A theory of optimal convex regularization for low-dimensional recovery

Yann Traonmilin, Rémi Gribonval, Samuel Vaiter

Noname manuscript No.
(will be inserted by the editor)

A theory of optimal convex regularization for low-dimensional recovery

Yann Traonmilin, Rémi Gribonval and Samuel Vaiter

Received: date / Accepted: date

Abstract We consider the problem of recovering elements of a low-dimensional model from under-determined linear measurements. To perform recovery, we consider the minimization of a convex regularizer subject to a data fit constraint. Given a model, we ask ourselves what is the "best" convex regularizer to perform its recovery. To answer this question, we define an optimal regularizer as a function that maximizes a compliance measure with respect to the model. We introduce and study several notions of compliance. We give analytical expressions for compliance measures based on the best-known recovery guarantees with the restricted isometry property. These expressions permit to show the optimality of the ℓ^1-norm for sparse recovery and of the nuclear norm for low-rank matrix recovery for these compliance measures. We also investigate the construction of an optimal convex regularizer using the example of sparsity in levels.

Keywords inverse problems · convex regularization · low dimensional modeling · sparse recovery · low rank matrix recovery

Mathematics Subject Classification (2020) 65K10 · 90C25 · 49N45

1 Introduction

In a finite-dimensional Hilbert space $\mathcal{H}$ (with associated inner product $\langle \cdot, \cdot \rangle$, and norm $\| \cdot \|_{\mathcal{H}}$), we consider the observation model:

$$y = Mx_0 \tag{1}$$

where y is an m-dimensional vector of measurements, M is an under-determined linear operator (from $\mathcal{H} = \mathbb{C}^n$, or $\mathbb{R}^n$, to $\mathbb{C}^m$), and $x_0 \in \mathcal{H}$ is the unknown vector we want to recover. The problem of recovering x_0 from y is typically ill-posed. It is thus necessary to use prior information on x_0 to recover it with a guarantee of success.

Y. Traonmilin
Univ. Bordeaux, Bordeaux INP, CNRS, IMB, UMR 5251,F-33400 Talence
yann.traonmilin@math.u-bordeaux.fr

R. Gribonval
Univ Lyon, ENS de Lyon, UCBL, CNRS, Inria, LIP, F-69342 Lyon, France.

S. Vaiter
CNRS, Université Côte d'Azur, LJAD, Nice, France.

In this work, we suppose that x_0 belongs to a low-dimensional cone Σ (a positively homogeneous set, *i.e.,* for every $x \in \Sigma$ and $\lambda \geq 0$, $\lambda x \in \Sigma$) that models known properties of the unknown. Examples of such models include sparse as well as low-rank models and many of their generalizations. Note that in these examples the models belong to the slightly less general class of models that are (finite or infinite) unions of subspaces (homogeneous sets).

To recover x_0, a classical method is to solve the minimization problem

$$x^* \in \arg\min_{Mx=y} R(x) \tag{2}$$

where R is a function – the regularizer – that ensures the existence of the minimum.

Many works [18,12,30,11] give practical regularizers ensuring that $x^* = x_0$ for several low-dimensional models (in particular sparse and low-rank models, see [22] for a most complete review of these results). A practical regularizer is a function that permits the effective calculation of x^*. Without computational constraint, the best possible regularizer would be $R = \iota_\Sigma$: the characteristic function of Σ defined by $\iota_\Sigma(x) = 0$ if $x \in \Sigma$, $\iota_\Sigma(x) = +\infty$ otherwise (see Section 2 for a review of this fact). Unfortunately, this function is generally not convex (unless Σ itself is a convex set) and can lead to an intractable optimization problem in general, even though recent works show that using $R = \iota_\Sigma$ and a dedicated minimization technique is a possible route for certain particular low-dimensional models that can be smoothly embedded in $\mathbb{R}^n$ [16,34,35].

In this work, we focus on continuous convex regularizers that guarantee the existence of a minimizer x^* and the existence of practical optimization algorithms to perform minimization (2) such as the primal-dual method [13] (provided their proximity operators can be calculated). Note that convexity in itself is not sufficient to guarantee the practical feasibility of minimization (2) ($R(x)$ could be NP-hard to calculate, *e.g.,* the nuclear norm for tensors [23], and/or the proximal operator of R could be NP-hard to compute).

Under conditions on the measurement operator M that typically involve the number of measurements and its structure (*e.g.,* random for compressed sensing), the fact that $x_0 \in \Sigma$ permits to give recovery guarantees when the convex regularizer R is well chosen. For example, when $\Sigma = \Sigma_k$ is the set of k-sparse vectors in $\mathbb{R}^n$ and $R(\cdot) = \|\cdot\|_1$ (ℓ^1-norm), or when $\Sigma = \Sigma_r$ is the set of matrices of rank lower than r in $\mathbb{R}^{p\times p}$ and $R(\cdot) = \|\cdot\|_*$ (nuclear norm), x_0 can be recovered as long as the number of measurements is of the order of the dimension of the model (up to some log factors) : $m \geq O(k\log(n/k))$ for sparse recovery or $m \geq O(rp)$ for low rank recovery.

Our approach to provide these results is to exhibit a regularizer R for a given model set Σ and to give the best possible recovery guarantees for the pair (R, Σ). Recent works aim at giving guidelines to obtain guarantees as tight as possible for general sparse models and convex regularizers [14,2,41,36,3,26]. With such frameworks, it becomes possible to compare the performance of different regularizers. This leads naturally to the following question which we address in this work: what is the "best" convex regularizer to recover a given low-dimensional model Σ?

To tackle this problem, it is necessary to define the notion of "best" based on recovery guarantees. We propose different possibilities and follow one route that permits us to give optimality results in the sparse and low-rank cases and show the difficulties that arise when considering more complex generalized sparsity models. This work can be viewed as a way to give meaning to the expression "convexification" of a low-dimensional model, that is often used and rarely defined.

1.1 Related works

Low-complexity models induced by convex regularization. Many regularizers encountered in signal processing and machine learning are known to induce a specific type of model. Without aiming for exhaustivity, the use of the ℓ^1 norm [15] induces a sparse pattern in the solution, while group regularization with mixed $\ell^1 - \ell^2$ norms restricts this sparse pattern to satisfy a specific block structure [42]. More advanced model sets, such as low-rank matrices are linked to the use of the nuclear norm [20]. For a wide class of regularizers, including decomposable norms [10], decomposable M-estimator [27], atomic norms [14] and partly smooth functions [39,40], the connection between nonsmooth convexity and model space can be made explicit. Note that all these works take the following stance: given a convex regularizer R, what is the model set Σ induced by minimizing $R(x)$?

Convexification of combinatorial functions. Given a real function f, it is known that its biconjugate f^{**} is a convex and closed function, whatever the initial properties of f. For instance, if f is the constant function equal to 1 except in 0 – that is the counting function ℓ^0 in dimension 1 – restricted to $[-1, 1]$, *i.e.,*

$$f(x) = \begin{cases} 1 & \text{if } x \in [-1,1] \setminus \{0\}, \\ 0 & \text{if } x = 0, \\ +\infty & \text{otherwise,} \end{cases}$$

then its biconjugate is the absolute value $|\cdot|$ restricted to $[-1, 1]$. Unfortunately, this construction is harder to generalize on an unbounded domain or in higher dimension. For instance, the biconjugate of the ℓ^0 counting function not restricted to a bounded set is the constant 0. Of interest, we can mention convex closures of submodular functions (functions of $\{0, 1\}^p$) that can be calculated explicitly using the Lovász extension [5] and convex closure of almost convex functions [24].

