Optimal frame completions with prescribed norms for majorization

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 characterize in a complete and explicit way the optimal completions of $\mathcal{F}_0$ obtained by adding a finite sequence of vectors with prescribed norms, where optimality is measured with respect to majorization (of the eigenvalues of the frame operators of the completed sequence). Indeed, we construct (in terms of a fast algorithm) a vector - that depends on the eigenvalues of the frame operator of the initial sequence $\mathcal{F}_0$ and the sequence of prescribed norms - that is a minimum for majorization among all eigenvalues of frame operators of completions with prescribed norms. Then, using the eigenspaces of the frame operator of the initial sequence $\mathcal{F}_0$ we describe the frame operators of all optimal completions for majorization. Hence, the concrete optimal completions with prescribed norms can be obtained using recent algorithmic constructions related with the Schur-Horn theorem.

The well known relation between majorization and tracial inequalities with respect to convex functions allow to describe our results in the following equivalent way: given a finite sequence of vectors $\mathcal{F}_0$ in $\mathbb{C}^d$ we show that the completions with prescribed norms that minimize the convex potential induced by a strictly convex function are structural minimizers, in the sense that they do not depend on the particular choice of the convex potential.

AMS subject classification: 42C15, 15A60.
Keywords: frames, frame completions, majorization, convex potentials, 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 (stable) linear encoding-decoding schemes. As opposed to bases, frame are not subject to linear independence; indeed, it turns out that the redundancy allowed in finite frames can be turned into robustness of the transmission scheme that they induce, which make frames a useful device for transmission of signals through noisy channels (see [5, 6, 7, 15, 24, 28, 27]).

On the other hand, the so-called tight frames allow for redundant linear representations of vectors 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. 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 [2, 8, 11, 13, 17, 18, 19, 26]. It turns out that in some cases it is not possible to find a frame fulfilling the previous demands.

An alternative approach to deal with the construction of frames with prescribed parameters and nice associated reconstruction formulas was posed in [3] by Benedetto and Fickus; they defined a functional, called the frame potential, and showed 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, 22, 25, 30] and [12, 31, 32] for related problems in the context of fusion frames). Moreover, in [30] it is shown that minimizers of the frame potential under suitable restrictions (considered in the literature) are structural minimizers in the sense that they coincide with minimizers of more general *convex potentials* (see Section 2.2).

Recently, the following frame completion problem was posed in [20] (in the vein of [3]): 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 [9, 19, 29] for completion problems for frames). In this setting, 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 functional defined on the set of frames, compute the frame completions with prescribed norms that minimize this functional. Moreover, this last problem raises the question of whether the completions that minimize these functionals coincide i.e., whether the minimizers are structural in this setting.

A first step towards the solution of the general version of the completion problem was made in [33]. There we showed that under certain hypothesis (feasible cases, see Section 3.3), optimal frame completions with prescribed norms are structural (do not depend on the particular choice of the convex functional), as long as we consider convex potentials, that contain the MSE and the frame potential. 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 were not known even for the MSE nor the frame potential. Recently, in some feasible cases the set of all optimal frame completions is characterized in [35, 21].

In [34] we considered the structure of completions that minimize a fixed convex potential (non feasible case). There, we showed that the eigenvalues of optimal completions with respect to a fixed convex potential are uniquely determined by the solution of an optimization problem in a compact convex subset of $\mathbb{R}^d$ for a convex objective function that is associated to the convex potential in a natural way. Then, we showed an important geometrical feature of optimal completions $\mathcal{F} = (\mathcal{F}_0, \mathcal{G})$ for a fixed convex potential, namely that the vectors in the completion $\mathcal{G}$ are eigenvectors of the frame operator of the completed sequence $\mathcal{F}$ (see Section 3.2 for a detailed exposition of these results). Based on these facts, we developed an algorithm that allowed us to compute the solutions of the completion problem for small dimensions. In this setting we conjectured some properties of the optimal frame completions in the general case, based on common features of the solutions of several examples obtained by this algorithm (see Section 4 for a detailed description of these conjectures).

In this paper, building on our previous work [33] and [34], we give a complete and explicit description of the spectral and geometrical structure of optimal completions with prescribed norms with respect to a convex potential induced by a strictly convex function. Our approach is constructive and allows to develop a fast and effective algorithm that computes the spectral structure of optimal completions. As we shall see, given an initial sequence $\mathcal{F}_0$ in $\mathbb{C}^d$ and a sequence of positive numbers $\mathbf{a}$, both the spectral and geometrical structure of optimal completions depend only on the frame operator of $\mathcal{F}_0$ and $\mathbf{a}$, but they do not depend on the particular choice of the convex potential. Hence, we show that in the general case the minimizers of convex potentials (induced by strictly convex functions) are structural.

In order to obtain the previous results, we begin by proving the properties of general optimal completions conjectured in [34]. These properties (that are structural, in the sense that they do not depend on the convex potential) are then used to compute several other structural parameters - that involve the notion of feasibility developed in [33] - that completely describe the spectral structure of optimal completions. As a consequence of this description, we conclude that optimal solutions have the same eigenvalues and hence, the eigenvalues of optimal completions are minimum for the so-called majorization preorder. Moreover, all the parameters involved in the description of the spectral structure of optimal completions can be computed in terms of fast algorithms. With the spectral data and results from [33] we completely describe the set positive matrices that correspond to the frame operators of sequences $\mathcal{G}$ with norms prescribed by $\mathbf{a}$ and such that $\mathcal{F} = (\mathcal{F}_0, \mathcal{G})$ are optimal. Finally, some optimal completions $\mathcal{G}$ can be also effectively computed by using recent results from [8] (see also [17] and [21]) that implements the Schur-Horn theorem.

The paper is organized as follows. In Section 2 we briefly recall the basic framework of finite frame theory, the notion of submajotization - that will play a central role in this note - and the relation of submajorization with tracial inequalities involving convex functions. In section 3 we describe the context of our main problem - namely, optimal completions with prescribed norms, where optimality is described in terms of majorization - and give a detailed account of several related results that were developed in our previous works [33] and [34] that we shall need in the sequel, in a way suitable for this note; in particular, we include a new construction of the spectra of optimal completions in the feasible cases. In Section 4 we introduce new structural parameters - that can be efficiently computed in terms of explicit algorithms - and show how to give a complete description of the spectra of optimal completions for strictly convex potentials, in terms of these parameters in the general case. This allow us to show that the spectra of such optimal completions do not depend on the choice of strictly convex potential, so that minimizers are then structural. The proofs of the technical results of this section is presented in Section 5. In particular, we settle in the affirmative some features of the structure of optimal completions for strictly convex potentials that were conjectured in [34]. As a byproduct we also settle in the affirmative a conjecture on local minimizers of strictly convex potentials with prescribed norms posed in [30].

2 Preliminaries

In this section we describe the basic notions that we shall consider throughout the paper. In Section 2.1 we describe some general notations and terminology. In Section 2.2 we describe some basic notions and facts of frame theory and we recall the notion of convex potential from [30]. In Section 2.3 we describe some aspects of submajorization that we shall need in the sequel.

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 $\operatorname{tr} x = \sum_{i\in\mathbb{I}_m} x_i$ and by $x^\downarrow$ (resp. $x^\uparrow$) the rearrangement of x in decreasing (resp. increasing) order. We denote by $(\mathbb{R}^m)^\downarrow = \{x \in \mathbb{R}^m : x = x^\downarrow\}$ the set of downwards ordered vectors, and similarly $(\mathbb{R}^m)^\uparrow$.

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 transformations $T : \mathcal{H} \to \mathcal{K}$. 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.

If $W \subseteq \mathcal{H}$ is a subspace we denote by $P_W \in L(\mathcal{H})^+$ the orthogonal projection onto W. 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 $x \otimes x = P_{\operatorname{span}\{x\}}$.

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 in $\mathcal{M}_d(\mathbb{C})$, $\mathcal{M}_d(\mathbb{C})^+$ the cone of positive semidefinite matrices, and $\mathcal{G}l(d)^+ = \mathcal{M}_d(\mathbb{C})^+ \cap \mathcal{G}l(d)$.

Given $S \in \mathcal{M}_d(\mathbb{C})^+$, we write $\lambda(S) = \lambda^\downarrow(S) \in (\mathbb{R}^d_{\geq 0})^\downarrow$ the vector of eigenvalues of S - counting multiplicities - arranged in decreasing order. Similarly we denote by $\lambda^\uparrow(S) \in (\mathbb{R}^d_{\geq 0})^\uparrow$ the reverse ordered vector of eigenvalues of S. If $\lambda = (\lambda_i)_{i\in\mathbb{I}_d} \in \mathbb{R}^d_{\geq 0}$ (not necessarily ordered), a system $\mathcal{B} = \{h_i\}_{i\in\mathbb{I}_d} \subseteq \mathbb{C}^d$ is a "ONB of eigenvectors for $S\,,\,\lambda$" 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

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

2.2 Basic framework of finite frames

In what follows we consider (n,d)-frames. See [3, 10, 16, 23, 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

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

is the **analysis** operator of $\mathcal{F}$. Its adjoint $T^*_\mathcal{F} \in L(\mathbb{C}^n\,,\,\mathcal{H})$ is called the **synthesis** operator and is given by $T^*_\mathcal{F}\, v = \sum_{i\in\mathbb{I}_n} v_i\, f_i$ for every $v = (v_i)_{i\in\mathbb{I}_n} \in \mathbb{C}^n$. The **frame operator** of $\mathcal{F}$ is

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

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 $\tau = \operatorname{tr} S_\mathcal{F} = \sum_{i\in\mathbb{I}_n} \|f_i\|^2$.

In their seminal work [3], Benedetto and Fickus introduced a functional defined, the so-called frame potential, given by

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

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 us denote by

$$\operatorname{Conv}(\mathbb{R}_{\ge 0}) = \{f : [0\,,\,\infty) \to [0\,,\,\infty) : f \ \text{ is a convex function } \ \}$$

and $\operatorname{Conv_s}(\mathbb{R}_{\ge 0}) = \{f \in \operatorname{Conv}(\mathbb{R}_{\ge 0}) : f$ is strictly convex $\}$. Following [30] we consider the (generalized) convex potential P_f associated to any $f \in \operatorname{Conv}(\mathbb{R}_{\ge 0})$, given by

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

where the matrix $f(S_\mathcal{F})$ is defined by means of the usual functional calculus. △

As shown in [30, Sec. 4] these convex potentials (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] and Remark 5.6 below).

Remark 2.2. The results that we shall develop in this work apply in the case of convex potentials P_f for any $f \in \operatorname{Conv_s}(\mathbb{R}_{\ge 0})$. 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 $\tilde f$. △

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 [4].

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 $x \prec y$. In case that $x \prec y$ but $y \not\prec x$ we say that y majorizes x strictly. If the two vectors x and y have different sizes, 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, if we let $x, y \in \mathbb{R}^d$ and let $I \subseteq \mathbb{R}$ be an interval such that $x, y \in I^d$ then (see for example [4]):

1. $x \prec y \Leftrightarrow \operatorname{tr} f(x) \stackrel{\text{def}}{=} \sum\limits_{i=1}^{d} f(x_i) \le \sum\limits_{i=1}^{d} f(y_i) = \operatorname{tr} f(y)$ for every convex function $f : I \to \mathbb{R}$.
2. If only $x \prec_w y$, but the map $f : I \to \mathbb{R}$ is convex and increasing, then $\operatorname{tr} f(x) \le \operatorname{tr} f(y)$.
3. If $x \prec y$ and $f : I \to \mathbb{R}$ 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 $i \in \mathbb{I}_d$.

As a consequence of item 3. above, if $x \prec y$ strictly and $f : I \to \mathbb{R}$ is a strictly convex function then $\operatorname{tr} f(x) < \operatorname{tr} f(y)$: indeed, if $\operatorname{tr} f(x) = \operatorname{tr} f(y)$ then by item 3. above we would have that $y_i = x_{\sigma(i)}$, $i \in \mathbb{I}_d$, for a permutation σ of $\mathbb{I}_d$ and hence that $y \prec x$.

The notion of vector submajorization can be extended to a preorder between selfadjoint matrices as follows: given $S_1 , S_2 \in \mathcal{H}(d)$ we say that S_1 is submajorized by S_2 , and write $S_1 \prec_w S_2$ (resp. $S_1 \prec S_2$) if $\lambda(S_1) \prec_w \lambda(S_2)$ (resp. $\lambda(S_1) \prec \lambda(S_2)$, i.e. $S_1 \prec_w S_2$ and $\operatorname{tr} S_1 = \operatorname{tr} S_2$).

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 [4]) 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

$$x_i \prec y_i \quad \text{for} \quad i = 1 , 2 \implies x = (x_1 , x_2) \prec y = (y_1 , y_2) \quad \text{in} \quad \mathbb{R}^{r+s} \ . \tag{4}$$

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 $D\, y = x$. △

3 Optimal completions with prescribed norms

In this section we give a detailed description of the optimal completion problem and recall some notions and results from our previous work [33, 34], in a way suitable for the exposition of the results herein. In particular, the exposition of the results in Section 3.3 differs from that of [34], since this new presentation is better suited for our present purposes.

