{j\in J_k}\lambda_j(A+B)-\lambda_j(A)=\sum_{j=1}^{k}\lambda_j(B) \quad \text{for every} \quad k\in\mathbb{I}_d . \tag{40}$$

Let $k\in\mathbb{I}_d$ be such that $\lambda_{k-1}(B)>\lambda_k(B)$ (recall that $B\neq\alpha I$ for $\alpha\in\mathbb{R}$). Let us denote by $B_k=B-\lambda_k(B)\,I$ and notice Eq. (40) also holds if we replace B by B_k.

By construction $\lambda_k(B_k)=0$ and the orthogonal projection onto the kernel of the positive part $B_k^+\in\mathcal{M}_d(\mathbb{C})^+$ coincides with the spectral projection of the B associated to the interval $(-\infty,\lambda_k(B)]$. Moreover, $\dim\ker B_k^+=d-k+1$.

Since $B_k^+\in\mathcal{M}_d(\mathbb{C})^+$ and $B_k\leq B_k^+$ then Weyl's monotonicity principle implies that

$$\lambda_j(A+B_k)\leq\lambda_j(A+B_k^+)\,,\ j\in\mathbb{I}_d \implies \sum_{j\in J_{k-1}}\lambda_j(A+B_k)\leq\sum_{j\in J_{k-1}}\lambda_j(A+B_k^+)\ .$$

Therefore

$$\begin{aligned}\sum_{j\in J_{k-1}}\lambda_j(A+B_k)-\lambda_j(A) &\leq \sum_{j\in J_{k-1}}\lambda_j(A+B_k^+)-\lambda_j(A)\\ &\leq \sum_{j\in\mathbb{I}_d}\lambda_j(A+B_k^+)-\lambda_j(A)\\ &= \operatorname{tr}(A+B_k^+)-\operatorname{tr}A=\sum_{j=1}^{k-1}\lambda_j(B_k)\end{aligned}$$

since $\lambda_j(A+B_k^+)\geq\lambda_j(A)$ for $j\in\mathbb{I}_d$ - again by Weyl's monotonicity principle - and since, by hypothesis, $\lambda_k(B_k)=0$. The inequalities above are the key part of the proof of Lidskii's Theorem 2.4 ($\lambda(A+B)-\lambda(A)\prec\lambda(B)$). But here they actually equalities, by Eq. (40).

Let $J_{k-1}^c=\mathbb{I}_d\setminus J_{k-1}$. Then, from the above equalities we get that $\lambda_j(A+B_k^+)=\lambda_j(A)$ for every $j\in J_{k-1}^c$. By Corollary 6.2 there exists an ONS $\{x_j\}_{j\in J_{k-1}^c}$ such that $A\,x_j=\lambda_j(A)\,x_j$ and $B_k^+\,x_j=0$ for every $j\in J_{k-1}^c$. All these facts together imply that

$$P_k \stackrel{\text{def}}{=} \sum_{j\in J_{k-1}^c} x_j\otimes x_j=P_{\ker B_k^+} \quad \text{and} \quad P_k\,A=A\,P_k\ .$$

Recall that P_k is also the spectral projection of B associated to the interval $(-\infty,\lambda_k(B)]$, for any $k\in\mathbb{I}_d$ such that $\lambda_{k-1}(B)>\lambda_k(B)$. Since the spectral projection of B associated with $(-\infty,\lambda_1(B)]$ equals the identity operator, and B is a linear combination of the projections P_k and I, we conclude that A and B commute. □

Now we are ready to prove that if $S_1\in\mathcal{M}_d(\mathbb{C})^+$ is as in Eq. (35) then $S_0\,S_1=S_1\,S_0$.