Convexification of objective function Many works intent to find a convex proxy to a non-convex objective function. In [7], adding a Lagrangian term to the regularization of a constrained non-convex minimization permits to build an equivalent minimization problem that is convex locally. Another possibility is to try to perform a regularization by infimal regularization [8] for lower semicontinuous objective functionals. In [28], in a function space setting, Pock *et al.* propose a high dimensional lifting of the Lagrangian formulation of (2) where the data-fit functional is non-convex. In the context of non-convex polynomial optimization, Lasserre's hierarchies [25] are used to recast the original problem in a hierarchy of convex semi-definite positive problems which provide global convergence results. The drawback of this method is the computational cost that makes it impractical for high-dimensional problems. Finally, convex closure of submodular functions also permits to cast sparsity inducing objective functions (where the regularizer is a submodular function of the support) into convex problems [5]. Note that if one aims to find a non-convex, but continuous, regularization, several works of interest may be cited, such as the use of ℓ^p minimization [21], SCAD [19], or CEL0 [32]. Nevertheless, in this paper, we focus on convex functions.

1.2 Contributions

In this paper, we define notions of *compliance measures* between a low-dimensional model and a regularizer in finite dimension. The compliance of a function R for a model Σ is a function

$$R \mapsto A_\Sigma(R) \tag{3}$$

that quantifies the recovery capabilities of Σ with R and minimization (2).

An optimal regularizer for a model Σ is defined as a regularizer that maximizes the compliance measure. In this article, we focus on the maximization of compliance measures on the set $\mathcal{C}$ of coercive continuous convex regularizers over $\mathcal{H}$.

- We introduce compliance measures in **Section 2** using tight recovery guarantees: a good regularizer is a regularizer that permits the recovery of Σ as often as possible. We discuss the elementary properties of these measures and show that optimal coercive continuous convex regularizers can be found in the smaller class of atomic norms with atoms included in the model set. While such compliance measures can be optimized in basic toy examples, they require to be approximated in order to deal with sparse and low-rank models.
- We propose in **Section 3** compliance measures exploiting best known uniform recovery guarantees based on the restricted isometry property (RIP). We give explicit formulations of such recovery guarantees and show that, indeed, the ℓ^1-norm and the nuclear norm are optimal for sparse and low-rank recovery (respectively).
- We study the case of a generalized sparsity model in **Section 4**: sparsity in levels. We show how an optimal regularizer can be explicitly constructed in a small family of convex regularizers (ℓ^1-norm weighted by levels). This example shows the difficulty of calculating optimal regularizers for new low-dimensional models and opens many questions for the study of optimal regularizers.

1.3 Notations

In $\mathcal{H}$, we denote $S(1) := \{z \in \mathcal{H} : \|z\|_\mathcal{H} = 1\}$ the unit sphere with respect to $\|\cdot\|_\mathcal{H}$. Given a linear operator $M : \mathcal{H} \to \mathbb{C}^m$, we denote M^H its Hermitian adjoint.

For $\Sigma \subseteq \mathcal{H}$ an arbitrary set, we denote ι_Σ its charateristic function defined by $\iota_\Sigma(x) = 0$ if $x \in \Sigma$, $\iota_\Sigma(x) = +\infty$ otherwise. We denote $\mathcal{E}(\Sigma) := \mathbb{R}_+ \cdot \overline{\text{conv}}(\Sigma)$, where $\overline{\text{conv}}(\Sigma)$ is the closure of the convex hull of Σ. We define $\bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$. Given a function $f : \mathcal{H} \to \bar{\mathbb{R}}$, we denote by $\texttt{dom}(f)$ its domain, *i.e.,* the set $\texttt{dom}(f) := \{x \in \mathcal{H} : f(x) < +\infty\}$.

2 Optimal regularizer for a low dimensional model

In this section, starting from the characterization of exact recovery of a model Σ, we introduce the notion of compliance measure and associated optimal convex regularizer.

2.1 Characterization of exact recovery using descent cones

Before defining an optimal regularizer, we must characterize when Σ can be recovered by solving (2). The fact that a given $x_0 \in \Sigma$ is recovered by solving (2) with regularizer R (*i.e.,* that the solution x^* of (2) is unique and satisfies $x^* = x_0$ when $y := Mx_0$) is equivalent to the fact that $R(x + z) > R(x)$ for every $z \in \ker(M) \setminus \{0\}$ (see *e.g.,* [14]). To summarize this, we use the following definition of symmetrized descent cones.

Definition 1 ((Symmetrized) descent cones) Consider a regularizer $R : \mathcal{H} \to \bar{\mathbb{R}}$. For any $x \in \mathtt{dom}(R)$, the descent cone of R at x is

$$\mathcal{T}_R(x) := \{\gamma z : \gamma \in \mathbb{R}, z \in \mathcal{H}, R(x+z) \leq R(x)\}\,. \tag{4}$$

For any set $\Sigma \subset \mathtt{dom}(R)$, we define $\mathcal{T}_R(\Sigma) := \bigcup_{x\in\Sigma} \mathcal{T}_R(x)$.

Other definitions of descent cones (*e.g.,* non-symmetric like in [14]) could be used. The symmetrization facilitates technical derivations in the following and permits to characterize recovery as well. For ease of reading, in the following, symmetrized descent cones will be referred to as descent cones.
Recovery guarantees with a regularizer R for a linear operator M generally come in two flavors (recall that x^* is the result of minimization (2)):

- **Non-uniform recovery**: If $x_0 \in \Sigma$, then $x^* = x_0$ is equivalent to $\mathcal{T}_R(x_0) \cap \ker M = \{0\}$.
- **Uniform recovery**: "For all $x_0 \in \Sigma$, $x^* = x_0$" is equivalent to

$$\mathcal{T}_R(\Sigma) \cap \ker M = \{0\}. \tag{5}$$

In the literature, recovery guarantees are obtained when the measurement operator M fulfills sufficient conditions that imply these characterizations. Distinguishing these two types of recovery guarantees especially makes sense in the context of compressed sensing when M is chosen at random. Typical results are then of the form:

- **Non-uniform recovery**: Given $x_0 \in \Sigma$, with high probability on the draw of M, $x^* = x_0$.
- **Uniform recovery**: With high probability on the draw of M, $x^* = x_0$ for all $x_0 \in \Sigma$.

The main techniques to obtain recovery guarantees using a condition on the number of measurements differ largely between these two cases (see Section 3). In this work, we mostly focus on uniform recovery guarantees to stay in a fully deterministic setting. For such uniform recovery guarantees, we see that the only interactions that matter between the model set Σ, the regularizer R, and the measurement operator M are summarized by equation (5).

2.2 Compliance measures and optimal regularization

To define a notion of optimal regularizer, we simply propose to say that an optimal regularizer is a function that optimizes a (hopefully well-chosen) criterion. We call such a criterion, a *compliance measure* and specifically aim at defining it such that it should be maximized. The objective is to define a compliance measure that *quantifies* the recovery capabilities of a given regularizer R given a model set Σ.

Starting from the characterization of exact recovery, we see that the kernel of M heavily influences the recovery capability of R. If we had some knowledge that $M \in \mathcal{M}$ where $\mathcal{M}$ is a set of linear operators, we would want to define a compliance measure $A_{\Sigma,\mathcal{M}}(R)$ that tells us how good is a regularizer in these situations and to maximize it. Such maximization might yield a function R^* that depends on $\mathcal{M}$ (*e.g.,* in [32], when looking for tight continuous relaxation of the ℓ^0 penalty a dependency on M appears).