3.1 Presentation of the problem

In several applied situations it is desired to construct a sequence $\mathcal{G}$ in such a way that the frame operator of $\mathcal{G}$ is given by some $B \in \mathcal{M}_d(\mathbb{C})^+$ and the squared norms of the frame elements are prescribed by a sequence of positive numbers $\mathbf{a} = (a_i)_{i\in\mathbb{I}_k} \in \mathbb{R}^k_{>0}$. That is, given a fixed $B \in \mathcal{M}_d(\mathbb{C})^+$ and $\mathbf{a} \in \mathbb{R}^k_{>0}$, we analyze the existence (and construction) of a sequence $\mathcal{G} = \{g_i\}_{i\in\mathbb{I}_k}$ such that $S_\mathcal{G} = B$ and $\|g_i\|^2 = a_i$, for $i \in \mathbb{I}_k$. This is known as the classical frame design problem. It has been treated by several research groups (see for example [2, 8, 11, 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 3.1 ([2, 29])**.** Let $B \in \mathcal{M}_d(\mathbb{C})^+$ with $\lambda(B) \in (\mathbb{R}^d_{\geq 0})^\downarrow$ and let $\mathbf{a} = (a_i)_{i\in\mathbb{I}_k} \in \mathbb{R}^k_{>0}$. 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 $\|g_i\|^2 = a_i$ for every $i \in \mathbb{I}_k$ if and only if $\mathbf{a} \prec \lambda(B)$ (completing with zeros if $k \neq d$). $\square$

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. Indeed, 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 $\mathbf{a} = (a_i)_{i\in\mathbb{I}_k} \in \mathbb{R}^k_{>0}$ 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

$$\mathcal{G} = \{g_i\}_{i\in\mathbb{I}_k} \in \mathcal{H}^k \quad \text{with} \quad \|g_i\|^2 = a_i \quad \text{for} \quad i \in \mathbb{I}_k \ ,$$

such that the resulting completed sequence $\mathcal{F} = (\mathcal{F}_0 \, , \, \mathcal{G}) \in \mathbf{F}(n \, , \, d)$ - obtained by juxtaposition of the two finite sequences - is a frame whose MSE, given by 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 $\mathcal{F} = (\mathcal{F}_0 \, , \, \mathcal{G}) \in \mathbf{F}(n \, , \, d)$, with $\|g_i\|^2 = a_i$ for $i \in \mathbb{I}_k$, and such that its frame potential FP $(\mathcal{F}) =$ 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 $\mathcal{F}_0 = \varnothing$ in [3, 14, 22, 25, 30]). More generally, we can measure robustness of the completed frame $\mathcal{F} = (\mathcal{F}_0, \mathcal{G})$ in terms of general convex potentials (see Definition 2.1).

In order to describe the main problems we first fix the notation that we shall use throughout the paper.

Definition 3.2. Let $\mathcal{F}_0 = \{f_i\}_{i\in\mathbb{I}_{n_o}} \in \mathcal{H}^{n_o}$ and $\mathbf{a} = (a_i)_{i\in\mathbb{I}_k} \in (\mathbb{R}^k_{>0})^\downarrow$ 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

$$\mathcal{C}_\mathbf{a}(\mathcal{F}_0) = \big\{ \, (\mathcal{F}_0, \mathcal{G}) \in \mathcal{H}^n : \ \mathcal{G} = \{g_i\}_{i\in\mathbb{I}_k} \quad \text{and} \quad \|g_i\|^2 = a_i \ \text{ for } \ i \in \mathbb{I}_k \big\} \ ,$$

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

When the initial data $(\mathcal{F}_0 \, , \, \mathbf{a})$ are fixed, we shall use the notations $S_0 = S_{\mathcal{F}_0}$ and $\lambda = \lambda^\uparrow(S_0)$.

We remark that we shall use the vector $\lambda = \lambda^\uparrow(S_0)$ instead of $\lambda^\downarrow(S_0)$ for convenience (see the comments at the beginning of Section 3). $\triangle$

Main problems: (Optimal completions with prescribed norms for majorization) Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP and let $f \in \text{Conv}_s(\mathbb{R}_{\geq 0})$.

P1. Give an explicit description (both spectral and geometrical) of $\mathcal{F} \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ that are the minimizers of P_f in $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$.

P2. Construct a fast algorithm that efficiently computes all possible $\mathcal{F} \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ that are the minimizers of P_f in $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$.

P3. Verify that the set of $\mathcal{F} \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ that are the minimizers of P_f in $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ is the same for every $f \in \text{Conv}_s(\mathbb{R}_{\geq 0})$.

$\triangle$

In previous works we have obtained some results related with the problems above. Indeed, in [33] we obtained a partial affirmative answer to P3, while in [34] we obtained some partial results related with P1. and a non-efficient algorithm as in P2. that worked in small examples (see Sections 3.2 and 3.3 below).

In this paper, building on our previous work, we completely solve the three problems above in terms of a constructive (algorithmic) approach.

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

In this section we collect results of [34] that we shall use in this paper. Throughout this section we fix the initial data $(\mathcal{F}_0\,,\,\mathbf{a})$ for the CP. Notice that we are using the following convention in Definition 3.2: we denote $\lambda = \lambda^{\uparrow}(S_{\mathcal{F}_0}) \in \mathbb{R}^d$, i.e. arranged in non-decreasing order. Thus we recast the results from [34] using this convention. Also notice that we are assuming that $\mathbf{a} = \mathbf{a}^{\downarrow} \in \mathbb{R}^k$.

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

Proposition 3.3. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. Then

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

Let $\mu \in (\mathbb{R}^d_{\geq 0})^{\downarrow}$ be such that $\mathbf{a} \prec \mu$, and let

$$\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu) \stackrel{\text{def}}{=} \{\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0) :\; \lambda(S_{\mathcal{G}}) = \mu\} \subseteq \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0) .$$

By Proposition 3.3 we get the following partition:

$$\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0) = \bigsqcup_{\mu \in \Gamma_d(\mathbf{a})} \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu) \quad \text{where} \quad \Gamma_d(\mathbf{a}) \stackrel{\text{def}}{=} \{\mu \in (\mathbb{R}^d_{\geq 0})^{\downarrow} :\; \mathbf{a} \prec \mu\} . \tag{5}$$

Building on Lidskii's inequality (see [4, III.4]) we obtained the following result:

Theorem 3.4. Consider the previous notations and fix $\mu = \mu^{\downarrow} \in \Gamma_d(\mathbf{a})$. Then,

1. The set $\Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)) \stackrel{\text{def}}{=} \{\lambda(S_{\mathcal{F}}) :\; \mathcal{F} \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)\}$ is convex.
2. Let $\nu \stackrel{\text{def}}{=} \lambda^{\uparrow}(S_{\mathcal{F}_0}) + \mu^{\downarrow}$. Then $\nu^{\downarrow}$ 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 $\lambda(S_{\mathcal{F}}) = \nu^{\downarrow}$ then $S_{\mathcal{F}_0}$ and $S_{\mathcal{G}}$ commute. $\square$

Remark 3.5. Consider the previous notations and fix $\mu = \mu^\downarrow \in \Gamma_d(\mathbf{a})$. Let $f \in \text{Conv}_s(\mathbb{R}_{\geq 0})$ and let P_f be the convex potential induced by f. By the results described in Section 2.3 and Theorem 3.4 we see that, if $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$ then

$$\mathcal{F} \in \text{argmin}\{P_f(\mathcal{F}') : \mathcal{F}' \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0\,,\,\mu)\} \iff \lambda(S_{\mathcal{F}}) = (\lambda + \mu)^\downarrow = (\,\lambda^\uparrow + \mu^\downarrow\,)^\downarrow\ . \tag{6}$$

That is, if we consider the partition of $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ described in Eq. (5), 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 (6). 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

$$\mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0) \stackrel{\text{def}}{=} \big\{\,\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0) :\ \lambda(S_{\mathcal{F}}) = \big(\,\lambda^\uparrow(S_{\mathcal{F}_0}) + \lambda^\downarrow(S_{\mathcal{G}})\,\big)^\downarrow\,\big\}\ . \tag{7}$$

Indeed, Eqs. (5) and (6) show that if $\mathcal{F}$ is a minimizer of P_f in $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ then $\mathcal{F} \in \mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)$. Since the potential $P_f(\mathcal{F})$ depends on $\mathcal{F}$ through the eigenvalues of $S_{\mathcal{F}}$ we introduce the sets

$$\mathcal{S}(\mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)) \stackrel{\text{def}}{=} \{S_{\mathcal{F}} :\ \mathcal{F} \in \mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)\} \quad \text{and} \quad \Lambda(\mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)\,) \stackrel{\text{def}}{=} \{\lambda(S_{\mathcal{F}}) :\ \mathcal{F} \in \mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)\}\ . \tag{8}$$

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

$$\Lambda_{\mathbf{a}}^{\text{op}}(\lambda) \stackrel{\text{def}}{=} \{\lambda^\uparrow + \mu : \mu \in \Gamma_d(\mathbf{a})\} = \{\lambda^\uparrow + \mu^\downarrow :\ \mu \in \mathbb{R}^d_{\geq 0} \quad \text{and} \quad \mathbf{a} \prec \mu\}\ . \tag{9}$$

△

Theorem 3.6. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. Denote by $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$. Then

1. The set $\Lambda_{\mathbf{a}}^{\text{op}}(\lambda)$ is compact and convex.
2. The spectral picture $\Lambda(\mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)\,) = \{\nu^\downarrow :\ \nu \in \Lambda_{\mathbf{a}}^{\text{op}}(\lambda)\}$.
3. If $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)$, with $\lambda^\downarrow(S_{\mathcal{G}}) = \mu$, then there exists $\{v_i : i \in \mathbb{I}_d\}$ an ONB of eigenvectors for $S_{\mathcal{F}_0}$, λ such that

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

For every $f \in \text{Conv}(\mathbb{R}_{\geq 0})$ we consider the convex map

$$F : \mathbb{R}^d_{\geq 0} \to \mathbb{R} \quad \text{given by} \quad F(\gamma) = \text{tr}\ f(\gamma) = \sum_{i\in\mathbb{I}_d} f(\gamma_i)\ , \quad \text{for} \quad \gamma \in \mathbb{R}^d_{\geq 0}\ . \tag{10}$$

Theorem 3.7. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP and let $f \in \text{Conv}_s(\mathbb{R}_{\geq 0})$. Then there exists a vector $\mu_f(\lambda\,,\,\mathbf{a}) = \mu = \mu^\downarrow \in \Gamma_d(\mathbf{a})$ such that:

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

$$\mu \in \Gamma_d(\mathbf{a}) \quad \text{and} \quad F(\lambda + \mu) = \min_{\gamma\in\Gamma_d(\mathbf{a})} F(\lambda + \gamma) = \min_{\nu\in\Lambda_{\mathbf{a}}^{\text{op}}(\lambda)} F(\nu)\ . \tag{11}$$

Hence, if we let $\nu_f(\lambda\,,\,\mathbf{a}) \stackrel{\text{def}}{=} \lambda + \mu_f(\lambda\,,\,\mathbf{a})$ then $\exists !$ argmin $\{F(x) : x \in \Lambda_{\mathbf{a}}^{\text{op}}(\lambda)\} = \nu_f(\lambda\,,\,\mathbf{a})$. $\square$

Theorem 3.8. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. Let $f \in \text{Conv}_s(\mathbb{R}_{\geq 0})$ and assume that $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G})$ is a global minimizer of P_f on $\mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)$. Then, there exists a partition $\{J_i\}_{i\in\mathbb{I}_p}$ of $\mathbb{I}_k$ and $c_1 > \ldots > c_p > 0$ such that

1. The subfamilies $\mathcal{G}_i = \{f_j\}_{j\in J_i}$ (for $i \in \mathbb{I}_p$) are mutually orthogonal, i.e. $S_\mathcal{G} = \oplus_{i\in\mathbb{I}_p} S_{\mathcal{G}_i}$.
2. The frame operators $S_{\mathcal{G}_i}$ and $S_{\mathcal{F}_0}$ commute, for every $i \in \mathbb{I}_p$.
3. We have that $S_\mathcal{F}\, f_j = c_i\, f_j$, for every $j \in J_i$ and every $i \in \mathbb{I}_p$.

The statement is still valid if we assume that $\mathcal{F}$ is just a local minimizer, but if we also assume as a hypothesis that $\mathcal{F}$ satisfies item 2 (for example if $S_{\mathcal{F}_0} = 0$). □

3.3 The feasible case of the CP

In this section we recall the results from [33] that we shall need in the sequel. Throughout this section we fix the initial data $(\mathcal{F}_0, \mathbf{a})$ for the CP. Denote by $S_0 = S_{\mathcal{F}_0}$, $\lambda = \lambda^\uparrow(S_0)$ and $t = \operatorname{tr} \lambda + \operatorname{tr} \mathbf{a}$. In [33] we introduced the following set

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

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 $\mu(\lambda\,,\,\mathbf{a}) \in (\mathbb{R}^d_{\ge 0})^\downarrow$ - that can be effectively computed by a fast algorithm - such that, if $\nu(\lambda\,,\,\mathbf{a}) \overset{\text{def}}{=} \lambda + \mu(\lambda\,,\,\mathbf{a}) \in \mathbb{R}^d_{>0}$ then $S \in U_t(S_0\,,\,m)$ is a $\prec$-minimizer if and only if $\lambda(S) = \nu(\lambda\,,\,\mathbf{a})^\downarrow$.

Notice that by construction $\nu(\lambda\,,\,\mathbf{a})$ is not a necessarily ordered vector (nor decreasing, nor increasing); yet, in terms of the terminology from [33], we have that $\nu_{\lambda\,,\,m}(t) = \nu(\lambda\,,\,\mathbf{a})^\downarrow$. Thus, we have reversed the order of the vector $\mu(\lambda\,,\,\mathbf{a})$ - accordingly with reversing the order of $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$ - and we have changed the description of the vector $\nu(\lambda\,,\,\mathbf{a})$ - while preserving all of their majorization properties - with respect to [33]. Nevertheless, we point out that the ordering of the entries of the vector $\nu(\lambda\,,\,\mathbf{a})$ presented here plays a crucial role in simplifying the exposition of the results herein, as it guaranties that $\mu(\lambda\,,\,\mathbf{a}) = \nu(\lambda\,,\,\mathbf{a}) - \lambda$.

The following definition and remark show the relevance of the notions introduced above for the computation of the spectral structure of solutions for the optimal completion problem.