Theorem 6.4. *Let S_0, $S_1\in\mathcal{H}(d)$ be such that $\lambda(S_0+S_1)=\big(\lambda(S_0)+\lambda^{\uparrow}(S_1)\big)^{\downarrow}$. Then S_0 and S_1 commute.*

Proof. Take $B=S_0+S_1$ and $A=-S_1$. Therefore $-\lambda(A)=\lambda^{\uparrow}(-A)=\lambda^{\uparrow}(S_1)$, so that $\lambda(A+B)-\lambda(A)=\lambda(S_0)+\lambda^{\uparrow}(S_1)$. Hence A and B satisfy the assumptions in Proposition 6.3 and they must commute. In this case S_0 and S_1 also commute. □

6.2 Characterization of optimal matching matrices

Let $S_0 \in \mathcal{M}_d(\mathbb{C})^+$ and let $S_1 \in \mathcal{M}_d(\mathbb{C})^+$ be an optimal matching matrix for S_0. Then, Theorem 6.4 implies that $S_0\, S_1 = S_1\, S_0$ and hence there exists a common ONB of eigenvectors for S_0 and S_1. In order to complete describe S_0 and S_1 we first consider some technical results.

We begin by fixing some notations. Let $\lambda \in \mathbb{R}^d_{>0}$. For every $j \in \mathbb{I}_d$ we define the set

$$L(\lambda\,,\, j) = \{i \in \mathbb{I}_d : \lambda_i = \lambda_j\} \ .$$

If we assume that $\lambda = \lambda^\downarrow$ or $\lambda = \lambda^\uparrow$ then the sets $L(j)$ are formed by consecutive integers. In the firs case we have that $\lambda_i < \lambda_j \implies k > l$ for every $k \in L(\lambda\,,\, i)$ and $l \in L(\lambda\,,\, j)$.

Given a permutation $\sigma \in \mathbb{S}_d$ and $\lambda \in \mathbb{R}^d_{>0}$ we denote by $\lambda_\sigma = (\lambda_{\sigma(1)}\,,\, \dots\,,\, \lambda_{\sigma(d)})$. Observe that

$$\lambda = \lambda_\sigma \iff \lambda = \lambda_{\sigma^{-1}} \iff \sigma\big(L(\lambda\,,\, j)\big) = L(\lambda\,,\, j) \quad \text{for every} \quad j \in \mathbb{I}_d \ . \tag{41}$$

The following inequality is well known (see for example [3, II.5.15]):

Proposition 6.5 (Rearrangement inequality for products of sums)**.** Let $\lambda\,,\, \mu \in \mathbb{R}^d_{>0}$ be such that $\lambda = \lambda^\downarrow$ and $\mu = \mu^\uparrow$. Then $\prod_{i=1}^d (\lambda_i + \mu_i) \geq \prod_{i=1}^d (\lambda_i + \mu_{\sigma(i)})$ for every permutation $\sigma \in \mathbb{S}_d$.

The following result deals with the case of equality in the last inequality.

Proposition 6.6. *Let λ, $\mu \in \mathbb{R}^d_{>0}$ be such that $\lambda = \lambda^\downarrow$ and $\mu = \mu^\uparrow$. Let $\sigma \in \mathbb{S}_d$ be such that*

$$(\lambda + \mu)^\downarrow = (\lambda + \mu_\sigma)^\downarrow \ .$$

Moreover, assume that σ also satisfies that:

$$\text{if} \ \ r\,,\, s \in \mathbb{I}_d \ \text{ are such that } \ \mu_{\sigma(r)} = \mu_{\sigma(s)} \ \text{ with } \ \sigma(r) < \sigma(s) \ \text{ then } \ r < s \ . \tag{42}$$

Then the permutation σ satisfies that $\lambda = \lambda_\sigma$.

Proof. For every $\tau \in \mathbb{S}_d$ let $F(\tau) = \prod_{i=1}^d (\lambda_i + \mu_{\tau(i)})$. By the hypothesis and Proposition 6.5,

$$F(\sigma) = F(\text{id}) = \max_{\tau \in \mathbb{S}_d} F(\tau) \ .$$

Assume that $\lambda \neq \lambda_{\sigma^{-1}}$. In this case there exists $j\,,\, k \in \mathbb{I}_d$ such that

$$\mu_j < \mu_k \quad \text{ and } \quad \lambda_{\sigma^{-1}(j)} < \lambda_{\sigma^{-1}(k)} \ . \tag{43}$$