In the following, we propose a more *universal* notion of optimal convex regularizer that does not depend on a particular class of linear operators $\mathcal{M}$: we propose compliance measures $A_\Sigma(R)$ that depend only on the set Σ and on the regularizer R, and consider their maximization on some set of convex functions $\mathcal{C}$ (that are coercive and continuous, see Section 2.4):

$$\sup_{R\in\mathcal{C}} A_\Sigma(R). \tag{6}$$

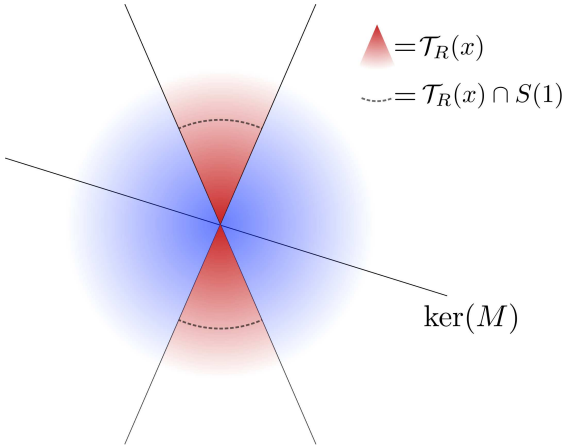

Fig. 1 A representation of recovery guarantees based on descent cones of a convex functions. Recovery of $x \in \Sigma$ fails if $\ker(M)$ intersects $\mathcal{T}_R(x)$ non trivially. The bigger is the descent cone (red) the more likely recovery will fail. The bigger the space left by the descent cone (blue), the more likely recovery will succeed

Of course, the existence of a maximizer of $A_\Sigma(R)$ is in itself a general question of interest: we could ask ourselves what conditions on $A_\Sigma(R)$ and $\mathcal{C}$ are necessary and sufficient for the existence of a maximizer, which is out of the scope of this article. In the sparse recovery and low-rank matrix recovery examples studied in this article, the existence of a maximizer of the considered compliance measures will be verified.

To build a compliance measure that does not depend on M, we define the optimal regularizer as the regularizer which guarantees recovery of Σ in as many situations as possible, *i.e.,* for "as many linear operators M as possible". Intuitively, a regularizer R is "good" if the set $\mathcal{T}_R(\Sigma)$ "leaves a lot of space" for $\ker M$ to not intersect it (trivially), see Figure 1). Among non-convex regularizers, the optimal one is the characteristic function of the model set Σ.

Lemma 1 *Consider an arbitrary non-empty set $\Sigma \subseteq \mathcal{H}$ and denote ι_Σ its characteristic function. The corresponding descent cone is*

$$\mathcal{T}_{\iota_\Sigma}(\Sigma) = \{\gamma z : \gamma \in \mathbb{R}, z \in \Sigma - \Sigma\} \supseteq \Sigma - \Sigma$$

where $\Sigma - \Sigma$ is the so-called secant set of Σ. For any regularizer R such that $\Sigma \subseteq \mathtt{dom}(R)$ we have $\mathcal{T}_{\iota_\Sigma}(\Sigma) \subseteq \mathcal{T}_R(\Sigma)$. Finally, if Σ is positively homogeneous then $\mathcal{T}_{\iota_\Sigma}(\Sigma) = \Sigma - \Sigma$.

Proof See Appendix A.2

This Lemma shows that ι_Σ is at least as successful as any regularizer R for the exact recovery of Σ (without any consideration of compliance measure). Moreover, $\mathcal{T}_{\iota_\Sigma}(\Sigma)$ is the smallest possible descent cone with respect to inclusion. Hence ι_Σ can be considered as the ideal regularizer [9] and indeed the optimal one with respect to any compliance measure defined as $A_\Sigma(R) = f(\mathcal{T}_R(\Sigma))$ where f is some function on subsets of $\mathcal{H}$ that is monotonic with respect to set inclusion. This is why the search for optimal regularizers only makes sense under some constraint on R.

2.3 A first compliance measure

As a first concrete example, we define here a theoretical compliance measure that reflects the idea that smaller descent cones are better. However, this compliance measure does not lead

to analytical expressions for the general study of sparse recovery. Our core results in the next sections rely on compliance measures based on recovery guarantees using the restricted isometry property (RIP).

Observe that exact recovery only depends on the orientation of the kernel of M, the natural uniform measure of orientation is the uniform measure on the unit sphere $S(1)$. In our setting, using this measure is a way of considering that we do not have prior information on the orientation of the kernel of M.

Using this measure, given a convex function R, the "amount of space left for the kernel of M" can be quantified by the "volume" of $\mathcal{T}_R(\Sigma) \cap S(1)$. Hence a compliance measure for uniform recovery can be defined as

$$A^U_\Sigma(R) := 1 - \frac{\mathrm{vol}\left(\mathcal{T}_R(\Sigma) \cap S(1)\right)}{\mathrm{vol}(S(1))}. \tag{7}$$

More precisely, here, the volume $\mathrm{vol}(E)$ of a set E is the measure of E with respect to the uniform measure on the sphere $S(1)$ (*i.e.,* the $n-1$-dimensional Hausdorff measure of $\mathcal{T}_R(\Sigma) \cap S(1)$). This measure is well defined as the descent cones of convex functions are symmetrized convex cones.

When looking at non-uniform recovery for random Gaussian measurements, the quantity defined by $\frac{\mathrm{vol}(\mathcal{T}_R(x_0) \cap S(1))}{\mathrm{vol}(S(1))}$ represents the probability that a randomly oriented kernel of dimension 1 with uniform probability on the sphere $S(1)$ intersects (non trivially) $\mathcal{T}_R(x_0)$. The highest probability of intersection with respect to x_0 quantifies the lack of compliance of R, hence we could define:

$$A^{NU}_\Sigma(R) := 1 - \sup_{x \in \Sigma} \frac{\mathrm{vol}\left(\mathcal{T}_R(x) \cap S(1)\right)}{\mathrm{vol}(S(1))}. \tag{8}$$

Note that this can be linked with the Gaussian width and statistical dimension theory of sparse recovery [14,2].

These compliance measures are completely dependent on the metric defining $S(1)$ (here the Hilbert norm $\|\cdot\|_{\mathcal{H}}$), other choices could be considered especially if measurement operators M showing a particular structure were considered.

In this article, we concentrate on compliance measures based on uniform recovery guarantees.

Remark 1 These compliance measures implicitly force $\Sigma \subset \mathtt{dom}(R)$, unless $A_\Sigma(R) = 0$. Indeed, suppose there exists $x \in \Sigma$ such that $R(x) = +\infty$, then for all $z \in \mathcal{H}$, we have $R(x+z) \leq +\infty = R(x)$. This implies $\mathcal{T}_R(x) = \mathcal{H}$ and $A^U_\Sigma(R) = A^{NU}_\Sigma(R) = 0$.

Remark 2 When studying convex regularization for low dimensional recovery in infinite dimensional separated Hilbert spaces, the noiseless recovery only depends on the behavior of the regularizer R on $\mathcal{E}(\Sigma)$ (defined in Section 1.3). The behavior of R outside of $\mathcal{E}(\Sigma)$ is only studied to obtain properties of robustness to modeling error [36]. In many examples of generalized sparsity and low-dimensional modeling in infinite dimension, the space $\mathcal{E}(\Sigma)$ has a finite dimension [1]. Our framework still applies in this case.

It is an open question to generalize our framework for low-dimensional recovery in more general settings such as Banach spaces (*e.g.,* for off-the-grid super-resolution).

2.4 Coercive continuous convex functions

As mentioned before we look for practical regularizers. We define $\mathcal{C}$ the set of all functions $R : \mathcal{H} \to \mathbb{R}$ (*i.e.,* with $\mathtt{dom}(R) = \mathcal{H}$) that are convex, continuous, and coercive.

The coercivity condition is typical in the context of convex regularization in order to avoid constant functions.