Definition 3.9. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP with $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$. We say that the pair $(\lambda\,,\,\mathbf{a})$ is **feasible** if $\mu(\lambda\,,\,\mathbf{a})$ satisfies that $\mathbf{a} \prec \mu(\lambda\,,\,\mathbf{a})$. △

Remark 3.10. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP with $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$. Assume that the pair $(\lambda\,,\,\mathbf{a})$ is feasible and denote $\mu = \mu(\lambda\,,\,\mathbf{a})$. In this case (see [33]) for any S which is a $\prec$-minimizer in $U_t(S_0\,,\,m)$ - where $m = d - k$ - it holds that $\lambda(S - S_0) = \mu$ and hence, by Proposition 3.3, we conclude that $S \in \mathcal{SC}_\mathbf{a}(\mathcal{F}_0)$. Moreover, Proposition 3.3 also shows that $\mathcal{SC}_\mathbf{a}(\mathcal{F}_0) \subseteq U_t(S_0\,,\,m)$. Then S is also a $\prec$-minimizer in $\mathcal{SC}_\mathbf{a}(\mathcal{F}_0)$. Therefore, as a consequence of the results in Section 2.3, any completion $\mathcal{F} = (\mathcal{F}_0, \mathcal{G}) \in \mathcal{C}_\mathbf{a}(\mathcal{F}_0)$ such that $S_\mathcal{F} = S$ is a minimizer of P_f for every $f \in \operatorname{Conv}(\mathbb{R}_{\ge 0})$.

On the other hand, as a consequence of the geometrical structure of $S = S_\mathcal{F}$ as above (see [33, 34]), we conclude that there exists $c > 0$ such that $S_\mathcal{F}\, g_i = c\, g_i$ for every $i \in \mathbb{I}_k$. That is, in this case the structure of the completing sequence $\mathcal{G}$ given in Theorem 3.8 is trivial: the partition of $\mathbb{I}_k$ has only one member and there exists a unique constant $c = c_1$. △

It is worth pointing out that it is easy to construct examples of initial data $(\mathcal{F}_0\,,\,\mathbf{a})$ for the CP such that the pair $(\lambda, \mathbf{a})$ is not feasible (see [33]), so that comments in Remark 3.10 do not apply in these cases.

Remark 3.11. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP with $k \geq d$ (so that $m = d - k \leq 0$), let $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$ and let $t = \text{tr }\mathbf{a} + \text{tr }\lambda$. In [33] we shown that there are two cases:

1. Since $\lambda = \lambda^\uparrow$ then $\lambda_d = \max\{\lambda_i : i \in \mathbb{I}_d\}$. If

$$\frac{t}{d} = \frac{\text{tr }\mathbf{a} + \text{tr }\lambda}{d} \geq \lambda_d \quad \text{then} \quad \lambda \preccurlyeq \frac{t}{d}\,\mathbb{1}_d = \nu(\lambda\,,\,\mathbf{a})\ . \tag{12}$$

2. If $\lambda_d > \frac{t}{d}$ then there exists $s \in \mathbb{I}_{d-1}$ such that

$$\nu(\lambda\,,\,\mathbf{a}) = (c\,\mathbb{1}_s\,,\,\lambda_{s+1}\,,\,\dots\,,\,\lambda_d) \quad \text{with} \quad \lambda_s \leq c < \lambda_{s+1} \tag{13}$$

so that $\lambda \preccurlyeq \nu(\lambda\,,\,\mathbf{a}) = \nu(\lambda\,,\,\mathbf{a})^\uparrow$, and in this case the index s also satisfies that

$$c = \frac{1}{s}\,\Big[\ \text{tr }\mathbf{a} + \sum_{i=1}^{s}\ \lambda_i\Big] \quad \text{so that} \quad \text{tr }\nu(\lambda\,,\,\mathbf{a}) = t = \text{tr }\lambda + \text{tr }\mathbf{a}\ .$$

In what follows we obtain an explicit description of the vector $\nu(\lambda\,,\,\mathbf{a})$ in case $d \leq k$ (so that $m \leq 0$) and $\frac{1}{d}\,[\text{tr }\mathbf{a} + \text{tr }\lambda\,] < \lambda_d$. Explicitly, we compute the parameters s and c of Eq. (13) in a way that is key for the developments of Section 4. Our present techniques differ substantially from those introduced in [33]. We begin by showing that the vector $\nu(\lambda\,,\,\mathbf{a})$ above is unique. Then, we show that the computation of $\nu(\lambda\,,\,\mathbf{a})$ for $m \in \mathbb{I}_{d-1}$ can be reduced to the case when $m = 0$. First we need to introduce some notations: $\triangle$

Definition 3.12. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. Assume that $d \leq k$. We denote by

$$\lambda = \lambda^\uparrow(S_{\mathcal{F}_0}) \quad \text{and} \quad h_i = \lambda_i + a_i \quad \text{for every} \quad i \in \mathbb{I}_d\ .$$

Given $j\,,\,r \in \mathbb{I}_d \cup \{0\}$ such that $j < r$, by $Q_{j\,,\,r}$ we denote the final averages:

$$Q_{j\,,\,r} = \frac{1}{r-j}\,\Big[\ \sum_{i=j+1}^{r}\ h_i + \sum_{i=r+1}^{k}\ a_i\Big] = \frac{1}{r-j}\,\Big[\ \sum_{i=j+1}^{k}\ a_i + \sum_{i=j+1}^{r}\ \lambda_i\Big]\ . \tag{14}$$

We shall abbreviate $Q_r = Q_{0\,,\,r}$. $\triangle$

Lemma 3.13. *Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP with $k \geq d$. Let $r \in \mathbb{I}_d$. Then*

1. *If $r < d$ and $Q_r < \lambda_{r+1}$ then $Q_r < Q_j$, for every j such that $r < j \leq d$.*

2. *If $r < d$ and $Q_r \leq \lambda_{r+1}$ then $Q_r \leq Q_j$, for every j such that $r < j \leq d$.*

3. *If $\lambda_r \leq Q_r$ then $Q_r \leq Q_j$, for every j such that $1 \leq j < r$.*

Proof. Denote by $c = Q_r$ for a fixed $r < d$. Recall that $\lambda = \lambda^\uparrow$. If $j > r$ then

$$c < \lambda_{r+1} \implies Q_j = \frac{1}{j}\left(\text{tr }\mathbf{a} + \sum_{i=1}^{r}\lambda_i + \sum_{i=r+1}^{j}\lambda_i\right) > \frac{1}{j}\,(\,r\,c + (j-r)\,c\,) = c\ .$$

The proof of item 2 is identical. On the other side, if $j < r$ then

$$\lambda_r \leq c \implies Q_j = \frac{1}{j}\left(\text{tr }\mathbf{a} + \sum_{i=1}^{r}\lambda_i - \sum_{i=j+1}^{r}\lambda_i\right) \geq \frac{1}{j}\,(\,r\,c - (r-j)\,c\,) = c\ . \qquad \square$$

Proposition 3.14. *Let* $(\mathcal{F}_0\,,\,\mathbf{a})$ *be initial data for the CP with* $k \geq d$ *(so that* $m \leq 0$*) and assume that* $\frac{1}{d}\,[\,\mathrm{tr}\ \mathbf{a} + \mathrm{tr}\ \lambda\,] < \lambda_d$*. Then*

1. *There exists a unique index* $s \in \mathbb{I}_d$ *such that* $\lambda_s \leq Q_s < \lambda_{s+1}$ *, and in this case*

$$s = \max\,\{w \in \mathbb{I}_{d-1} : Q_w = \min_{j\in\mathbb{I}_d}\ Q_j\,\} \quad \text{and} \quad \nu(\lambda\,,\,\mathbf{a}) = (Q_s\,\mathbb{1}_s\,,\,\lambda_{s+1}\,,\,\dots\,,\,\lambda_d) \tag{15}$$

2. *If another index* $r \in \mathbb{I}_{d-1}$ *satisfies that* $\lambda_r \leq Q_r \leq \lambda_{r+1}$ *, then*
   (a) $Q_r = \min_{j\in\mathbb{I}_d}\ Q_j = Q_s$ *and* $r \leq s$*.*
   (b) *If* $r < s$*, then* $Q_r = \lambda_{r+1} = \lambda_s$ *and also* $\nu(\lambda\,,\,\mathbf{a}) = (Q_r\,\mathbb{1}_r\,,\,\lambda_{r+1}\,,\,\dots\,,\,\lambda_d)$*.*
3. *Given* $\rho = (c\,\mathbb{1}_r\,,\,\lambda_{r+1}\,,\,\dots\,,\,\lambda_d)$ *(or* $\rho = c\,\mathbb{1}_d$*) such that* $\lambda \leqslant \rho = \rho^\uparrow$ *and* $\mathrm{tr}\ \rho = \mathrm{tr}\ \nu(\lambda\,,\,\mathbf{a})$ *then* $\rho = \nu(\lambda\,,\,\mathbf{a})$*.*

Proof. The existence of an index s such as in item 1 is guaranteed by the properties of $\nu(\lambda\,,\,\mathbf{a})$ stated in [33]. Nevertheless, it is easy to see that the index s described in Eq. (15) satisfies that $\lambda_s \leq Q_s < \lambda_{s+1}$. The formula given in Eq. (15), which shows the uniqueness of $\nu(\lambda\,,\,\mathbf{a})$, is a direct consequence of Lemma 3.13. Assume that $\lambda_r \leq Q_r \leq \lambda_{r+1}$. Then $Q_r = \min_{j\in\mathbb{I}_d}\ Q_j = Q_s$ and $r \leq s$ by Lemma 3.13. If $r < s$, then $Q_s = \frac{1}{s}\,(r\,Q_r + \sum_{i=r+1}^{s}\lambda_i) = Q_r$. This clearly implies all the equalities of item (b). Finally, observe that item 2 $\implies$ item 3. □

Remark 3.15. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP with $m = d - k > 0$. Then if

$$\tilde{\lambda} = (\lambda_1\,,\,\dots\,,\,\lambda_k) \in (\mathbb{R}^k)^\uparrow \quad \text{then} \quad \nu(\lambda\,,\,\mathbf{a}) = (\nu(\,\tilde{\lambda}\,,\,\mathbf{a})\,,\,\lambda_{k+1}\,,\,\dots\,,\,\lambda_d)\ , \tag{16}$$

and $\nu(\,\tilde{\lambda}\,,\,\mathbf{a})$ is constructed as in Proposition 3.14.

The proof is direct by observing that, extracting the entries $\lambda_{k+1}\,,\,\dots\,,\,\lambda_d$ of the vector $\nu(\lambda\,,\,\mathbf{a})$ as described in [33, Def. 4.13], the vector that one obtains (with the reverse order) satisfies the conditions of item 3 of Proposition 3.14 relative to the pair $(\,\tilde{\lambda}\,,\,\mathbf{a})$. △

The following result is in a sense a converse to Remark 3.10. It establishes that if there exists $f \in \mathrm{Conv_s}(\mathbb{R}_{\geq 0})$ and a minimizer $\mathcal{F} = (\mathcal{F}_0, \mathcal{G})$ of P_f in $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ such that the structure of the completing sequence $\mathcal{G}$ as described in Theorem 3.8 is trivial, the the underlying pair $(\lambda\,,\,\mathbf{a})$ is feasible. Recall the notation $\nu_f(\lambda\,,\,\mathbf{a})$ given in Theorem 3.7.

Lemma 3.16. *Let* $(\mathcal{F}_0\,,\,\mathbf{a})$ *be initial data for the CP,* $k \geq d$*, and let* $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}^{\mathrm{op}}(\mathcal{F}_0)$ *be a minimum for* P_f *on* $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ *for a* $f \in \mathrm{Conv_s}(\mathbb{R}_{\geq 0})$*. Suppose that, for some* $c > 0$*,*

$$W = R(S_{\mathcal{G}}) \neq \mathcal{H} \quad \text{and} \quad S_{\mathcal{F}}\big|_W \in L(W) = c\,I_W\ .$$

Let $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$*,* $\mu = \lambda^\downarrow(S_{\mathcal{G}})$ *and* $s \overset{def}{=} \dim W = \max\{i \in \mathbb{I}_d : \mu_i \neq 0\}$*. Then*

$$\lambda_s < c \leq \lambda_{s+1} \quad \text{so that} \quad (\lambda\,,\,\mathbf{a}) \quad \text{is feasible and} \quad \nu_f(\lambda\,,\,\mathbf{a}) = \nu(\lambda\,,\,\mathbf{a})\ .$$

The same final conclusion trivially holds if $s = \dim W = d$ *and* $S_{\mathcal{F}} = c\,I$*.*

Proof. Suppose that $s < d$. By hypothesis $\nu_f(\lambda\,,\,\mathbf{a}) = \lambda^\uparrow + \mu^\downarrow = \big(c\,\mathbb{1}_s\,,\,\lambda_{s+1}\,,\,\dots\,,\,\lambda_d\big)$ and it satisfies that $\lambda(S_{\mathcal{F}}) = \nu_f(\lambda\,,\,\mathbf{a})^\downarrow$. Since $\mathbf{a} \prec \mu = \mu^\downarrow$ then $\mathrm{tr}\ \mu = \mathrm{tr}\ \mathbf{a} > \sum_{i=1}^{s} a_i$, because $s < d \leq k$. Suppose now that $c > \lambda_{s+1}$. For small $t > 0$ consider the vector

$$\gamma(t) = \big(\,c\,\mathbb{1}_{s-1}\,,\,(c-t)\,,\,\lambda_{s+1} + t\,,\,\lambda_{s+2}\,,\,\dots\,,\,\lambda_d\big) \in \mathbb{R}^d \quad \text{with} \quad \mathrm{tr}\ \gamma(t) = \mathrm{tr}\ S_{\mathcal{F}}\ .$$

Let $\mu(t) = \gamma(t) - \lambda$. For every t we have that tr $\mu(t) = $ tr μ. On the other hand, if

$$t < \frac{\mu_s}{2} \implies \mu(t) = (\mu_1 \, , \, \dots \, , \, \mu_{s-1} \, , \, \mu_s - t \, , \, t \, , \, 0 \, \mathbb{1}_{d-s-1}) = \mu(t)^\downarrow \in (\mathbb{R}^d_{\geq 0})^\downarrow \, .$$

It is easy to see that if also $t < \sum_{i=s+1}^k a_i$ then still $\mathbf{a} \prec \mu(t)$. So there exists $\mathcal{F}' \in \mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0)$ such that $\lambda(S_{\mathcal{F}'}) = \gamma(t)^\downarrow$. Notice that, since $(c - t \, , \, \lambda_{s+1} + t) \prec (c \, , \, \lambda_{s+1})$ strictly, then $P_f(\mathcal{F}') =$ tr $f(\gamma(t)\,) <$ tr $f(\nu_f(\lambda \, , \, \mathbf{a})\,) = P_f(\mathcal{F})$, a contradiction. Hence $c \leq \lambda_{s+1}$.