Indeed, let j_0 be the smallest index such that σ^{-1} does not restrict to a permutation on $L(\lambda\,,\, j_0)$. Then, there exists $j \in L(\lambda\,,\, j_0)$ such that $\sigma^{-1}(j) \notin L(\lambda\,,\, j_0)$. As $\sigma^{-1}(L(\lambda\,,\, j_0) \setminus \{j\}) \neq L(\lambda\,,\, j_0)$ there also exists $k \notin L(\lambda\,,\, j_0)$ such that $\sigma^{-1}(k) \in L(\lambda\,,\, j_0)$. They have the required properties:

- First note that $\lambda_{\sigma^{-1}(j)} < \lambda_{j_0} = \lambda_{\sigma^{-1}(k)}$ (and then also $\sigma^{-1}(j) > \sigma^{-1}(k)$) because $\sigma^{-1}(j)$ can not be in $L(\lambda\,,\, j_0)$ nor in $L(\lambda\,,\, r)$ for any $r < j_0$ (where σ^{-1} acts as a permutation).
- A similar argument shows that $j < k$. We have used in both cases that the sets $L(\lambda\,,\, j)$ are formed by consecutive integers, since the vector λ is decreasingly ordered.
- Observe that $j < k \implies \mu_j \leq \mu_k$. So it suffices to show that $\mu_j \neq \mu_k$. Let us denote by $r = \sigma^{-1}(j)$ and $s = \sigma^{-1}(k)$. The previous items show that $r > s$ and $\sigma(r) < \sigma(s)$. Hence the equality $\mu_j = \mu_{\sigma(r)} = \mu_{\sigma(s)} = \mu_k$ is forbidden by our hypothesis (42).

So Eq. (43) is proved. Consider now the permutation $\tau = \sigma^{-1} \circ (j\,,\,k)$, where $(j\,,\,k)$ stands for the transposition of the indexes j and k. Straightforward computations show that

$$(\lambda_{\sigma^{-1}(j)} + \mu_j)\,(\lambda_{\sigma^{-1}(k)} + \mu_k) - (\lambda_{\sigma^{-1}(j)} + \mu_k)\,(\lambda_{\sigma^{-1}(k)} + \mu_j) = (\lambda_{\sigma^{-1}(j)} - \lambda_{\sigma^{-1}(k)})\,(\mu_k - \mu_j) \overset{(43)}{<} 0\ .$$

From the previous inequality we conclude that $F(\mathrm{id}) = F(\sigma) < F(\tau) \le F(\mathrm{id})$. This contradiction arises from the assumption $\lambda \neq \lambda_{\sigma^{-1}}$. Therefore $\lambda = \lambda_{\sigma^{-1}} \overset{(41)}{=} \lambda_\sigma$ as desired. □

Remark 6.7. Let $\lambda\,,\,\mu \in \mathbb{R}^d_{>0}$ be such that $\lambda = \lambda^\downarrow$ and $\mu = \mu^\uparrow$. Let $\tau \in \mathbb{S}_d$ be such that $(\lambda+\mu)^\downarrow = (\lambda + \mu_\tau)^\downarrow$. Then, by considering convenient permutations of the sets $L(\mu\,,\,j)$ we can always replace τ by σ in such a way that $\mu_\sigma = \mu_\tau$ and such that this σ satisfies the condition (42) of Proposition 6.6. Hence, in this case $(\lambda+\mu)^\downarrow = (\lambda+\mu_\sigma)^\downarrow$ and the previous result applies. △

Theorem 6.8 (Equality in Lidskii's inequality)**.** *Let $S_0 \in \mathcal{M}_d(\mathbb{C})^+$ and let $S_1 \in \mathcal{M}_d(\mathbb{C})^+$ be an optimal matching matrix for S_0. Let $\lambda = \lambda(S_0)$ and $\mu = \lambda^\uparrow(S_1)$. Then there exists $\{v_i : i \in \mathbb{I}_d\}$ a ONB for S_0 and λ such that*

$$S_1 = \sum_{i\in\mathbb{I}_d} \mu_i \cdot v_i \otimes v_i \quad \textit{and} \quad S_0 + S_1 = \sum_{i\in\mathbb{I}_d} (\lambda_i + \mu_i)\, v_i \otimes v_i\ . \tag{44}$$