With any proper lower semi-continuous regularizer (hence, with any regularizer in $\mathcal{C}$) the convergence of the primal dual algorithm is guaranteed [13]. This guarantees the existence of practical algorithms (for the problem $\min_x \frac{1}{2}\|Mx-y\|^2 + \lambda R(x)$) when the proximity operator

$$y \mapsto \texttt{prox}_{\lambda R}(y) := \arg\min \tfrac{1}{2}\|u-y\|_{\mathcal{H}}^2 + \lambda R(u) \tag{9}$$

can be approximated efficiently.

2.5 Elementary properties and reduction to atomic "norms"

As compliance measures based on uniform recovery guarantees can be written as functions of descent cones $\mathcal{T}_R(\Sigma)$, we have the following immediate Lemma.

Lemma 2 *Let R_1, R_2 be two functions such that $\mathcal{T}_{R_1}(\Sigma) \subset \mathcal{T}_{R_2}(\Sigma)$ then $A^U_\Sigma(R_1) \geq A^U_\Sigma(R_2)$.*

In other words, the compliance measure is decreasing with respect to the inclusion of descent cones. We also verify that multiplying a regularizer by a scalar does not change the compliance measure which is consistent with recovery guarantees.

Lemma 3 *Let $\lambda > 0$. Then*

$$\begin{aligned} A^U_\Sigma(\lambda R) &= A^U_\Sigma(R), \\ A^{NU}_\Sigma(\lambda R) &= A^{NU}_\Sigma(R). \end{aligned} \tag{10}$$

Proof Let $x \in \Sigma$. We remark that

$$\begin{aligned} \mathcal{T}_{\lambda R}(x) &= \{\gamma z : \gamma \in \mathbb{R}, \lambda R(x+z) \leq \lambda R(x)\} \\ &= \{\gamma z : \gamma \in \mathbb{R}; R(x+z) \leq R(x)\} \\ &= \mathcal{T}_R(x). \end{aligned} \tag{11}$$

This shows directly that $A^{NU}_\Sigma(\lambda R) = A^{NU}_\Sigma(R)$. This also implies that $\mathcal{T}_{\lambda R}(\Sigma) = \mathcal{T}_R(\Sigma)$ and $A^U_\Sigma(\lambda R) = A^U_\Sigma(R)$.

□

More generally, any operation on R that leaves $\mathcal{T}_R(\Sigma)$ invariant does not change the compliance measure. This is typically the case of the post-composition of R with an increasing function.

We now recall the notion of atomic "norm" and show that optimal regularizers can be found in a set of atomic norms.

Definition 2 (Atomic norm) The atomic "norm" induced by a set $\mathcal{A}$ is defined as:

$$\|x\|_{\mathcal{A}} := \inf\{t \in \mathbb{R}_+ : x \in t \cdot \overline{\text{conv}}(\mathcal{A})\} \tag{12}$$

where $\overline{\text{conv}}(\mathcal{A})$ is the closure of the convex hull $\text{conv}(\mathcal{A})$ in $\mathcal{H}$. This "norm" is finite only on

$$\mathcal{E}(\mathcal{A}) := \mathbb{R}_+ \cdot \overline{\text{conv}}(\mathcal{A}) = \{x = t \cdot y, t \in \mathbb{R}_+, y \in \overline{\text{conv}}(\mathcal{A})\} \subset \mathcal{H}. \tag{13}$$

It is extended to $\mathcal{H}$ by setting $\|x\|_{\mathcal{A}} = +\infty$ if $x \notin \mathcal{E}(\mathcal{A})$.

Classical convex regularizer for sparse and low rank models are atomic norms:

– The ℓ^1-norm $\|\cdot\|_1$ is the atomic norm induced by signed canonical basis vectors.
– The nuclear norm $\|\cdot\|_*$ is the atomic norm induced by unitary rank one matrices.

Atomic norms are convex gauges induced by the convex hull of atoms. Their properties can be linked with the properties of the set $\mathcal{A}$ with classical results on convex gauge functions (see Appendix A.1).

It is possible to define an atomic norm, denoted $\|\cdot\|_\Sigma$, specifically induced by the model Σ.

Definition 3 (Atomic norm induced by the model) Given a cone Σ, we define

$$\|\cdot\|_\Sigma := \|\cdot\|_{\Sigma\cap S(1)}. \tag{14}$$

Known as the k-support norm for sparse model, it is not adapted to perform uniform recovery (it cannot recover 1-sparse vectors).

In [36, Lemma 2.1], it was shown that there is always an atomic norm with smaller descent cones than the descent sets of a proper coercive continuous function with convex level sets. We give a version of this result for our definition of cones and specify the properties of the obtained atomic norm.

Lemma 4 *Let Σ be a cone such that $\mathcal{E}(\Sigma) = \mathcal{H}$ and R be a coercive continuous convex function. Then there exists $\mathcal{A} \subset \Sigma$ such that:*

$$\mathcal{T}_{\|\cdot\|_\mathcal{A}}(\Sigma) \subseteq \mathcal{T}_R(\Sigma). \tag{15}$$

and $\|\cdot\|_\mathcal{A}$ is a coercive, continuous, positively homogeneous convex function.

Proof See Appendix A.2.2.

With Lemma 4, for all coercive continuous convex functions R it is possible to find an atomic norm R' with atoms included in Σ such that $\mathcal{T}_{R'}(\Sigma) \subset \mathcal{T}_R(\Sigma)$. We define

$$\mathcal{C}_\Sigma := \{\text{coercive continuous positively homogeneous atomic ``norms'' } \|\cdot\|_\mathcal{A} : \mathcal{A} \subset \Sigma\}. \tag{16}$$

As a consequence of this Lemma, we have the following immediate result.

Theorem 1 *Let Σ be a cone such that $\mathcal{E}(\Sigma) = \mathcal{H}$. Suppose A_Σ is a compliance measure that is a decreasing function of $\mathcal{T}_R(\Sigma)$ with respect to set inclusion. Then*

$$\sup_{R\in\mathcal{C}} A_\Sigma(R) = \sup_{R\in\mathcal{C}_\Sigma} A_\Sigma(R). \tag{17}$$

In particular

$$\sup_{R\in\mathcal{C}} A^U_\Sigma(R) = \sup_{R\in\mathcal{C}_\Sigma} A^U_\Sigma(R). \tag{18}$$

Proof Let $R \in \mathcal{C}$, with Lemma 4, there exists $\|\cdot\|_\mathcal{A} \in \mathcal{C}_\Sigma$ such that $\mathcal{T}_{\|\cdot\|_\mathcal{A}}(\Sigma) \subset \mathcal{T}_R(\Sigma)$. This implies $\mathcal{T}_{\|\cdot\|_\mathcal{A}}(\Sigma) \cap S(1) \subset \mathcal{T}_R(\Sigma) \cap S(1)$ and $A_\Sigma(R) \leq A_\Sigma(\|\cdot\|_\mathcal{A})$. □

Theorem 1 shows that we can limit ourselves to atomic norms if our only objective is to maximize the compliance measure.

With such measures, we can calculate optimal regularizers for elementary linear transformations of models.