The condition $\lambda_s < c$ follows from the fact that $c - \lambda_s = \mu_s > 0$. These facts show that $\lambda = \lambda^\uparrow \leqslant \nu_f(\lambda \, , \, \mathbf{a}) = \nu_f(\lambda \, , \, \mathbf{a})^\uparrow \implies \nu_f(\lambda \, , \, \mathbf{a}) = \nu(\lambda \, , \, \mathbf{a})$ (by item 3 of Proposition 3.14). In particular, $\mathbf{a} \prec \lambda(S_{\mathcal{G}}) = \mu = \nu(\lambda \, , \, \mathbf{a}) - \lambda = \mu(\lambda \, , \, \mathbf{a})$ so that $(\lambda \, , \, \mathbf{a})$ is feasible. □

4 Uniqueness and characterization of the minimum

In this section we shall state the main results of th paper. For the sake of clarity of the exposition, we postpone the more technical proofs until Section 5.

4.1 (Fixed data, notations and terminolgy)**.** Let $(\mathcal{F}_0 \, , \, \mathbf{a})$ be initial data for the CP. Until Theorem 4.8, we shall assume that $k \geq d$, so that $m = d - k \leq 0$. Recall that

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

Fix $f \in {\rm Conv_s}(\mathbb{R}_{\geq 0})$ and a minimizer $\mathcal{F} = (\mathcal{F}_0 \, , \, \mathcal{G}) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ be for P_f on $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$.

1. By Theorem 3.7, we know that $\mathcal{F} \in \mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0)$ and, if we denote by $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$, then $\lambda^\downarrow(S_{\mathcal{G}}) = \mu_f(\lambda \, , \, \mathbf{a}) = \nu_f(\lambda \, , \, \mathbf{a}) - \lambda$. By Theorem 3.6 there exists $\{v_i : i \in \mathbb{I}_d\}$ an ONB of eigenvectors for $S_{\mathcal{F}_0}$, λ such that

$$S_{\mathcal{G}} = \sum_{i \in \mathbb{I}_d} \mu_i \cdot v_i \otimes v_i \quad \text{ and } \quad S_{\mathcal{F}} = S_{\mathcal{F}_0} + S_{\mathcal{G}} = \sum_{i \in \mathbb{I}_d} (\lambda_i + \mu_i) \, v_i \otimes v_i \, . \tag{17}$$

2. Let $s_{\mathcal{F}} = \max \{ i \in \mathbb{I}_d : \mu_i \neq 0 \} = {\rm rk} \, S_{\mathcal{G}}$. Denote by $W = R(S_{\mathcal{G}})$, which reduces $S_{\mathcal{F}}$.

3. Let $S = S_{\mathcal{F}}\big|_W \in L(W)$ and $\sigma(S) = \{c_1 \, , \, \dots \, , \, c_p\}$ (where $c_1 > c_2 > \cdots > c_p > 0$).

4. Let $K_j = \{ i \in \mathbb{I}_{s_{\mathcal{F}}} : \lambda_i + \mu_i = c_j \}$ and $J_j = \{ i \in \mathbb{I}_k : S \, g_i = c_j \, g_i \}$. By Theorem 3.8,

$$\mathbb{I}_{s_{\mathcal{F}}} = \bigsqcup_{j \in \mathbb{I}_p} K_j \quad \text{ and } \quad \mathbb{I}_k = \bigsqcup_{j \in \mathbb{I}_p} J_j \, .$$

   Observe that $s_{\mathcal{F}}$, $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$ and the sets K_j completely describe the vector $\mu = \lambda(S_{\mathcal{G}})$.

5. Since $R(S_{\mathcal{G}}) = {\rm span}\{g_i : i \in \mathbb{I}_k\} = W = \oplus_{i \in \mathbb{I}_p} \ker (S - c_i \, I_W \,)$ then for every $j \in \mathbb{I}_p$,

$$W_j \stackrel{\rm def}{=} {\rm span}\{g_i : i \in J_j\} = \ker (S - c_j \, I_W \,) = {\rm span}\{v_i : i \in K_j\} \, , \tag{18}$$

   because $g_i \in \ker (S - c_j \, I_W \,)$ for every $i \in J_j$. Note that, by Theorem 3.8, each W_j reduces both $S_{\mathcal{F}_0}$ and $S_{\mathcal{G}}$.

6. If $p = 1$ then $J_1 = \mathbb{I}_k$ and $S = c_1 \, I_W$. Hence the minimum $\mathcal{F}$ satisfies the hypothesis of Lemma 3.16, so that the pair $(\lambda \, , \, \mathbf{a})$ is feasible.

7. We denote by $h_i = \lambda_i + a_i$ for every $i \in \mathbb{I}_d$. Given $j\,,\, r \in \mathbb{I}_d$ such that $j \le r$, let

$$P_{j\,,\,r} = \frac{1}{r-j+1}\ \sum_{i=j}^{r} h_i = \frac{1}{r-j+1}\ \sum_{i=j}^{r} \lambda_i + a_i\ ,$$

be the initial averages. We abbreviate $P_{1\,,\,r} = P_r$. △

Remark 4.2 (A reduction procedure)**.** Consider the data, notations and terminology fixed in 4.1. For any $j \in \mathbb{I}_{p-1}$ denote by

$$I_j = \mathbb{I}_d \setminus \bigcup_{i\le j} K_i\ \ ,\ \ L_j = \mathbb{I}_k \setminus \bigcup_{i\le j} J_i\ \ ,\ \ \lambda^{(j)} = (\lambda_i)_{i\in I_j}\ \ ,\ \ \mathcal{G}_j = (g_i)_{i\in L_j}\ ,\ \mathbf{a}^{(j)} = (a_i)_{i\in L_j}$$

and take some sequence $\mathcal{F}_0^{(j)}$ in $\mathcal{H}_j = \big[\bigoplus_{i\le j} W_i\big]^\perp$ such that $S_{\mathcal{F}_0^{(j)}} = S_0|_{\mathcal{H}_j}$ (notice that, by construction, $\mathcal{H}_j$ reduces S_0).

Then, it is straightforward to show that $\mathcal{F}_j = (\mathcal{F}_0^{(j)}\,,\,\mathcal{G}_j)$ is a (global) minimizer of P_f on $\mathcal{C}_{\mathbf{a}_j}(\mathcal{F}_0^{(j)})$ in $\mathcal{H}_j$, i.e. an optimal completion for the reduced problem. Indeed, recall that the minimality is computed in terms of the map F defined in Eq. (10), which works independently in each entry of $\lambda(S_{\mathcal{F}}) = \nu_f(\lambda\,,\,\mathbf{a})^\downarrow$.

The importance of the previous remarks lies in the fact that they provide a powerful reduction method to compute the structure of the sets $\mathcal{G}_i$, K_i and J_i for $i \in \mathbb{I}_p$ as well as the set of constants $c_1 > \ldots > c_p > 0$. Indeed, assume that we are able to describe the sets $\mathcal{G}_1$, K_1 , J_1 and the constant c_1 in some structural sense, using the fact that these sets are extremal (e.g. these sets are built on $c_1 > c_j$ for $2 \le j \le p$).

Then, in principle, we could apply these structural arguments to find $\mathcal{G}_2$, K_2 , J_2 and the constant c_2 , using the fact that these are now extremal sets of $\mathcal{F}_1$, which is a P_f minimizer of the reduced CP for $(\mathcal{F}_0^{(1)}, \mathbf{a}^{(1)})$. On the other hand, the minimality of the final reduction $\mathcal{F}_{p-1}$ produces a pair $(\lambda^{(p-1)}\,,\,\mathbf{a}^{(p-1)})$ which must be feasible by item 6 of 4.1, because it has an unique constant c_p associated to the unique set K_p . As we shall see, this strategy can be implemented to obtain (inductively) a precise description of the sets above. △

Remark 4.3. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP with $d \le k$, $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$ and $\mathbf{a} = \mathbf{a}^\downarrow$. Fix $f \in \mathrm{Conv_s}(\mathbb{R}_{\ge 0})$ and let $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}^{\mathrm{op}}(\mathcal{F}_0)$ be a global minimum for P_f on $\mathcal{C}_{\mathbf{a}}^{\mathrm{op}}(\mathcal{F}_0)$. In section 5.1 we shall prove the following properties (conjectured in [34]) of the sets J_j and K_j defined in item 4. of 4.1 describing $\mu_f(\lambda\,,\,\mathbf{a})$ and $\nu_f(\lambda\,,\,\mathbf{a})$:

1. Each set J_j and K_j consists of consecutive indexes, for $j \in \mathbb{I}_p$.

2. The sets K_j and J_j have the same number of elements, for $j \in \mathbb{I}_{p-1}$.

3. Moreover, $J_1 < \ldots < J_p$ (i.e. if $l \in J_i$ and $h \in J_j$ with $i < j \Rightarrow l < h$) and $K_1 < \ldots < K_p$. In particular, by items 1 and 2 above, $K_j = J_j$ for $j \in \mathbb{I}_{p-1}$.

△

We state the properties of the sets J_j and K_j, $j \in \mathbb{I}_p$ described in Remark 4.3 in the following:

Theorem 4.4. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP with $d \le k$. With the notations of Remark 4.3, assume that $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$, $\mu = \mu^\downarrow = \mu_f(\lambda\,,\,\mathbf{a})$ and $\mathbf{a} = \mathbf{a}^\downarrow$. Then

1. There exist $0 = s_0 < s_1 < s_2 < \cdots < s_{p-1} < s_p = s_{\mathcal{F}} = \max\{j \in \mathbb{I}_d : \mu_j \neq 0\}$ such that

$$K_j = J_j = \{s_{j-1}+1\,,\,\ldots\,,\,s_j\}\ , \quad , \quad \text{for} \quad j \in \mathbb{I}_{p-1}\ ,$$

$$K_p = \{s_{p-1}+1\,,\,\ldots\,,\,s_p\}\ ,\ J_p = \{s_{p-1}+1\,,\,\ldots\,,\,k\}\ .$$

2. The vector $\nu_f(\lambda\,,\,\mathbf{a}) = \big(c_1\,\mathbb{1}_{s_1}\,,\,\ldots\,,\,c_p\,\mathbb{1}_{s_p-s_{p-1}}\,,\,\lambda_{s_p+1}\,,\,\ldots\,,\,\lambda_d\big)$, where

$$c_r = \frac{1}{s_r - s_{r-1}} \sum_{i=s_{r-1}+1}^{s_r} h_i = P_{s_{r-1}+1\,,\,s_r} \quad \text{for} \quad r \in \mathbb{I}_{p-1}\ ,$$

   or also $c_r = \lambda_j + \mu_j$ for every $j \in K_r = J_r$ for $r \in \mathbb{I}_{p-1}$.

3. The constant c_p is the one defined by the feasible final part i.e., $c_p = Q_{s_{p-1}\,,\,s_p}$ and the indexes s_{p-1} and s_p are determined by the last block (recall Lemma 3.16).

Proof. See Section 5.2. ☐

Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. Assume that $\nu_f(\lambda\,,\,\mathbf{a}) = \big(c_1\,\mathbb{1}_{s_1}\,,\,\ldots\,,\,\lambda_{s_1+1}\,,\,\ldots\,,\,\lambda_d\big)$ i.e. with $p = 1$, in the notations of Theorem 4.4. Then, by Lemma 3.16, the pair $(\lambda\,,\,\mathbf{a})$ is feasible and $\nu_f(\lambda\,,\,\mathbf{a}) = \nu(\lambda\,,\,\mathbf{a})$.

In what follows we shall need the following notion, that allow us to show feasibility in the more general case in which, in the notations of Theorem 4.4, $p > 1$.

Definition 4.5. Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. Let $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0}) \in (\mathbb{R}^d_{>0})^\uparrow$. Suppose that $k \geq d$. Given $s \in \mathbb{I}_{d-1}$ denote by

$$\lambda^s = (\lambda_{s+1}\,,\,\ldots\,,\,\lambda_d) \in \mathbb{R}^{d-s} \quad \text{and} \quad \mathbf{a}^s = (a_{s+1}\,,\,\ldots\,,\,a_k) \in \mathbb{R}^{k-s}\ ,$$

the truncations of the original vectors λ and $\mathbf{a}$. We say that the index s is **feasible** if the pair $(\lambda^s\,,\,\mathbf{a}^s)$ is feasible for the CP. Note that $(d-s)-(k-s) = d-k = m \leq 0$. Therefore

$$\nu_s \overset{\text{def}}{=} \nu(\lambda^s\,,\,\mathbf{a}^s) \overset{(13)}{=} \big(c\,\mathbb{1}_{r-s}\,,\,\lambda_{r+1}\,,\,\ldots\,,\,\lambda_d\big) \quad \text{where} \quad c = Q_{s\,,\,r}$$

for the unique $r > s$ such that $\lambda_r \leq c < \lambda_{r+1}$ (or $\nu_s = Q_{s\,,\,d}\,\mathbb{1}_{d-s}$ if $\lambda_d \leq Q_{s\,,\,d}$). This means that $\lambda_s \leqslant \nu_s \in (\mathbb{R}^{d-s}_{>0})^\uparrow$. △

By Remark 4.2 and Lemma 3.16 we know that, with the notations of 4.1, the index s_{p-1} associated to the minimum $\nu = \nu_f(\lambda\,,\,\mathbf{a})$ is feasible - in the sense of Definition 4.5 - because the last block of ν is constructed with the final feasible parts of λ and $\mathbf{a}$, and $\nu_{s_{p-1}} = \big(c_p\,\mathbb{1}_{s_p-s_{p-1}}\,,\,\lambda_{s_p+1}\,,\,\ldots\,,\,\lambda_d\big)$.