Proof. Let us assume further that $S_0\,,\ S_1$ are invertible matrices so that $\lambda\,,\,\mu \in \mathbb{R}^d_{>0}$. By Theorem 6.4 we see that S_0 and S_1 commute. Then, there exists $\mathcal{B} = \{w_i : i \in \mathbb{I}_d\}$ an ONB for S_0 and λ such that $S_1\, w_i = \mu_{\tau(i)}\, w_i$ for every $i \in \mathbb{I}_d$, and for some permutation $\tau \in \mathbb{S}_d$. Therefore

$$\big(\,\lambda + \mu\,\big)^\downarrow \overset{(35)}{=} \lambda(S_0 + S_1) = \big(\,\lambda + \mu_\tau\,\big)^\downarrow\ .$$

By Remark 6.7 we can replace τ by $\sigma \in \mathbb{S}_d$ in such a way that $\mu_\tau = \mu_\sigma$, $(\lambda+\mu)^\downarrow = (\lambda+\mu_\sigma)^\downarrow$ and σ satisfies the hypothesis (42). Hence, by Proposition 6.6, we deduce that $\lambda_{\sigma^{-1}} = \lambda$. Therefore one easily checks that the ONB formed by the vectors $v_i = w_{\sigma^{-1}(i)}$ for $i \in \mathbb{I}_d$ (i.e. the rearrangement $\mathcal{B}_{\sigma^{-1}}$ of $\mathcal{B}$) is still a ONB for S_0 and λ, but it now satisfies Eq. (44).

In case S_0 or S_1 are not invertible, we can argue as above with the matrices $\tilde{S}_0 = S_0 + I$ and $\tilde{S}_1 = S_1 + I$. These matrices are invertible and such that $\tilde{S}_1$ is an optimal matching for $\tilde{S}_0$. Further, $\lambda(\tilde{S}_0) = \lambda(S_0) + \mathbb{1}$ and $\lambda(\tilde{S}_1) = \lambda(S_1) + \mathbb{1}$. Hence, if $\{v_i :\ i \in \mathbb{I}_d\}$ has the desired properties for $\tilde{S}_0$ and $\tilde{S}_1$ then this ONB also has the desired properties for S_0 and S_1. □

References

[1] J. Antezana, P. Massey, M. Ruiz and D. Stojanoff, The Schur-Horn theorem for operators and frames with prescribed norms and frame operator, Illinois J. Math., 51 (2007), 537-560.

[2] J.J. Benedetto, M. Fickus, Finite normalized tight frames, Adv. Comput. Math. 18, No. 2-4 (2003), 357-385 .

[3] R. Bhatia, Matrix Analysis, Berlin-Heildelberg-New York, Springer 1997.

[4] B.G. Bodmann, Optimal linear transmission by loss-insensitive packet encoding, Appl. Comput. Harmon. Anal. 22, no. 3, (2007) 274-285.

[5] B.G. Bodmann, D.W. Kribs, V.I. Paulsen, Decoherence-Insensitive Quantum Communication by Optimal C^*-Encoding, IEEE Transactions on Information Theory 53 (2007) 4738-4749.

[6] B.G. Bodmann, V.I. Paulsen, Frames, graphs and erasures, Linear Algebra Appl. 404 (2005) 118-146.

[7] J. Cahill, M. Fickus, D.G. Mixon, M.J. Poteet, N.K. Strawn, Constructing finite frames of a given spectrum and set of lengths, preprint arXiv:1106.0921, 2011.

[8] R. Calderbank, P.G. Casazza, A. Heinecke, G. Kutyniok, A. Pezeshki, Fusion frames: existence and construction, preprint arXiv:0906.5606, 2009.

[9] P.G. Casazza, The art of frame theory, Taiwanese J. Math. 4 (2000), no. 2, 129-201.

[10] P.G. Casazza, Custom building finite frames. In Wavelets, frames and operator theory, volume 345 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2004, 61-86.

[11] P.G. Casazza, M. Fickus, Minimizing fusion frame potential. Acta Appl. Math. 107 (2009), no. 1-3, 7-24.