Lemma 5 *Consider a compliance measure defined as: $A_\Sigma(R) := f(\mathcal{T}_R(\Sigma))$ with f some scalar valued function defined on non-empty subsets of $\mathcal{H}$. For any invertible linear map F on $\mathcal{H}$, any model set Σ and any regularizer R, we have*

$$\mathcal{T}_R(F\Sigma) = F(\mathcal{T}_{R\circ F}(\Sigma)) \tag{19}$$
$$A_{F\Sigma}(R) = f[F(\mathcal{T}_{R\circ F}(\Sigma))]. \tag{20}$$

Proof First $\gamma z \in \mathcal{T}_R(F\Sigma)$ if, and only if, there exists $x \in \Sigma$ such that $R(Fx + z) \leq R(Fx)$, *i.e.,* such that $(R \circ F)(x + F^{-1}z) \leq (R \circ F)(x)$. This is in turn equivalent to $\gamma F^{-1}z \in \mathcal{T}_{R\circ F}(\Sigma)$, *i.e.,* $\gamma z \in F(\mathcal{T}_{R\circ F}(\Sigma))$. Second, it follows that $A_{F\Sigma}(R) = f(\mathcal{T}_R(F\Sigma)) = f[F(\mathcal{T}_{R\circ F}(\Sigma))]$.

A consequence of Lemma 5 is that we can build optimal regularizers from other optimal regularizers when the model set is obtained from another one by applying a linear isometry.

Corollary 1 *Consider a compliance measure defined as: $A_\Sigma(R) := f(\mathcal{T}_R(\Sigma))$ with f some scalar valued function on subsets of $\mathcal{H}$. Assume that f is invariant to a family $\mathcal{F}$ of invertible linear maps on $\mathcal{H}$,* i.e., *for any $F \in \mathcal{F}$ and any non-empty set $\mathcal{T} \subseteq \mathcal{H}$, $f(F\mathcal{T}) = f(\mathcal{T})$. Then, for any $F \in \mathcal{F}$, any regularizer R and any model set Σ, we have*

$$A_{F\Sigma}(R \circ F^{-1}) = A_\Sigma(R). \tag{21}$$

Proof By Lemma 5, $A_{F\Sigma}(R\circ F^{-1}) = f[F(\mathcal{T}_{(R\circ F^{-1})\circ F}(\Sigma))] = f(F\mathcal{T}_R(\Sigma)) = f(\mathcal{T}_T(\Sigma)) = A_\Sigma(R)$. □

Corollary 2 *Consider F an isometry on $\mathcal{H}$, R a regularizer and Σ a model set. We have*

$$A^U_{F\Sigma}(R \circ F^{-1}) = A^U_\Sigma(R). \tag{22}$$

Proof The volume is invariant to isometries, hence $A^U_\Sigma(R) = f^U(\mathcal{T}_R(\Sigma))$ where $f^U(\cdot)$ is invariant to isometries. □

2.6 An exact result in 3D: the most we can do?

Compliance measures $A^U_\Sigma(R)$ and $A^{NU}_\Sigma(R)$ where effectively optimized [38] in the case of 1-sparse recovery in dimension 3, *i.e.,* for $\Sigma = \Sigma_1$ the set of 1-sparse vectors in $\mathbb{R}^3$. In this case, $\mathcal{C}_\Sigma = \{\|\cdot\|_\mathcal{A} : \mathcal{A} \subset \Sigma_1\}$. It was shown that for the set $\mathcal{C}'_\Sigma = \{\|\cdot\|_\mathcal{A} : \mathcal{A} \subset \Sigma_1, \mathcal{A} = -\mathcal{A}\}$ (which is the set of weighted ℓ^1-norms) that

$$\arg\max_{R\in\mathcal{C}'_\Sigma} A^U_\Sigma(R) = \arg\max_{R\in\mathcal{C}'_\Sigma} A^{NU}_\Sigma(R) = \{\lambda\|\cdot\|_1 : \lambda > 0\}. \tag{23}$$

To show this, the solid angles of all descent cones of arbitrary weighted ℓ^1-norms were calculated exactly and their size minimized with respect to the weights.

Unfortunately, calculating exactly these solid angles in dimension d seems out of reach for any sparsity and atomic norm in $\mathcal{C}_\Sigma$ even if some progress in bounds of these quantities [26] in some particular cases (non-uniform recovery with ℓ^1-norm in probability with random matrices). To the best of our knowledge, no general bound is available for the volume of descent cones of arbitrary atomic norms in $\mathcal{C}_\Sigma$ for sparse recovery. To build a compliance measure that we could optimize, we would need to first to establish such general bounds with some tightness.

In the next section, we propose to build compliance measures based on best-known uniform recovery guarantees that have some "tightness" properties. This will enable us to manipulate analytical expressions and give results for sparse recovery and low-rank recovery.

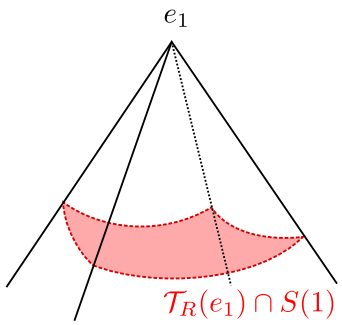

Fig. 2 Solid angle of a half descent cone of a weighted ℓ^1-norm

3 Compliance measures based on the restricted isometry property

The most used tool for the study of uniform recovery is the restricted isometry property (RIP). This property is adequate for multiple models [36], to be tight in some sense [17] for sparse and low-rank recovery, to be necessary in some sense [9], and to be well adapted to the study of random operators [29]. In [36], given a regularizer R, an explicit constant $\delta_\Sigma(R)$ is given, such that $\delta_\Sigma(M) < \delta_\Sigma(R)$ guarantees the exact recovery of elements of Σ by minimization (2). Hence, using $\delta_\Sigma(R)$ as a compliance measure, the higher the value of $\delta_\Sigma(R)$, the less stringent the RIP condition on M to ensure recovery of all elements of Σ using R as a regularizer.

To formalize further this idea, we start by recalling definitions and results about RIP recovery guarantees then apply our methodology. We also give results that emphasize the relevant quantity (depending on the geometry of the regularizer and the model) to optimize.

Definition 4 (RIP constant) Consider an arbitrary non-empty set $\Sigma \subset \mathcal{H}$ and M a linear map from $\mathcal{H}$ to $\mathbb{C}^m$. The RIP constant of M is defined as

$$\delta_\Sigma(M) = \sup_{x \in \Sigma - \Sigma} \left| \frac{\|Mx\|_2^2}{\|x\|_{\mathcal{H}}^2} - 1 \right|, \tag{24}$$

where $\Sigma - \Sigma$ (differences of elements of Σ) is called the secant set. For brevity we will simply denote $\delta(M)$ when the model set Σ is clear from context.

This coincides with the usual notion of RIP for sparse recovery when $\Sigma = \Sigma_k$ is the set of vectors with at most k nonzero entries (and $\Sigma - \Sigma = \Sigma_{2k}$); and with the RIP for low-rank recovery when $\Sigma = \Sigma_r$ is the set of matrices of rank at most r (then, $\Sigma - \Sigma = \Sigma_{2r}$).

A natural and sharp RIP-based compliance measure is $A_\Sigma^{RIP,\mathtt{sharp}}(R) = \delta_\Sigma^{\mathtt{sharp}}(R)$ defined as:

$$\delta_\Sigma^{\mathtt{sharp}}(R) := \inf_{M:\ker M \cap \mathcal{T}_R(\Sigma) \neq \{0\}} \delta_\Sigma(M). \tag{25}$$

This is the smallest RIP constant of measurement operators where uniform recovery fails [17], hence the following immediate theorem.

Theorem 2 *Consider an arbitrary non-empty set $\Sigma \subseteq \mathcal{H}$. Suppose M has RIP with constant $\delta_\Sigma(M) < \delta_\Sigma^{\mathtt{sharp}}(R)$, then for all $x_0 \in \Sigma$ and x^* the result of minimization (2) satisfies*

$$x^* = x_0. \tag{26}$$

Conversely, there exists M with $\delta_\Sigma(M) \geq \delta_\Sigma^{\mathtt{sharp}}(R)$ and $x_0 \in \Sigma$ such that $x^* \neq x_0$.