Proposition 4.6. *Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. With the notations of Theorem 4.4, the global minimum $\nu_f(\lambda\,,\,\mathbf{a})$ satisfies that*

1. *The index s_{p-1} (where the feasible part begins) is determined by*

$$s_{p-1} = \min\ \{\ s \in \mathbb{I}_d\ :\ s\ \ \text{is feasible}\ \}\ .$$

2. *The following recursive method allow to describe the vector $\nu_f(\lambda\,,\,\mathbf{a})$ as in Theorem 4.4:*

   (a) The index $s_1 = \max\big\{j \leq s_{p-1}\ :\ P_{1\,,\,j} = \max_{i \leq s_{p-1}} P_{1\,,\,i}\big\}$, and $c_1 = P_{1\,,\,s_1}$.

   (b) If the index s_j is already computed and $s_j < s_{p-1}$, then

$$s_{j+1} = \max\left\{ s_j < r \le s_{p-1} \ : \ P_{s_j+1\,,\,j} = \max_{s_j < i \le s_{p-1}} P_{s_j+1\,,\,i} \right\} \quad \textit{and} \quad c_{j+1} = P_{s_j+1\,,\,s_{j+1}} \ .$$

Proof. See Propositions 5.14 and 5.11. □

The following are the main results of the paper. In order to state them, we introduce the spectral picture of the completions with prescribed norms, given by

$$\Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)\,) \stackrel{\text{def}}{=} \{\lambda(S_{\mathcal{F}}) : \ \mathcal{F} \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)\} \ .$$

Theorem 4.7. *Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP with $m = d - k \le 0$. Then the vector $\nu = \nu_f(\lambda\,,\,\mathbf{a})$ is the same for every $f \in \text{Conv}_\text{s}(\mathbb{R}_{\ge 0})$. Therefore,*

$$\nu^\downarrow \in \Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)\,) \quad \textit{and} \quad \nu^\downarrow \prec \gamma \quad \textit{for every} \quad \gamma \in \Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)\,) \ . \tag{19}$$

Proof. By Proposition 4.6, the minima $\nu = \nu_f(\lambda\,,\,\mathbf{a})$ are completely characterized by the data $(\lambda\,,\,\mathbf{a})$ without interference of the map f. Therefore, given any $\gamma \in \Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0))$,

$$\text{tr } f(\nu) \le \text{tr } f(\gamma) \quad \text{for every} \quad f \in \text{Conv}_\text{s}(\mathbb{R}_{\ge 0}) \implies \nu \prec \gamma \ . \qquad \square$$

The following result shows that the structure of optimal completions in $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ in case $m = d-k > 0$ can be obtained from the case in which $m = 0$.

Theorem 4.8. *Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP with $m = d - k > 0$. If we let*

$$\lambda' = (\lambda_1\,,\,\dots\,,\,\lambda_k) \in (\mathbb{R}_{\ge 0}^k)^\uparrow \quad \textit{then} \quad \nu_f(\lambda\,,\,\mathbf{a}) = (\nu_f(\lambda'\,,\,\mathbf{a})\,,\,\lambda_{k+1}\,,\,\dots\,,\,\lambda_d) \ ,$$

where $\nu_f(\lambda'\,,\,\mathbf{a})$ is constructed as in Proposition 4.6 (since $d' = k$, by construction of $\lambda' \in (\mathbb{R}_{\ge 0}^{d'})^\uparrow$). In this case the vector $\nu_f(\lambda\,,\,\mathbf{a})$ is the same for every $f \in \text{Conv}_\text{s}(\mathbb{R}_{\ge 0})$ and also satisfies Eq. (19).

Proof. Since $k = d - m$ and $\mathbf{a} \in \mathbb{R}^k$ we deduce that any $\delta = \delta^\downarrow \in \mathbb{R}_{\ge 0}^d$ such that $\mathbf{a} \prec \delta$ must have $\delta_{k+1} = \ldots = \delta_d = 0$. It is easy to see that this fact implies that

$$\Lambda_{\mathbf{a}}^{\text{op}}(\lambda) = \{\lambda^\uparrow + \delta^\downarrow : \ \delta \in \mathbb{R}_{\ge 0}^d \quad \text{and} \quad \mathbf{a} \prec \delta\} = \{(\gamma\,,\,\lambda_{k+1}\,,\,\dots\,,\,\lambda_d) : \gamma \in \Lambda_{\mathbf{a}}^{\text{op}}(\lambda')\} \tag{20}$$

We know that $\nu_f(\lambda\,,\,\mathbf{a}) - \lambda = \mu = \mu^\downarrow$ and that $\mathbf{a} \prec \mu \implies \mu_{k+1} = \ldots = \mu_d = 0$. Recall the map $F : \mathbb{R}_{\ge 0}^d \to \mathbb{R}$ defined in Eq. (10) for each $f \in \text{Conv}_\text{s}(\mathbb{R}_{\ge 0})$. Therefore

$$\nu_f(\lambda\,,\,\mathbf{a}) \in \Lambda_{\mathbf{a}}^{\text{op}}(\lambda) \quad \text{and} \quad \nu_f(\lambda\,,\,\mathbf{a}) = \mu^\downarrow + \lambda^\uparrow \implies \nu_f(\lambda\,,\,\mathbf{a}) = (\rho\,,\,\lambda_{k+1}\,,\,\dots\,,\,\lambda_d) \ , \tag{21}$$

for some $\rho \in \Lambda_{\mathbf{a}}^{\text{op}}(\lambda')$. Then $F(\nu_f(\lambda\,,\,\mathbf{a})\,) = F(\rho) + F(\lambda_{k+1}\,,\,\dots\,,\,\lambda_d)$. By Eq. (11),

$$F(\nu_f(\lambda\,,\,\mathbf{a})\,) \stackrel{(11)}{=} \min_{\nu \in \Lambda_{\mathbf{a}}^{\text{op}}(\lambda)} F(\nu) \stackrel{(20)}{=} \left[\min_{\gamma \in \Lambda_{\mathbf{a}}^{\text{op}}(\lambda')} F(\gamma) \right] + F(\lambda_{k+1}\,,\,\dots\,,\,\lambda_d) \ .$$

Using Eq. (11) again we deduce that $\rho = \nu_f(\lambda'\,,\,\mathbf{a})$. Since $\nu_f(\lambda'\,,\,\mathbf{a})$ is constructed as in Proposition 4.6, then it is the same vector for every strictly convex map f and the same happens with $\nu_f(\lambda\,,\,\mathbf{a})$, so that $\nu_f(\lambda\,,\,\mathbf{a})^\downarrow$ is a minimum for majorization on $\Lambda(\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)\,)$. □

Remark 4.9. The construction of the minimum $\nu_f(\lambda\,,\,\mathbf{a})$ given by Proposition 4.6 is algorithmic, an it can be easily implemented in Mathlab. It only depends on - an already available, see [33] - routine for checking feasibility, which is fast and efficient. △

5 Proofs of some technical results.

In this section we present detailed proofs of several statements in section 4. All these results assume that the initial data $(\mathcal{F}_0\,,\,\mathbf{a})$ for the CP satisfies that $k \geq d$. As already explained, the general case can be reduced to this situation.

5.1 Description of the sets K_i and J_i .

5.1. We begin by recalling the notations of 4.1: Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP, with $k \geq d$. Fix a convex map $f \in \text{Conv}_s(\mathbb{R}_{\geq 0})$. We consider the following objects:

1. Let $\mathcal{F} = (\mathcal{F}_0\,,\,\mathcal{G}) \in \mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)$ be a global minimum for P_f on $\mathcal{C}_{\mathbf{a}}^{\text{op}}(\mathcal{F}_0)$ (or a local minimum if $\mathcal{F}_0 = \varnothing$). If $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$ and $\mu \stackrel{\text{def}}{=} \lambda^\downarrow(S_{\mathcal{G}})$, then there exists $\{v_i : i \in \mathbb{I}_d\}$ an ONB of eigenvectors for $S_{\mathcal{F}_0}$, λ such that
$$S_{\mathcal{G}} = \sum_{i\in\mathbb{I}_d} \mu_i \cdot v_i \otimes v_i \quad \text{ and } \quad S_{\mathcal{F}} = S_{\mathcal{F}_0} + S_{\mathcal{G}} = \sum_{i\in\mathbb{I}_d} (\lambda_i + \mu_i)\, v_i \otimes v_i \ .$$
2. Let $s_{\mathcal{F}} = \max\{i \in \mathbb{I}_d : \mu_i \neq 0\} = \text{rk}\, S_{\mathcal{G}}$. Denote by $W = R(S_{\mathcal{G}})$, which reduces $S_{\mathcal{F}}$.
3. Let $S = S_{\mathcal{F}}\big|_W \in L(W)$ and $\sigma(S) = \{c_1\,,\,\dots\,,\,c_p\}$ (where $c_1 > c_2 > \cdots > c_p$).
4. Let $K_j = \{i \in \mathbb{I}_s : \lambda_i + \mu_i = c_j\}$ and $J_j = \{i \in \mathbb{I}_k : S\, g_i = c_j\, g_i\}$. Then
$$\mathbb{I}_{s_{\mathcal{F}}} = \bigcup_{j\in\mathbb{I}_p}^{D} K_j \quad \text{ and } \quad \mathbb{I}_k = \bigcup_{k\in\mathbb{I}_p}^{D} J_k \ .$$

We remark that, if $\mathcal{F}_0 = \varnothing$, these facts are still valid for local minima by Theorem 3.8. The next three Propositions give a complete proof of Theorem 4.4. The first of them justifies the convention that $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$. △

Proposition 5.2. *Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP with $\lambda = \lambda^\uparrow(\,S_{\mathcal{F}_0}\,)$, and consider the notations of 5.1. If $p > 1$, then*
$$i \in K_1 \quad \Longrightarrow \quad i < j \ (\Longrightarrow \ \lambda_i \leq \lambda_j\,) \quad \textit{for every} \quad j \in \bigcup_{r>1} K_r = \mathbb{I}_{s_{\mathcal{F}}} \setminus K_1 \ .$$
Inductively, by means of Remark 4.2, we deduce that all sets K_j consist on consecutive indexes, and that $K_i < K_j$ (in terms of their elements) if $i < j$.

Proof. Suppose that there are $i \in K_1$ and $j \in K_r$ (for some $r > 1$) such that $j < i$. Then $\lambda_j \leq \lambda_i$ and $\mu_i \leq \mu_j$. For $t > 0$ very small, let $\mu_i(t) = \mu_i - t > 0$ and $\mu_j(t) = \mu_j + t$. Consider the vector $\mu(t)$ obtained by changing in μ the entries μ_i by $\mu_i(t)$ and μ_j by $\mu_j(t)$. Observe that not necessarily $\mu(t) = \mu(t)^\downarrow$, but we are indeed sure that $c_1 > c_r$.

Nevertheless, by Remark 2.3, $(\mu_i\,,\,\mu_j) \prec (\mu_i(t)\,,\,\mu_j(t)\,) \implies \mathbf{a} \prec \mu \prec \mu(t)$. Therefore there exists $\mathcal{F}' = (\mathcal{F}_0\,,\,\mathcal{G}') \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ such that, using the ONB of Eq. (17),
$$S_{\mathcal{G}'} = \sum_{h\in\mathbb{I}_d} \mu_h(t) \cdot v_h \otimes v_h \quad \text{ and } \quad S_{\mathcal{F}'} = S_{\mathcal{F}_0} + S_{\mathcal{G}'} = \sum_{h\in\mathbb{I}_d} (\lambda_h + \mu_h(t)\,)\, v_h \otimes v_h \ .$$

Denote by $V = \text{span}\{v_i\,,\,v_j\}$, which reduces both $S_{\mathcal{F}}$ and $S_{\mathcal{F}'}$. Also $S_{\mathcal{F}'}|_{V^\perp} = S_{\mathcal{F}}|_{V^\perp}$. Considering the restrictions to V as operators in $L(V) \cong \mathcal{M}_2(\mathbb{C})$ we get that
$$\lambda(S_{\mathcal{F}'}|_V) = (\lambda_i + \mu_i(t)\,,\,\lambda_j + \mu_j(t)\,) = (c_1 - t\,,\,c_r + t) \prec (c_1\,,\,c_r) = \lambda(S_{\mathcal{F}}|_V) \quad \text{strictly} \ ,$$

for t small enough in such a way that $c_1 - t > c_r + t$, so that $(c_1 - t\,,\, c_r + t) = (c_1 - t\,,\, c_r + t)^{\downarrow}$. Then the map F of Eq. (10), considered both on $\mathbb{R}^2_{\geq 0}$ and $\mathbb{R}^d_{\geq 0}$, satisfies that

$$F\big(\,\lambda(S_{\mathcal{F}'}|_V)\,\big) < F\big(\,\lambda(S_{\mathcal{F}}|_V)\,\big) \implies P_f(\mathcal{F}') = F\big(\,\lambda(S_{\mathcal{F}'})\,\big) < F(\lambda(S_{\mathcal{F}})\,) = P_f(\mathcal{F})\ ,$$

a contradiction. The inductive argument follows from Remark 4.2. □

5.3. In the following two statements we assume that, for some $f \in \operatorname{Conv_s}(\mathbb{R}_{\geq 0})$, the sequence $\mathcal{F} = (\mathcal{F}_0\,,\, \mathcal{G}) \in \mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0)$ is a global minimum for P_f , or it is a local minimum if $S_{\mathcal{F}_0} = 0$ and $\lambda = 0$. In both cases 5.1 applies. △

Proposition 5.4. *Let $(\mathcal{F}_0\,,\, \mathbf{a})$ be initial data for the CP, and let $\mathcal{F} = (\mathcal{F}_0\,,\, \mathcal{G}) \in \mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0)$ as in 5.1 and 5.3. Suppose that $p > 1$. Given $h \in J_i$ and $l \in J_r$ then*

$$i < r \implies a_h - a_l \geq c_i - c_r > 0\ .$$

In particular, the sets J_i consist of consecutive indexes, and $J_1 < J_2 < \ldots < J_p$ (in terms of their elements). □

Proof. Let us assume that $i < r \in \mathbb{I}_p$, $h \in J_i$ and $l \in J_r$, but $l < h$ (even less: that $a_l \geq a_h$). Then

$$g_l \otimes g_l \leq S_{\mathcal{G}} \leq S_{\mathcal{F}} \quad \text{and} \quad S_{\mathcal{F}}\ g_l = c_r\, g_l \implies a_h = \|g_h\|^2 \leq \|g_l\|^2 = a_l \leq c_r < c_i\ .$$

We also know that $\langle g_l\,,\, g_h\rangle = 0$. Denote by $w_h = \frac{g_h}{\|g_h\|} = a_h^{-1/2}\, g_h$ and $w_l = \frac{g_l}{\|g_l\|} = a_l^{-1/2}\, g_l$. Let

$$g_h(t) = \cos(t)\ g_h + \sin(t)\ \|g_h\|\ w_l \quad \text{and} \quad g_l(t) = \cos(\gamma t)\ g_l + \sin(\gamma t)\ \|g_l\|\ w_h \quad \text{for} \quad t \in \mathbb{R}$$

for some convenient $\gamma > 0$ that we shall find later. Let $\mathcal{F}_\gamma(t)$ be the sequence obtained by changing in $\mathcal{F}$ the vectors g_h by $g_h(t)$ and g_l by $g_l(t)$, for every $t \in \mathbb{R}$. Notice that $\|g_h(t)\|^2 = a_h$ and $\|g_l(t)\|^2 = a_l$ for every $t \in \mathbb{R}$, so that all the sequences $\mathcal{F}_\gamma(t) \in \mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$.