[12] P.G. Casazza, M. Fickus, J.C. Tremain, and E. Weber, The Kadison-Singer problem in mathematics and engineering: a detailed account. Operator theory, operator algebras, and applications, 299-355, Contemp. Math., 414, Amer. Math. Soc., Providence, RI, 2006.

[13] P.G. Casazza, and M.T. Leon, Existence and construction of finite frames with a given frame operator. Int. J. Pure Appl. Math. 63 (2010), no. 2, 149-157.

[14] P.G. Casazza, M. Fickus, J. Kovacevic, M. T. Leon,J. C. Tremain, A physical interpretation of tight frames, Harmonic analysis and applications, 51-76, Appl. Numer. Harmon. Anal., Birkhäuser Boston, MA, 2006.

[15] P.G. Casazza, J. Kovacevic, Equal-norm tight frames with erasures. Adv. Comput. Math. 18 (2003), no. 2-4, 387-430.

[16] O. Christensen, An introduction to frames and Riesz bases. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, Inc., Boston, MA, 2003. xxii+440 pp.

[17] I.S. Dhillon, R.W. Heath Jr., M.A. Sustik, J.A. Tropp, Generalized finite algorithms for constructing Hermitian matrices with prescribed diagonal and spectrum, SIAM J. Matrix Anal. Appl. 27 (1) (2005) 61-71.

[18] K. Dykema, D. Freeman, K. Kornelson, D. Larson, M. Ordower, E. Weber, Ellipsoidal tight frames and projection decomposition of operators: Illinois J. Math. 48 (2004), 477-489.

[19] D. J. Feng, L. Wang and Y. Wang, Generation of finite tight frames by Householder transformations. Adv Comput Math 24 (2006), 297-309.

[20] M. Fickus, D.G. Mixon and M.J. Poteet, Frame completions for optimally robust reconstruction, preprint arXiv:1107.1912, 2011.

[21] M. Fickus, B.D. Johnson, K. Kornelson, K.A. Okoudjou, Convolutional frames and the frame potential. Appl. Comput. Harmon. Anal. 19 (2005), no. 1, 77-91.

[22] D. Han and D.R. Larson, Frames, bases and group representations. Mem. Amer. Math. Soc. 147 (2000), no. 697, x+94 pp.

[23] R.B. Holmes, V.I. Paulsen, Optimal frames for erasures, Linear Algebra Appl. 377 (2004) 31-51.

[24] B.D. Johnson, K.A. Okoudjou, Frame potential and finite abelian groups. Radon transforms, geometry, and wavelets, 137-148, Contemp. Math., 464, Amer. Math. Soc., Providence, RI, 2008.

[25] A. Klyachko, Stable bundles, representation theory and Hermitian operators. Selecta Math. (N.S.) 4 (1998), no. 3, 419-445.

[26] K. A. Kornelson, D. R. Larson, Rank-one decomposition of operators and construction of frames. Wavelets, frames and operator theory, Contemp. Math., 345, Amer. Math. Soc., Providence, RI, 2004, 203-214.

[27] J. Leng, D. Han, Optimal dual frames for erasures II. Linear Algebra Appl. 435 (2011), 1464-1472.

[28] J. Lopez, D. Han, Optimal dual frames for erasures. Linear Algebra Appl. 432 (2010), 471-482.

[29] P. Massey, M.A. Ruiz, Tight frame completions with prescribed norms. Sampl. Theory Signal Image Process. 7 (2008), no. 1, 1-13.

[30] P. Massey and M. Ruiz, Minimization of convex functionals over frame operators, Adv. Comput. Math. 32 (2010), 131-153.

[31] P. Massey, M. Ruiz and D. Stojanoff, The structure of minimizers of the frame potential on fusion frames, J Fourier Anal. Appl. 16 N° 4 (2010) 514-543.

[32] P. Massey, M. Ruiz and D. Stojanoff, Duality in reconstruction systems. Linear Algebra Appl. 436 (2012), 447-464.

[33] P. Massey, M. Ruiz and D. Stojanoff, Optimal dual frames and frame completions for majorization, Appl. Comput. Harmon. Anal. (in press).