Despite this sharp property with respect to recovery, $\delta_\Sigma^{\texttt{sharp}}(R)$ is not endowed with any known analytic expression more explicit than its definition, and it is an open question to derive closed-form formulations of this constant for a general R, even for the particular case of sparse or low-rank models. This limits the possibility to conduct an exact optimization with respect to R, and motivates the investigation of more explicit RIP-based compliance measures, with two flavors:

- Compliance measures $\delta_\Sigma^{\texttt{nec}}(R)$ based on necessary RIP conditions [17] which yield sharp recovery constants for particular set of operators M, *e.g.,*

$$\delta_\Sigma^{\texttt{nec}}(R) := \inf_{z \in \mathcal{T}_R(\Sigma)\setminus\{0\}} \delta_\Sigma(I - \Pi_z). \tag{27}$$

  where Π_z is the orthogonal projection onto the one-dimensional subspace span(z) (other intermediate necessary RIP constants can be defined). Such measures upper bound $\delta_\Sigma^{\texttt{sharp}}(R)$ ($\delta_\Sigma^{\texttt{nec}}(R)$ characterizes RIP recovery guarantees of measurement operators having the shape $I - \Pi_z$).
- Compliance measures $\delta_\Sigma^{\texttt{suff}}(R)$ based on sufficient RIP constants for recovery (*e.g.,* the explicit sufficient RIP constant $\delta_\Sigma(R)$ from [36], which is explained in Section 3.3), which are lower bounds to $\delta_\Sigma^{\texttt{sharp}}(R)$.

Note that we have the relation

$$\delta_\Sigma^{\texttt{suff}}(R) \leq \delta_\Sigma^{\texttt{sharp}}(R) \leq \delta_\Sigma^{\texttt{nec}}(R). \tag{28}$$

The next sections aim at showing that the ℓ^1-norm (resp. the nuclear norm) maximizes the (best known) upper and lower bounds of $\delta_\Sigma^{\texttt{sharp}}(R)$ for k-sparse model (resp. low rank models).

First, in Section 3.1, we recall that when Σ is a non-trivial cone, the compliance measures associated to RIP constants can be cast to equivalent compliance measures associated to a restricted conditioning (RC), which can be written more explicitly.

Second, in Section 3.2, we use the expression of the RC-based compliance measure associated to $\delta_\Sigma^{\texttt{nec}}(\cdot)$ (from Equation (27)) to show that the ℓ^1 norm (resp. the trace-norm) optimizes $\delta_\Sigma^{\texttt{nec}}(\cdot)$ for k-sparse vectors (resp. for matrices of rank at most r), with $\delta_\Sigma^{\texttt{nec}}(R^\star) \approx 1/\sqrt{2}$ when n is large enough.

Finally, in Section 3.3, we give a characterization of $\delta_\Sigma^{\texttt{suff}}(R)$ and show the optimality of the ℓ^1-norm (resp. the nuclear norm) with $\delta_\Sigma^{\texttt{suff}}(R^\star) = 1/\sqrt{2}$.

3.1 Restricted conditioning as a compliance measure

When working with a model set Σ that is a cone, instead of working with actual RIP constants, it is easier to use (equivalently) the restricted conditioning.

Definition 5 (Restricted conditioning) Consider a cone $\Sigma \subset \mathcal{H}$ and a linear operator M from $\mathbb{R}^n$ to $\mathbb{C}^m$, and define

$$\gamma_\Sigma(M) := \frac{\sup_{x \in (\Sigma-\Sigma)\cap S(1)} \|Mx\|_2^2}{\inf_{x \in (\Sigma-\Sigma)\cap S(1)} \|Mx\|_2^2} \in [1, \infty] \tag{29}$$

where by convention here $a/0 = +\infty$ for any $a \geq 0$. For brevity we will simply denote $\gamma(M)$ when Σ is clear from context.

As shown below, the RIP constant $\delta_\Sigma(M)$ is a monotonically increasing function of $\gamma_\Sigma(M)$. The advantage of the latter is that it is invariant by rescaling M (rescaling leaves unchanged the recovery properties when measuring x_0 with M).

Lemma 6 *Consider a cone $\Sigma \subseteq \mathcal{H}$. For any M such that $\gamma_\Sigma(M) < \infty$, there is a unique $\lambda > 0$ such that*

$$\gamma_\Sigma(M) = \frac{1 + \delta_\Sigma(\lambda M)}{1 - \delta_\Sigma(\lambda M)} \tag{30}$$

or equivalently

$$\delta_\Sigma(\lambda M) = \frac{\gamma_\Sigma(M) - 1}{\gamma_\Sigma(M) + 1}. \tag{31}$$

Proof See Appendix A.3.

Thus, for cones, RIP-based compliance measures have equivalent RC-based compliance measures such that

$$\gamma_\Sigma(R) = \frac{1 + \delta_\Sigma(R)}{1 - \delta_\Sigma(R)} \qquad \text{and} \qquad \delta_\Sigma(R) = \frac{\gamma_\Sigma(R) - 1}{\gamma_\Sigma(R) + 1}. \tag{32}$$

The sharp RIP constant (25), the necessary RIP constant (27) and the sufficient RIP constant $\delta_\Sigma^{\mathtt{suff}}(R)$ of [36] are associated to

$$\gamma_\Sigma^{\mathtt{sharp}}(R) := \inf_{M:\ker M \cap \mathcal{T}_R(\Sigma) \neq \{0\}} \gamma_\Sigma(M) = \frac{1 + \delta_\Sigma^{\mathtt{sharp}}(R)}{1 - \delta_\Sigma^{\mathtt{sharp}}(R)}, \tag{33}$$

$$\gamma_\Sigma^{\mathtt{nec}}(R) := \inf_{z \in \mathcal{T}_R(\Sigma) \setminus \{0\}} \gamma_\Sigma(I - \Pi_z) = \frac{1 + \delta_\Sigma^{\mathtt{nec}}(R)}{1 - \delta_\Sigma^{\mathtt{nec}}(R)}, \tag{34}$$

$$\gamma_\Sigma^{\mathtt{suff}}(R) := \frac{1 + \delta_\Sigma^{\mathtt{suff}}(R)}{1 - \delta_\Sigma^{\mathtt{suff}}(R)}. \tag{35}$$

We deduce from (28) the inequalities

$$\gamma_\Sigma^{\mathtt{suff}}(R) \leq \gamma_\Sigma^{\mathtt{sharp}}(R) \leq \gamma_\Sigma^{\mathtt{nec}}(R). \tag{36}$$

The following result shows that $\gamma_\Sigma^{\mathtt{sharp}}(R)$ can be simplified.