Let $W = \operatorname{span}\{w_h\,,\, w_l\}$, a subspace which reduces $S_{\mathcal{F}}$ and $S_{\mathcal{F}_\gamma(t)}$. Note that $g_h(t), g_l(t) \in W$. In the matrix representation with respect to this basis of W we get that

$$g_h \otimes g_h = \begin{bmatrix} a_h & 0 \\ 0 & 0 \end{bmatrix} \begin{matrix} w_h \\ w_l \end{matrix} \quad , \quad g_h(t) \otimes g_h(t) = a_h \begin{bmatrix} \cos^2(t) & \cos(t)\,\sin(t) \\ \cos(t)\,\sin(t) & \sin^2(t) \end{bmatrix} \begin{matrix} w_h \\ w_l \end{matrix} \quad ,$$

$$g_l \otimes g_l = \begin{bmatrix} 0 & 0 \\ 0 & a_l \end{bmatrix} \begin{matrix} w_h \\ w_l \end{matrix} \quad \text{and} \quad g_l(t) \otimes g_l(t) = a_l \begin{bmatrix} \sin^2(\gamma\, t) & \cos(t)\,\sin(t) \\ \cos(t)\,\sin(t) & \cos^2(\gamma\, t) \end{bmatrix} \begin{matrix} w_h \\ w_l \end{matrix}$$

If we denote by $S(t) = S_{\mathcal{F}_\gamma(t)}$, we get that

$$S(t) = S_{\mathcal{F}} - g_h \otimes g_h - g_l \otimes g_l + g_h(t) \otimes g_h(t) + g_l(t) \otimes g_l(t)\ .$$

Therefore $S(t)|_{W^\perp} = S_{\mathcal{F}}|_{W^\perp}$. On the other hand, $S_{\mathcal{F}}|_W = \begin{bmatrix} c_i & 0 \\ 0 & c_r \end{bmatrix}$. Then

$$S(t)|_W = \begin{bmatrix} c_i + a_h\,(\cos^2(t) - 1) + a_l\,\sin^2(\gamma t) & a_h\,\cos(t)\,\sin(t) + a_l\,\cos(\gamma t)\,\sin(\gamma t) \\ a_h\,\cos(t)\,\sin(t) + a_l\,\cos(\gamma t)\,\sin(\gamma t) & c_r + a_h\,\sin^2(t) + a_l^2\,(\cos^2(\gamma t) - 1) \end{bmatrix} \overset{\text{def}}{=} A_\gamma(t)\ .$$

Note that $\operatorname{tr} A_\gamma(t) = c_i + c_r$ for every $t \in \mathbb{R}$. Therefore $\lambda(A_\gamma(t)\,) \prec (c_i\,,\, c_r)$ strictly $\iff \|A_\gamma(t)\|_2^2 < c_i^2 + c_r^2$. Hence we consider the map $m_\gamma : \mathbb{R} \to \mathbb{R}$ given by

$$m_\gamma(t) = \|A_\gamma(t)\|_2^2 = \operatorname{tr}\,(A_\gamma(t)^2) \quad \text{for every} \quad t \in \mathbb{R}\ .$$

Note that $S(0) = S_{\mathcal{F}} \implies m_\gamma(0) = c_i^2 + c_r^2$. We shall see that, for a convenient choice of γ, it holds that $m'_\gamma(0) = 0$ but $m''_\gamma(0) < 0$. This will contradict the (local) minimality of $\mathcal{F}$, because m_γ would have in this case a maximum at $t = 0$, so that $\lambda(A_\gamma(t)) \prec (c_i\, , c_r)$ strictly $\overset{(4)}{\implies} \lambda(S_{\mathcal{F}_\gamma(t)}) \prec \lambda(S_{\mathcal{F}})$ strictly $\implies P_f(\mathcal{F}_\gamma(t)) < P_f(\mathcal{F})$ for every t near 0.

Indeed, we first compute the derivatives of the entries a_{ij} of $A_\gamma(t)$:

$$\begin{array}{rcl} a'_{11} &=& -a_h \sin(2t) + \gamma\ a_l \sin(2\gamma t) \\ a'_{12} &=& a_h \cos(2t) + \gamma\ a_l \cos(2\gamma t) \\ a'_{22} &=& a_h \sin(2t) - \gamma\ a_l \sin(2\gamma t) \end{array} \quad \text{and} \quad \begin{array}{rcl} a''_{11} &=& 2\,[-a_h \cos(2t) + \gamma^2\ a_l \cos(2\gamma t)] \\ a''_{22} &=& 2\,[a_h \cos(2t) - \gamma^2 a_l \cos(2\gamma t)] \end{array} .$$

So $a'_{11}(0) = 0\,,\ a'_{22}(0) = 0$ and $a_{12}(0) = 0$. Then, for $i\, ,\, j \in \mathbb{I}_2$ we have that

$$(a_{ij}^2)'(0) = 2\ a_{ij}(0)\ a'_{ij}(0) = 0 \quad \text{and} \quad (a_{ij}^2)''(0) = 2\big(\,(a'_{ij})^2(0) + a_{ij}(0)\ a''_{ij}(0)\,\big)\ .$$

Therefore $(a_{11}^2)''(0) = 4c_i(-a_h + \gamma^2\, a_l)$, $(a_{12}^2)''(0) = 2\,(a_h + \gamma\, a_l)^2$ and $(a_{22}^2)''(0) = -4c_r\,(-a_h + \gamma^2\, a_l)$. We conclude that $m'_\gamma(0) = 0$ (for every $\gamma \in \mathbb{R}$) and that

$$m''_\gamma(0) = 4\,\Big[\,c_i(-a_h + \gamma^2\ a_l) + (a_h + \gamma\ a_l)^2 - c_r\ (-a_h + \gamma^2\ a_l)\Big]\ ,$$

which is quadratic polynomial on γ with discriminant (if we drop the factor 4) given by

$$D = a_h\ a_l\ \Big[\ a_h\ a_l - \big(\,a_l + (c_i - c_r)\big)\ \big(a_h - (c_i - c_r)\,\big)\ \Big]\ .$$

As we are assuming that $a_l \geq a_h$ then $D > 0$, because

$$\big(\,a_l + (c_i - c_r)\big)\ \big(a_h - (c_i - c_r)\,\big) = a_l\ a_h - (c_i - c_r)(a_l - a_h) - (c_i - c_r)^2 < a_l\ a_h\ .$$

Hence there exists $\gamma \in \mathbb{R}$ such that $m''_\gamma(0) < 0$. Observe that as long as $0 < (c_i - c_r)(a_l - a_h) + (c_i - c_r)^2$ ($\iff a_h - a_l < c_i - c_r$) we arrive at the same contradiction. □

The following result is inspired on some ideas from [1].

Proposition 5.5. *Let $(\mathcal{F}_0\, , \mathbf{a})$ be initial data for the CP, and let $\mathcal{F} = (\mathcal{F}_0\, , \mathcal{G}) \in \mathcal{C}^{\rm op}_{\mathbf{a}}(\mathcal{F}_0)$ as in 5.1 and 5.3. For every $j < p$, the subsequence $\{g_i\}_{i \in J_j}$ of $\mathcal{G}$ is linearly independent.*

Proof. Suppose that there exists $j \in \mathbb{I}_{p-1}$ such that $\{g_i\}_{i \in J_j}$ is linearly dependent. Hence there exists coefficients $z_l \in \mathbb{C}$, $l \in J_j$ (not all zero) such that $|z_l| \leq 1/2$ and

$$\sum_{l \in J_j} \overline{z_l}\ a_l\ g_l = 0\ . \tag{22}$$

Let $I_j \subseteq J_j$ be given by $I_j = \{l \in J_j :\ z_l \neq 0\}$ and let $h \in \mathbb{C}^d$ such that $\|h\| = 1$ and $S_{\mathcal{F}} h = c_p\, h$. For $t \in (-1, 1)$ let $\mathcal{F}(t) = (\mathcal{F}_0, \mathcal{G}(t))$ where $\mathcal{G}(t) = \{g_i(t)\}_{i \in \mathbb{I}_k}$ is given by

$$g_l(t) = \begin{cases} (1 - t^2\,|z_l|^2)^{1/2} g_l + t\, z_l\, a_l h & \text{if } \ l \in I_j \\ g_l & \text{if } \ l \in \mathbb{I}_k \setminus I_j \end{cases} .$$

Fix $l \in I_j$. Let $\text{Re}(A) = \frac{A + A^*}{2}$ denote the real part of each $A \in L(\mathcal{H})$. Then

$$g_l(t) \otimes g_l(t) = (1 - t^2\,|z_l|^2)\ g_l \otimes g_l + t^2\,|z_l|^2\, a_l^2\ h \otimes h + 2\,(1 - t^2\,|z_l|^2)^{1/2}\, t\ \ \text{Re}(h \otimes a_l\, z_l\, g_l)$$

Let $S(t)$ denote the frame operator of $\mathcal{F}(t)$ and notice that $S(0) = S_\mathcal{F}$. Note that

$$S(t) = S_\mathcal{F} + t^2 \sum_{l\in I_j} |z_l|^2 \left(-g_l \otimes g_l + a_l^2 \ h \otimes h\right) + R(t)$$

where $R(t) = 2 \sum\limits_{l\in I_j} (1 - t^2\,|z_l|^2)^{1/2}\, t \ \operatorname{Re}(h \otimes a_l\, z_l\, g_l)$. Then $R(t)$ is a smooth function such that

$$R(0) = 0 \quad , \quad R'(0) = \sum_{l\in I_j} \operatorname{Re}(h \otimes a_l\, z_l\, g_l) = \operatorname{Re}(h \otimes \sum_{l\in I_j} a_l\, z_l\, g_l) = 0 \ ,$$

and such that $R''(0) = 0$. Therefore $\lim\limits_{t\to 0} \ t^{-2} \ R(t) = 0$. We now consider

$$W = \operatorname{span} \left(\{g_l : \ l \in I_j\} \cup \{h\}\right) = \operatorname{span} \left\{ g_l : \ l \in I_j \right\} \perp \mathbb{C} \cdot h \ .$$

Then $\dim W = s + 1$, for $s = \dim \operatorname{span}\{g_l : \ l \in I_j\} \geq 1$. By construction, the subspace W reduces $S_\mathcal{F}$ and $S(t)$ for $t \in \mathbb{R}$, in such a way that $S(t)|_{W^\perp} = S_\mathcal{F}|_{W^\perp}$ for $t \in \mathbb{R}$. On the other hand

$$S(t)|_W = S_\mathcal{F}|_W + t^2 \sum_{l\in I_j} |z_l|^2 \left(-g_l \otimes g_l + a_l^2 \ h \otimes h\right) + R(t) = A(t) + R(t) \in L(W) \ , \qquad (23)$$

where we use the fact that the ranges of the selfadjoint operators in the second and third term in the formula above clearly lie in W. Then $\lambda\big(S_\mathcal{F}|_W \big) = \big(c_j \, \mathbb{1}_s \, , \, c_p \big) \in (\mathbb{R}_{>0}^{s+1})^\downarrow$ and

$$\lambda\Big(\textstyle\sum_{l\in I_j} |z_l|^2 g_l \otimes g_l \Big) \ = (\gamma_1 \, , \, \dots \, , \, \gamma_s \, , \, 0) \in (\mathbb{R}_{\geq 0}^{s+1})^\downarrow \quad \text{with} \quad \gamma_s > 0 \ ,$$

where we have used the definition of s and the fact that $|z_l| > 0$ for $l \in I_j$. Hence, for sufficiently small t, the spectrum of the operator $A(t) \in L(W)$ defined in (23) is

$$\lambda\big(A(t) \big) \ = \big(c_j - t^2\, \gamma_s \, , \, \dots \, , \, c_j - t^2\, \gamma_1 \, , \, c_p + t^2 \ \textstyle\sum_{l\in I_j} a_l^2\, |z_l|^2 \big) \in (\mathbb{R}_{\geq 0}^{s+1})^\downarrow \ ,$$

where we have used the fact that $\langle g_l \, , \, h \rangle = 0$ for every $l \in I_j$. Let us now consider

$$\lambda\big(R(t) \big) = \big(\delta_1(t) \, , \, \dots \, , \, \delta_{s+1}(t) \big) \in (\mathbb{R}_{\geq 0}^{s+1})^\downarrow \quad \text{for} \quad t \in \mathbb{R} \ .$$

Recall that in this case $\lim\limits_{t\to 0} t^{-2} \delta_j(t) = 0$ for $1 \leq j \leq s + 1$. Using Weyl's inequality on Eq. (23), we now see that $\lambda\big(S(t)|_W \big) \prec \lambda\big(A(t) \big) + \lambda\big(R(t) \big) \ \stackrel{\text{def}}{=} \ \rho(t) \in (\mathbb{R}_{\geq 0}^{s+1})^\downarrow$. We know that

$$\begin{aligned} \rho(t) \ &= \big(c_j - t^2\, \gamma_s + \delta_1(t) \, , \, \dots \, , \, c_j - t^2\, \gamma_1 + \delta_s(t) \, , \, c_p + t^2 \ \textstyle\sum_{l\in I_j} a_l^2\, |z_l|^2 + \delta_{s+1}(t) \big) \\ &= \Big(c_j - t^2\, (\gamma_s - \tfrac{\delta_1(t)}{t^2}) \, , \, \dots \, , \, c_j - t^2\, (\gamma_1 - \tfrac{\delta_s(t)}{t^2}) \, , \, c_p + t^2\, (\textstyle\sum_{l\in I_j} a_l^2\, |z_l|^2 + \tfrac{\delta_{s+1}(t)}{t^2}) \Big) \ . \end{aligned}$$

A direct test shows that, for small t, this $\rho(t) \prec \lambda(S_\mathcal{F}|_W) = \big(c_j \, \mathbb{1}_s \, , \, c_p \big)$ strictly. Then, since f is strictly convex, for every sufficiently small t we have that

$$P_f\big(\mathcal{F}(t) \big) \leq \operatorname{tr} \, f\big(\lambda(S_\mathcal{F}|_{W^\perp}) \big) + \operatorname{tr} \, f\big(\rho(t) \big) < \operatorname{tr} \, f\big(\lambda(S_\mathcal{F}|_{W^\perp}) \big) + \operatorname{tr} \, f\big(\lambda(S_\mathcal{F}|_W) \big) = P_f(\mathcal{F}) \ .$$

This last fact contradicts the assumption that $\mathcal{F}$ is a local minimizer of P_f in $\mathcal{C}_\mathbf{a}^{\text{op}}(\mathcal{F}_0)$. $\square$

Remark 5.6. Proposition 5.5 allows to show that in case $\mathcal{F}_0 = \emptyset$ then local and global minimizers of a convex potential P_f, induced by $f \in \text{Conv}_s(\mathbb{R}_{\geq 0})$, on $\mathcal{C}_{\mathbf{a}}(\mathcal{F}_0)$ - endowed with the product topology - coincide, as conjectured in [30].