Proposition 1 *Consider a cone $\Sigma \subseteq \mathcal{H}$. Let $\mathcal{P}$ be the set of symmetric positive semi-definite (PSD) linear operators on $\mathcal{H}$: $N \in \mathcal{P}$ if and only if $N^H = N$ and $N \succeq 0$. For $z \in \mathcal{H} \setminus \{0\}$ define*

$$f_\Sigma^{RC}(z) := \inf_{N \in \mathcal{P}:\ker N = \mathrm{span}(z)} \gamma_\Sigma(N) \tag{37}$$

and for any non-empty set $\mathcal{T} \subseteq \mathcal{H}$ such that $\mathcal{T} \neq \{0\}$ define

$$f_\Sigma^{RC}(\mathcal{T}) := \inf_{z \in \mathcal{T} \setminus \{0\}} f_\Sigma^{RC}(z). \tag{38}$$

We have

$$\inf_{M:\ker M \cap \mathcal{T} \neq \{0\}} \gamma_\Sigma(M) = f_\Sigma^{RC}(\mathcal{T}). \tag{39}$$

Proof This is an immediate consequence of Lemma 12 in Appendix A.3. □

Using $\mathcal{T} = \mathcal{T}_\Sigma(R)$, the sharp RC (or RIP) constant is the smallest RC constant of positive symmetric definite PSD operators with kernels of dimension 1 for which recovery of Σ fails:

$$\gamma_\Sigma^{\texttt{sharp}}(R) = f_\Sigma^{RC}(\mathcal{T}_R(\Sigma)). \tag{40}$$

Since $I - \Pi_z \in \mathcal{P}$ for any nonzero z, we have $f_\Sigma^{RC}(z) \leq \gamma_\Sigma(I-\Pi_z)$ hence we recover the inequality

$$\gamma_\Sigma^{\texttt{sharp}}(R) \leq \inf_{z \in \mathcal{T}_R(\Sigma)\setminus\{0\}} \gamma_\Sigma(I - \Pi_z) = \gamma_\Sigma^{\texttt{nec}}(R),$$

however it is an open question to determine whether this is an equality in particular cases or in general. The constant $\gamma_\Sigma^{\texttt{nec}}$ is already sharp in the following sense: for sparse recovery with the ℓ^1-norms, as well as for low-rank recovery with the nuclear norm, it is the biggest possible RIP constant ($\delta_\Sigma^{\texttt{suff}}(R) = \frac{1}{\sqrt{2}}$) that guarantees uniform recovery with $\|\cdot\|_1$ (respectively with the nuclear norm) for *all* sparsities k (for all rank levels r respectively) [17].

Similarly, to the compliance measures from Section 2, we can deduce an optimal regularizer after an isometric linear transformation of the model.

Lemma 7 *Consider a cone $\Sigma \subseteq \mathcal{H}$, an arbitrary regularizer R such that $\Sigma \subseteq \texttt{dom}(R)$, and a (linear) isometry F. We have*

$$\gamma_{F\Sigma}^{\texttt{sharp}}(R \circ F^{-1}) = \gamma_\Sigma^{\texttt{sharp}}(R). \tag{41}$$

Hence, for any class $\mathcal{C}'$ of regularizers,

$$R^* \in \arg\max_{R \in \mathcal{C}'} \gamma_\Sigma^{\texttt{sharp}}(R) \Leftrightarrow R^* \circ F^{-1} \in \arg\max_{R' \in \mathcal{C}'} \gamma_{F\Sigma}^{\texttt{sharp}}(R'). \tag{42}$$

Proof See Appendix A.3.

3.2 Compliance measures based on necessary RC conditions

In this section, we characterize the compliance measure

$$\gamma_\Sigma^{\texttt{nec}}(R) = \inf_{z \in \mathcal{T}_R(\Sigma)\setminus\{0\}} \gamma_\Sigma(I - \Pi_z). \tag{43}$$

To show optimality of the ℓ^1-norm for sparse recovery and of the nuclear norm for low-rank recovery, we will use the following characterization of $\gamma_\Sigma^{\texttt{nec}}(R)$ when Σ is a cone.

Lemma 8 *Consider a cone $\Sigma \subseteq \mathcal{H}$ such that $\Sigma \neq \{0\}$ and R an arbitrary regularizer such that $\Sigma \subseteq \texttt{dom}(R)$.*

1. *If there is $x \in \mathcal{H}$ such that $\Sigma \subseteq \text{span}(x)$, then*

$$\gamma_\Sigma^{\texttt{nec}}(R) = \begin{cases} +\infty & \text{if } \mathcal{T}_R(\Sigma) \subseteq \Sigma, \\ 1 & \text{otherwise.} \end{cases} \tag{44}$$

2. *If $\Sigma \subsetneq \text{span}(x)$ for every $x \in \mathcal{H}$, then*

$$\gamma_\Sigma^{\texttt{nec}}(R) = \frac{1}{1 - \inf_{z \in \mathcal{T}_R(\Sigma)\setminus\{0\}} \sup_{x \in (\Sigma-\Sigma)\cap S(1)} \frac{\langle x, z\rangle^2}{\|z\|_{\mathcal{H}}^2}}. \tag{45}$$

Proof See Appendix A.4.

To go further, we exploit two assumptions related to orthogonal projections on certain sets.

Definition 6 (Orthogonal projection) For any set E we define, for all $z \in \mathcal{H}$

$$P_E(z) = \arg\min_{y \in E} \|z - y\|_{\mathcal{H}}. \tag{46}$$

This is a set-valued operator, and $P_E(z)$ may be empty if the minimum is not achieved.

Some assumptions on E ensure that $P_E(z)$ is not empty for any z.

Lemma 9 *Consider a union of subspaces $E \subseteq \mathcal{H}$, and assume that $E \cap S(1)$ is compact. Then for every $z \in \mathcal{H}$, $P_E(z) \neq \emptyset$. Moreover, for every $x, x' \in P_E(z)$ we have $\|z-x\|_{\mathcal{H}}^2 = \|z-x'\|_{\mathcal{H}}^2$ and $\langle z, x\rangle = \|x\|_{\mathcal{H}}^2 = \|x'\|_{\mathcal{H}}^2 = \langle z, x'\rangle$, hence the notations $\|z - P_E(z)\|_{\mathcal{H}}^2$, $\langle z, P_E(z)\rangle$ and $\|P_E(z)\|_{\mathcal{H}}^2$ are unambiguous. We also have $\|z\|_{\mathcal{H}}^2 = \|z - P_E(z)\|_{\mathcal{H}}^2 + \|P_E(z)\|_{\mathcal{H}}^2$ and*

$$\langle z, P_E(z)\rangle = \|P_E(z)\|_{\mathcal{H}}^2 = \sup_{x \in E \cap S(1)} |\langle x, z\rangle|^2.$$

Proof See Appendix A.4.

Even if E is a union of subspaces and $E \cap S(1)$ is compact, $P_E(z)$ may not always be a singleton. For example, consider E the set of k-sparse vectors and z the vector with all entries equal to one.

Thanks to Lemma 9, we have the following characterization of the maximizers of $\delta_{\Sigma}^{\mathtt{nec}}$.

Corollary 3 *Consider a cone $\Sigma \subset \mathcal{H}$ and assume that $\Sigma - \Sigma$ is a union of subspaces, $(\Sigma - \Sigma) \cap S(1)$ is compact, and $\Sigma \neq \operatorname{span}(x)$ for each $x \in \Sigma$. For any class $\mathcal{C}'$ of regularizers such that $\Sigma \subseteq \mathtt{dom}(R)$ for every $R \in \mathcal{C}'$, the set of maximizers of $\delta_{\Sigma}^{\mathtt{nec}}(\cdot)$ satisfies (whether or not this set of maximizers is empty)*

$$\arg\max_{R \in \mathcal{C}'} \delta_{\Sigma}^{\mathtt{nec}}(R) = \arg\min_{R \in \mathcal{C}'} B_{\Sigma}(R) \qquad \textit{with} \qquad B_{\Sigma}(R) := \sup_{z \in \mathcal{T}_R(\Sigma) \setminus \{0\}} \frac{\|z - P_{\Sigma-\Sigma}(z)\|_{\mathcal{H}}^2}{\|P_{\Sigma-\Sigma}(z)\|_{\mathcal{H}}^2}. \tag{47}$$

For any optimal regularizer $R^\star$ we have

$$\delta_{\Sigma}^{\mathtt{nec}}(R^\star) = (1 + 2B_{\Sigma}(R^\star))^{-1}. \tag{48}$$

Proof See Appendix A.4.

We now have the tools to study optimality for sparse and low rank models.