Recall that a local minimizer $\mathcal{F}$ is a juxtaposition of tight frame sequences $\{\mathcal{F}_i\}_{i\in\mathbb{I}_p}$ which generate pairwise orthogonal subspaces of $\mathcal{H}$. Notice that by [34, Lemma 4.9] $\mathcal{F}$ is a frame for $\mathcal{H}$. Moreover, by Proposition 5.4, it is constructed using a partition of $\mathbf{a}$ with consecutive indexes.

Now by inspection of the proof of Proposition 5.5 we see that only one of such frame sequences can be a linearly dependent set: that with the smallest tight constant c_p. This forces that the (ordered) spectrum ν of a local minimizer must be either $\nu = c\mathbb{1}_d$ or

$$\nu = (a_1 , a_2 , \ldots , a_r , c , \cdots , c) , \qquad \text{where} \qquad a_r > c \geq a_{r+1} ,$$

and c is the constant of the unique tight subframe constructed with a linear dependent sequence of vectors with norms given by $\{a_i\}_{i=r+1}^k$ (notice that this forces $c \geq a_{r+1}$). But it is not difficult to see that this vector can be constructed in a unique way, that is, there is only one r such that

$$a_{r+1} \leq c = \frac{1}{d} \Big(\text{tr}(\mathbf{a}) - \sum_{i=1}^{r} a_i \Big) < a_r .$$

That is, the spectrum of local minimizers is unique and therefore local and global minimizers of P_f coincide, for every potential P_f as above. $\triangle$

5.2 Several proofs.

Let $(\mathcal{F}_0 , \mathbf{a})$ be initial data for the CP with $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$, $\mathbf{a} = \mathbf{a}^\downarrow$ and $d \leq k$. Recall that we denote by $h_i = \lambda_i + a_i$ for every $i \in \mathbb{I}_d$ and, given $j , r \in \mathbb{I}_d$ such that $j \leq r$, we denote by

$$P_{j\,,r} \ = \tfrac{1}{r-j+1} \sum_{i=j}^{r} h_i = \tfrac{1}{r-j+1} \sum_{i=j}^{r} \lambda_i + a_i .$$

We shall abbreviate $P_{1\,,r} = P_r$.

5.7 (Proof of Theorem 4.4). We rewrite its statement: Let $(\mathcal{F}_0 , \mathbf{a})$ be initial data for the CP with $d \geq k$. Let $\mathcal{F} = (\mathcal{F}_0 , \mathcal{G}) \in \mathcal{C}^{\text{op}}_{\mathbf{a}}(\mathcal{F}_0)$ be a **global** minimum for P_f on $\mathcal{C}^{\text{op}}_{\mathbf{a}}(\mathcal{F}_0)$. Using the notations of 4.1, assume that $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$, $\mu = \mu^\downarrow = \mu_f(\lambda , \mathbf{a})$ and $\mathbf{a} = \mathbf{a}^\downarrow$. Then

1. There exist indexes $0 = s_0 < s_1 < \cdots < s_{p-1} < s_p = s_{\mathcal{F}} = \max\{j \in \mathbb{I}_d : \mu_j \neq 0\}$ such that

$$\begin{aligned} K_j \ &= J_j = \{s_{j-1} + 1 , \ldots , s_j\} \ , \quad , \quad \text{for} \quad j \in \mathbb{I}_{p-1} \ , \\ K_p \ &= \{s_{p-1} + 1 , \ldots , s_p\} \ , \ J_p = \{s_{p-1} + 1 , \ldots , k\} \ . \end{aligned} \tag{24}$$

2. The vector $\nu_f(\lambda , \mathbf{a}) = \big(c_1 \mathbb{1}_{s_1} , \ldots , c_p \mathbb{1}_{s_p - s_{p-1}} , \lambda_{s_p+1} , \ldots , \lambda_d \big)$, where

$$c_r = \frac{1}{s_r - s_{r-1}} \sum_{i=s_{r-1}+1}^{s_r} h_i = P_{s_{r-1}+1\,,\,s_r} \quad \text{for} \quad r \in \mathbb{I}_{p-1} \ , \tag{25}$$

   or also $c_r = \lambda_j + \mu_j$ for every $j \in K_r = J_r$ for $r \in \mathbb{I}_{p-1}$.

3. The constant c_p is the one defined by the feasible final part i.e., $c_p = Q_{s_{p-1}\,,\,s_p}$ of (14) and the indexes s_{p-1} and s_p are determined by the last block, which is feasible.

Proof. Recall from Eq. (18) that for every $j \in \mathbb{I}_{p-1}$

$$W_j \overset{\text{def}}{=} \ker(S - c_j I_W) = \operatorname{span}\{v_i : i \in K_j\} = \operatorname{span}\{g_i : i \in J_j\}$$

By Proposition 5.5 $|J_j| = \dim W_j = |K_j|$ for $j < p$. Using now Propositions 5.2 and 5.4, we deduce that there exist indexes $0 = s_0 < s_1 < s_2 < \cdots < s_{p-1} < s_p = s_{\mathcal{F}} = \max\{j \in \mathbb{I}_d : \mu_j \neq 0\}$ such that the sets K_j and J_j satisfy Eq. (24). Using Eq. (18) again,

$$S_{\mathcal{G}}|_{W_j} = \sum_{i\in J_j} g_i \otimes g_i \implies \operatorname{tr} S_{\mathcal{G}}|_{W_j} = \sum_{i\in K_j} \mu_i = \sum_{i\in J_j} a_i \ . \tag{26}$$

Therefore $(s_j - s_{j-1})\, c_j = \operatorname{tr} S|_{W_j} = \operatorname{tr} S_{\mathcal{F}_0}|_{W_i} + \operatorname{tr} S_{\mathcal{G}}|_{W_i} = \sum_{i\in K_j} h_i$, for every $j < p$. Then the vector $\nu_f(\lambda\,,\mathbf{a}) = \big(c_1 \mathbb{1}_{s_1}\,, \ldots, c_p \mathbb{1}_{s_p - s_{p-1}}\,, \lambda_{s_p+1}\,, \ldots, \lambda_d\big)$, where the constants c_r are given by Eq. (25) for $r < p$. Item 3 follows from Remark 4.2 and Lemma 3.16. □

Lemma 5.8. *Let $(\mathcal{F}_0\,,\mathbf{a})$ be initial data for the CP with $k \geq d$. Given $m \in \mathbb{I}_d$,*

$$(a_j)_{j\in\mathbb{I}_m} \prec (P_m - \lambda_j)_{j\in\mathbb{I}_m} \iff P_m \geq P_i \quad \textit{for every} \quad i \in \mathbb{I}_m \iff P_{1\,,m} = \max_{i\in\mathbb{I}_m}\{P_{1\,,i}\} \ .$$

Proof. Straightforward. □

Remark 5.9. Let $(\mathcal{F}_0\,,\mathbf{a})$ be initial data for the CP with $k \geq d$ and recall the description of a minimum $\nu_f(\lambda\,,\mathbf{a})$ given in Theorem 4.4. As in Lemma 5.8 (or by an inductive argument using Remark 4.2) we can assure that for every $r \leq p-1$, the constants

$$c_r = P_{s_{r-1}+1\,,s_r} \geq P_{s_{r-1}+1\,,j} \quad \text{for every} \quad j \quad \text{such that} \quad s_{r-1}+1 \leq j \leq s_r \ . \tag{27}$$

It uses that $(a_j)_{j=s_{r-1}+1}^{s_r} \prec (\mu_j)_{j=s_{r-1}+1}^{s_r} = (c_r - \lambda_j)_{j=s_{r-1}+1}^{s_r}$, a consequence of Eq. (26). △

Lemma 5.10. *Let $(\mathcal{F}_0\,,\mathbf{a})$ be initial data for the CP. With the notations of Theorem 4.4, the global minimum $\nu_f(\lambda\,,\mathbf{a})$, its constants c_j and the indexes s_j (for $j \in \mathbb{I}_p$) satisfy the following properties:*

1. *Suppose that $p > 1$. For every $j \in \mathbb{I}_{p-1}$ such that $j > 1$, the constant c_j satisfies that*

$$c_j = P_{s_{j-1}+1\,,s_j} = \frac{1}{s_j - s_{j-1}} \sum_{i=s_{j-1}+1}^{s_j} h_i < \frac{1}{s_j}\sum_{i=1}^{s_j} h_i = P_{1\,,s_j} \ . \tag{28}$$

2. *Fix $j \in \mathbb{I}_{p-1}$ such that $j > 1$. Then*

$$P_{1\,,t} < P_{1\,,s_{j-1}} \quad \textit{for every} \quad s_{j-1} < t \leq s_{p-1} \ . \tag{29}$$

3. *In particular the averages $P_{1\,,s_j} = \dfrac{1}{s_j}\displaystyle\sum_{i=1}^{s_j} h_i < \dfrac{1}{s_{j-1}}\displaystyle\sum_{i=1}^{s_{j-1}} h_i = P_{1\,,s_{j-1}}$ for $2 \leq j \leq p-1$.*

Proof. The inequality of item 1 follows since

$$\begin{aligned}\sum_{i=1}^{s_j} h_i &= \sum_{i=1}^{s_1} h_i + \sum_{i=s_1+1}^{s_2} h_i + \ldots + \sum_{i=s_{j-1}+1}^{s_j} h_i \\ &= s_1\, c_1 + (s_2 - s_1)\, c_2 + \ldots + (s_j - s_{j-1})\, c_j > s_j\, c_j \ .\end{aligned}$$

Now we prove the inequality of Eq. (29): Given an index t such that $s_{j-1} < t \le s_j$,

$$\begin{aligned} t\,P_{1\,,\,t} &= s_{j-1}\,P_{1\,,\,s_{j-1}} + \textstyle\sum_{i=s_{j-1}+1}^{t} h_i \\ &= s_{j-1}\,P_{1\,,\,s_{j-1}} + (t-s_{j-1})\,\frac{1}{(t-s_{j-1})}\,\textstyle\sum_{i=s_{j-1}+1}^{t} h_i \\ &\overset{(27)}{\le} s_{j-1}\,P_{1\,,\,s_{j-1}} + (t-s_{j-1})\,c_j \\ &< s_{j-1}\,P_{1\,,\,s_{j-1}} + (t-s_{j-1})\,c_{j-1} \\ &\le s_{j-1}\,P_{1\,,\,s_{j-1}} + (t-s_{j-1})\,P_{1\,,\,s_{j-1}} = t\,P_{1\,,\,s_{j-1}}\ , \end{aligned}$$

where we used the fact that $c_{j-1} \le P_{1\,,\,s_{j-1}}$ for $1 \le j-1 \le p-1$, which follows from item 1. In particular we have proved item 3, and this also proves that Eq. (29) holds for $s_j < t \le s_{p-1}$. $\square$

Proposition 5.11. *Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. With the notations of Theorem 4.4, the global minimum $\nu = \nu_f(\lambda\,,\,\mathbf{a})$, its constants c_j and the indexes s_j (for $j \in \mathbb{I}_p$) satisfy the following properties: Suppose we know the index s_{p-1} , and that $p > 1$. Then we have a recursive method to reconstruct ν:*

1. *The index $s_1 = \max\big\{ j \le s_{p-1} \,:\, P_{1\,,\,j} = \max\limits_{i \le s_{p-1}} P_{1\,,\,i} \big\}$, and $c_1 = P_{1\,,\,s_1}$.*
2. *If we already compute the index s_j and $s_j < s_{p-1}$, then*
$$s_{j+1} = \max\big\{ s_j < r \le s_{p-1} \,:\, P_{s_j+1\,,\,r} = \max_{s_j < i \le s_{p-1}} P_{s_j+1\,,\,i} \big\} \quad \textit{and} \quad c_{j+1} = P_{s_j+1\,,\,s_{j+1}}\ .$$

Proof. The formula $P_{1\,,\,s_1} = \max\limits_{i \le s_{p-1}} P_{1\,,\,i}$ follows from Lemma 5.8 and Eq. (29) of Lemma 5.10, which also implies that s_1 must be the greater index (before s_{p-1}) satisfying this property.

The iterative program works by applying the last fact to the successive truncations of ν which are still minima in their neighborhood, by Remark 4.2. $\square$

Recall that $h_i = \lambda_i + a_i$ and that, for $0 \le j < r \le d$, we denoted by

$$Q_{j\,,\,r} = \frac{1}{r-j}\Big[\ \textstyle\sum_{i=j+1}^{r}\ h_i + \sum_{i=r+1}^{k}\ a_i\Big] = \frac{1}{r-j}\Big[\ \sum_{i=j+1}^{k}\ a_i + \sum_{i=j+1}^{r}\ \lambda_i\Big]\ , \tag{30}$$

and we abbreviate $Q_{1\,,\,r} = Q_r$. Recall also the notion of feasible indexes given in Definition 4.5: Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP with $\lambda = \lambda^\uparrow(S_{\mathcal{F}_0})$ and $k \ge d$. Given $s \in \mathbb{I}_{d-1}$ denote by $\lambda^s = (\lambda_{s+1}\,,\,\dots\,,\,\lambda_d) \in \mathbb{R}^{d-s}$ and $\mathbf{a}^s = (a_{s+1}\,,\,\dots\,,\,a_k)$, the truncations of the original vectors λ and $\mathbf{a}$. Recall that the index s is feasible if the pair $(\lambda^s\,,\,\mathbf{a}^s)$ is feasible for the CP. In any case we denote by

$$\nu_s = \nu(\lambda^s\,,\,\mathbf{a}^s) = \big(c\,\mathbb{1}_{r-s}\,,\,\lambda_{r+1}\,,\,\dots\,,\,\lambda_d\big) \quad \text{where} \quad c = Q_{s\,,\,r}$$

for the unique $r > s$ such that $\lambda_r \le c < \lambda_{r+1}$. This means that $\lambda_s \leqslant \nu_s \in (\mathbb{R}_{>0}^{d-s})^\uparrow$ and that $\operatorname{tr}\nu_s = \operatorname{tr}\,\lambda^s + \operatorname{tr}\,\mathbf{a}^s$.

Lemma 5.12. *Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. Fix an index $s \in \mathbb{I}_{d-1} \cup \{0\}$. Then*

1. *The index r associated to ν_s as in the previous notations is given by*
$$r = \max\,\{w \in \mathbb{I}_d : w > s \quad \textit{and} \quad Q_{s\,,\,w} = \min_{j>s}\ Q_{s\,,\,j}\ \}\ .$$
In other words, r is the unique index which satisfies: Given $j > s$,
$$Q_{s\,,\,r} < Q_{s\,,\,j} \quad \textit{if} \quad j > r \quad \textit{and} \quad Q_{s\,,\,r} \le Q_{s\,,\,j} \quad \textit{if} \quad j < r\ . \tag{31}$$

2. *Given an index $l \in \mathbb{I}_{d-1}$,*

$$l > s \quad \text{and} \quad Q_{s\,,\,l} < \lambda_{l+1} \implies l \geq r \ , \tag{32}$$

where r is the index associated to ν_s of item 1.

Proof. Item 1 follows from Proposition 3.14 applied to λ^s and $\mathbf{a}^s$.

Item 2 : Assume that $l < l+1 \leq r$. Then $Q_{s\,,\,l} < \lambda_{l+1} \leq \lambda_r \leq Q_{s\,,\,r}$. In this case

$$\begin{aligned}
\operatorname{tr} \lambda^s + \operatorname{tr} \mathbf{a}^s &\overset{(30)}{=} (l-s)\, Q_{s\,,\,l} \ + \sum_{i=l+1}^{d} \lambda_i \\
&= (l-s)\, Q_{s\,,\,l} \ + \sum_{l+1 \leq i \leq r} \lambda_i + \sum_{i=r+1}^{d} \lambda_i \\
&< (r-s)\, Q_{s\,,\,r} + \sum_{i=r+1}^{d} \lambda_i \overset{(30)}{=} \operatorname{tr} \lambda^s + \operatorname{tr} \mathbf{a}^s \ ,
\end{aligned}$$

a contradiction. Hence $l \geq r$. □

Proposition 5.13. *Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP which is not feasible, with $k \geq d$. Let*

$$s^* = \min \{ \ s \in \mathbb{I}_d \ : \ s \ \text{ is feasible} \ \} \ .$$

Let ν^* be constructed using the recursive method of Proposition 5.11, by using s^* instead of s_{p-1} (which can always be done). Then if we get the constants $c_1 > \ldots > c_{q-1}$, and we define c_q as the feasibility constant of λ^{s^*} and $\mathbf{a}^{s^*}$, then $c_{q-1} > c_q$.

Proof. For simplicity of the notations, by working with the pair $(\lambda^{s_{q-2}}\,,\,\mathbf{a}^{s_{q-2}})$, we can assume that $q = 2$. Denote by $s_1 = s^* < s_2$ and $c_1\,,\,c_2$ the indexes and constants given by:

$$c_1 = \frac{1}{s_1} \sum_{i=1}^{s_1} h_i = P_{1\,,\,s_1} \quad \text{and} \quad c_2 = Q_{s_1\,,\,s_2} = \frac{1}{s_2 - s_1} \left(\sum_{i=s_1+1}^{s_2} h_i + \sum_{i=s_2+1}^{k} a_i \right) \ , \tag{33}$$

and we must show that $c_1 > c_2$. Recall that $h_i = \lambda_i + a_i$. We can assume that:

- By Proposition 5.11, $c_1 \geq \frac{1}{p} \sum\limits_{i=1}^{p} h_i = P_{1\,,\,p}$ for every $p \in \mathbb{I}_{s_1}$.
- $c_2 \geq \frac{1}{p-s_1} \sum\limits_{i=s_1+1}^{p} h_i = P_{s_1+1\,,\,p}$ for every $s_1 + 1 \leq p \leq s_2$.
- $\lambda_{s_2} \leq c_2 < \lambda_{s_2+1}$,

where the second item follows by the feasibility of s^* and the last item says that c_2 is the feasible constant for the second block.

Suppose that $c_1 \leq c_2$ and we will arrive to a contradiction by showing that, in such case, the pair $(\lambda\,,\,\mathbf{a})$ would be feasible (that is, $s^* = 0$ or s_{q-2}). In order to do that, let

$$t \in \mathbb{I}_d \qquad \text{and} \qquad b \overset{\text{def}}{=} Q_t = \tfrac{1}{t} \ \left(\ \textstyle\sum_{i=1}^{t} h_i + \sum_{i=t+1}^{k} a_i \ \right)$$

be the unique constant such that $\lambda_t \leq b < \lambda_{t+1}$, which appears in $\nu(\lambda\,,\,\mathbf{a})$. Then

$$c \stackrel{\text{def}}{=} Q_{s_2} \quad = \frac{1}{s_2}\left(\sum_{i=1}^{s_2} h_i + \sum_{i=s_2+1}^{k} a_i\right) = \frac{1}{s_2}(s_1\, c_1 + (s_2 - s_1)\, c_2) \leq c_2 < \lambda_{s_2+1} \ .$$

By Eq. (32) we can deduce that $t \leq s_2$. Moreover, by item 1 of Lemma 5.12 we know that

$$b = Q_t \quad = \frac{1}{t}\left(\sum_{i=1}^{t} h_i + \sum_{i=t+1}^{k} a_i\right) \leq \frac{1}{p}\left(\sum_{i=1}^{p} h_i + \sum_{i=p+1}^{k} a_i\right) = Q_p \qquad \text{for every} \qquad p \in \mathbb{I}_d \ . \tag{34}$$

In particular, $b \leq c \leq c_2$. On the other side, $c_1 \leq b$. Indeed, if $\nu = \nu(\lambda\,,\,\mathbf{a})$ then

$$\lambda \leqslant \nu^* \quad \text{and} \quad t = \text{tr}\ \nu^* = \text{tr}\ \nu \implies \nu \prec \nu^* \implies b = \nu_1 \ \geq \ \nu_1^* = c_1 \ ,$$

because $c_1 \leq c_2 \implies \nu^* = (\nu^*)^\uparrow$ and since $\nu = \nu^\uparrow$ is the $\prec$-minimum of the set

$$\{\lambda^\uparrow(S) : S_{\mathcal{F}_0} \leq S \quad \text{and} \quad \text{tr}\ S = t\} = \{\rho = \rho^\uparrow : \lambda \leqslant \rho \quad \text{and} \quad \text{tr}\ \rho = t\} \ ,$$

by the remarks at the beginning of Section 3.3 and Proposition 3.14.

To show the feasibility, by Lemma 5.8 we must show that $b \geq P_{1\,,\,p}$ for every $p \in \mathbb{I}_t$. First, if we are in the case $t \leq s_1$, this is clear since $b \geq c_1 \geq P_{1\,,\,p}$ for every $p \leq s_1$. Finally, suppose that $t \geq s_1 + 1$. As before, $b \geq c_1$ implies $b \geq P_{1\,,\,p}$ for every $p \leq s_1$. On the other hand, if $s_1 < p \leq t$ then Lemma 5.12 applied to ν_{s_1} (whose "r" is s_2) assures that

$$c_2 < Q_{s_1\,,\,t} \quad \implies (t - s_1)\, c_2 \leq \sum_{i=s_1+1}^{t} h_i + \sum_{i=t+1}^{k} a_i \stackrel{(33)}{=} t\, b - s_1\, c_1 \ .$$

Since $p \leq t$ and $b \leq c_2$, this implies that $(p - s_1)\, c_2 \leq p\, b - s_1\, c_1$. Therefore

$$p\, P_{1\,,\,p} = s_1\, c_1 + (p - s_1)\, P_{s_1+1\,,\,p} \leq s_1\, c_1 + (p - s_1)\, c_2 \leq p\, b \ . \qquad \square$$

Proposition 5.14. *Let $(\mathcal{F}_0\,,\,\mathbf{a})$ be initial data for the CP. With the notations of Theorem 4.4, the global minimum $\nu_f(\lambda\,,\,\mathbf{a})$ satisfies that*

$$s_{p-1} = \min\ \{\ s \in \mathbb{I}_d \ : \ s \ \text{ is feasible}\ \} \ .$$

Proof. Denote by s^* the minimum of the statement. Since s_{p-1} is feasible (recall the remark after Definition 4.5), then $s^* \leq s_{p-1}$. On the other hand, let us construct the vector ν^* of Proposition 5.13, using the iterative method of Proposition 5.11 with respect to the index $s = s^*$, and the solution for the feasible pair $(\lambda^{s^*}\,,\,\mathbf{a}^{s^*})$ after s^*. Write $\nu^* = (\nu_1^*\,,\,\dots\,,\,\nu_s^*\,,\,c\,\mathbb{1}_{r-s}\,,\,\lambda_{r+1}\,,\,\dots\,,\,\lambda_d)$, where c is the constant of the feasible part of ν^*. Observe that Proposition 5.13 assures that $c < \min\{\nu_i^* : 1 \leq i \leq s\}$.

Using this fact and Proposition 5.11 it is easy to see that the vector $\mu = \nu^* - \lambda^\uparrow$ satisfies that $\mu = \mu^\downarrow$. On the other hand Lemma 5.8 and Remark 5.9 assure that $\mathbf{a} \prec \mu$ (using the majorization in each block and the fact that $\mathbf{a} = \mathbf{a}^\downarrow$). Then $\mu \in \Gamma_d(\mathbf{a})$ and $\nu^* \in \Lambda_{\mathbf{a}}^{\text{op}}(\lambda)$, the set defined in Eq. (9).

Moreover, in each step of the construction of the minimum $\nu = \nu_f(\lambda\,,\,\mathbf{a})$ we have to get the same index $s_j = s_j(\nu^*)$ of ν^* or there exists a step where the maximum which determines s_j (for $\nu_f(\lambda\,,\,\mathbf{a})$) satisfies that $s_j > s^*$ (in the eventual case in which $s_{p-1} > s^*$).

In both cases, we get that $\nu_i^* \le \nu_i$ for every index $1 \le i \le s^*$. Consider the subvector of ν^* given by $\rho = (\nu_1^*, \dots, \nu_s^*, \lambda_{r+1}, \dots, \lambda_d) \in \mathbb{R}^{s+d-r}$, and the respective part of $\nu_f(\lambda, \mathbf{a})$ given by $\xi = (\nu_1, \dots, \nu_s, \nu_{r+1}, \dots \nu_d)$. Since tr $\nu^* =$ tr ν, the previous remarks show that

$$\rho \leqslant \xi \implies \rho \prec_w \xi \implies (\rho, c\mathbb{1}_{r-s}) \prec (\xi, \nu_{s+1}, \dots, \nu_r),$$

where the final majorization follows using Lemma 4.6 of [33], which can be used since the constant $c < \min\{\nu_i^* : 1 \le i \le s\}$ by Proposition 5.13 (and because $c < \lambda_{r+1}$). Since majorization is invariant under rearrangements, we deduce that $\nu^* \prec \nu$.

Finally, using Theorem 3.7 we know that $\nu = \nu_f(\lambda, \mathbf{a})$ is the unique minimum for the map tr $f(\cdot)$ in the set $\Lambda_{\mathbf{a}}^{\rm op}(\lambda)$. This implies that $\nu^* = \nu$, and therefore $s_{p-1} = s^*$. □

References

[1] 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, 2006.

[2] 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.

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

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

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

[6] 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.

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

[8] J. Cahill, M. Fickus, D.G. Mixon, M.J. Poteet, N. Strawn, Constructing finite frames of a given spectrum and set of lengths, Appl. Comput. Harmon. Anal. (in press).

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

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

[11] 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.

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

[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, Proceedings of SPIE, 8138: 81380Q/1-8 (2011).

[21] M. Fickus, D. G. Mixon and M. J. Poteet, Constructing finite frames of a given spectrum, Finite frames, 55-107, Appl. Numer. Harmon. Anal. Birkhäuser/Springer, New York, 2013.

[22] 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.

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

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

[25] 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.

[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^o 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. 34 (2013), no. 2, 201-223.

[34] P. Massey, M. Ruiz and D. Stojanoff, Optimal completions of a frame, preprint. See arxiv version in arXiv:1206.3588.

[35] M. J. Poteet, Parametrizing finite frames and optimal frame completions, Doctoral thesis, Graduate School of Engineering and Management, Air Force Institute of Technology, Air University.