Optimal regularization for sparse recovery and for low-rank recovery

Consider now $\Sigma = \Sigma_k$ the set of k-sparse vectors in $\mathcal{H} = \mathbb{R}^n$ (associated with the canonical scalar product $\langle \cdot, \cdot \rangle$ and the ℓ^2-norm $\|\cdot\|_{\mathcal{H}} = \|\cdot\|_2$), where $1 \leq k \leq n/2$, $n \geq 3$. We have $\Sigma - \Sigma = \Sigma_{2k}$ (for $n \leq 2k$, in particular for $n \leq 2$ and any $k \geq 1$, uniform recovery is not possible for non-invertible M: as $\Sigma - \Sigma = \mathbb{R}^n$, by Lemma 1 we have $\mathcal{T}_R(\Sigma) = \mathbb{R}^n$ for any regularizer, thus $\mathcal{T}_R(\Sigma) \cap \ker M = \{0\}$ implies $\ker M = \{0\}$). It is well-known that Σ and $\Sigma - \Sigma$ are both unions of subspaces (hence Σ is a cone), and it is straightforward that $(\Sigma - \Sigma) \cap S(1)$ is compact and Σ is not reduced to a single line. Moreover, for any nonzero $z \in \mathbb{R}^n$, $P_{\Sigma-\Sigma}(z)$ contains the restriction z_{T_2} of z to any set $T_2 = T_2(z) \subseteq \{1, \ldots, n\}$ of size $2k$ such that $\min_{i \in T_2} |z_i| \geq \max_{j \in T_2^c} |z_j|$. Hence, we can invoke Corollary 3 to replace the maximization of $\delta_{\Sigma}^{\mathtt{nec}}(R)$ by the minimization of

$$B_{\Sigma}(R) = \sup_{z \in \mathcal{T}_R(\Sigma) \setminus \{0\}} \frac{\|z_{T_2^c}\|_2^2}{\|z_{T_2}\|_2^2}. \tag{49}$$

Similarly, We consider $\Sigma = \Sigma_r$ the set of matrices of rank at most r in the Hilbert space $\mathcal{H}$ of $n \times n$ symmetric matrices (we study the symmetric case for simplicity, but conjecture that our result can be extended to the non-symmetric case) with $\|\cdot\|_{\mathcal{H}} = \|\cdot\|_F$ (the Frobenius norm). We have again $\Sigma - \Sigma = \Sigma_{2r}$ and all conditions are satisfied such that Corollary 3 can be invoked. Denoting $\Delta = \texttt{eig}(z)$ the vector of eigenvalues of matrix $z \in \mathcal{H}$ sorted by decreasing absolute value, so that $z = U^T \Delta U$ for some unitary matrix U, and defining $z_T := z = U^T \Delta_T U$ for every index set T, we have $P_{\Sigma-\Sigma}(z) = z_{T_2}$ and $z - P_{\Sigma-\Sigma}(z) = z_{T_2^c}$ where $T_2 = T_2(z) \subseteq [1, n]$ is any index set containing $2k$ largest eigenvalues (in magnitude) of z, *i.e.*, such that $\min_{i \in T_2} |\Delta_i| \geq \max_{j \in T_2^c} |\Delta_j|$. With these observations and notations, we are again left to optimize (49).

Specializing to the class $\mathcal{C}$ of convex, coercive, continuous regularizers, we obtain the following theorems.

Theorem 3 *With k-sparse vectors, $\Sigma = \Sigma_k \subseteq \mathcal{H} = \mathbb{R}^n$, $k < \frac{n}{2}$, and $R^\star(\cdot) = \|\cdot\|_1$, we have*

$$\delta_\Sigma^{\texttt{nec}}(R^\star) = \sup_{R \in \mathcal{C}} \delta_\Sigma^{\texttt{nec}}(R) = (2B_{k,n}^\star + 1)^{-1} \qquad \text{with} \qquad B_{k,n}^\star := \max_{1 \leq L \leq n-2k} \frac{L/k}{(L/k+1)^2 + 1}. \tag{50}$$

Moreover, for $k = 1$, the unique functions $R \in \mathcal{C}_\Sigma$ maximizing $\delta_\Sigma^{\texttt{nec}}$ are of the form $R(\cdot) = \lambda\|\cdot\|_1$ with $\lambda > 0$.

Theorem 4 *With the set of rank-r matrices, $\Sigma = \Sigma_r$, in the space $\mathcal{H}$ of symmetric $n \times n$ matrices, $r < \frac{n}{2}$, and with $R^\star(\cdot) = \|\cdot\|_*$ (the nuclear norm), we have*

$$\delta_\Sigma^{\texttt{nec}}(R^\star) = \sup_{R \in \mathcal{C}} \delta_\Sigma^{\texttt{nec}}(R) = (2B_{r,n}^\star + 1)^{-1} \qquad \text{with} \qquad B_{r,n}^\star := \max_{1 \leq L \leq n-2r} \frac{L/r}{(L/r+1)^2 + 1}. \tag{51}$$

The proof of the two theorems exploits technical lemmas detailed in Appendix A.4.1 and Appendix A.4.2.

Proof We give the proof for the case of sparse recovery. To express it for low-rank recovery simply replace the notation k by r. For $1 \leq s \leq n$ and any regularizer R we define

$$B_\Sigma^s(R) := \sup_{z \in \mathcal{T}_R(\Sigma) \setminus \{0\}, z \in \Sigma_s} \frac{\|z_{T_2^c}\|_2^2}{\|z_{T_2}\|_2^2}. \tag{52}$$

For $s \leq 2k$ and any $z \in \Sigma_s$ we have $z_{T_2^c} = 0$ hence $B_\Sigma^s(R) = 0$, thus $B_\Sigma(R) = \max_{1 \leq L \leq n-2k} B_\Sigma^{2k+L}(R)$. First consider $R \in \mathcal{C}_\Sigma$. Since R is positively homogeneous and subadditive, by Lemma 15 for Σ_k / Lemma 19 for Σ_r,

$$B_\Sigma^{2k+L}(R) \geq \frac{\frac{L}{k}}{\left(\frac{L}{k}+1\right)^2 + 1}, \qquad \text{for each } 1 \leq L \leq n - 2k.$$

For $R^\star$ and $1 \leq L \leq n - 2k$ we also have (Lemma 17 / Lemma 20, inspired by [17]) that

$$B_\Sigma(R^\star) = \max_{1 \leq L \leq n-2k} \frac{\frac{L}{k}}{\left(\frac{L}{k}+1\right)^2 + 1}.$$

As a result,

$$B_\Sigma(R) \geq B_\Sigma(R^\star) = \max_{1 \leq L \leq n-2k} \frac{\frac{L}{k}}{\left(\frac{L}{k}+1\right)^2 + 1} =: B_{k,n}^\star$$

Finally, remark that $B_\Sigma(R)$ is an increasing function of $\mathcal{T}_R(\Sigma)$. Using Lemma 4, for any $R \in \mathcal{C}$ there is $R' \in \mathcal{C}_\Sigma$ such that

$$B_\Sigma(R) \geq B_\Sigma(R') \geq B^\star_{k,n}.$$

For $k = 1$, unicity comes from the fact that on a given orthant for $R \in \mathcal{C}_\Sigma$, R is a weighted ℓ^1 norm: $R((x_1, \dots, x_n)) = \sum_i w_i |x_i|$ and the equality case in Lemma 15 forces $w_i = \max_i w_i$. □

Because of the uniqueness result for $k = 1$, the ℓ^1-norm is the unique convex regularizer in $\cap \mathcal{C}_{\Sigma_k}$ that jointly maximizes $\delta^{\mathtt{nec}}_{\Sigma_k}$ for all $k < \frac{n}{2}$ (the proof of Theorem 3 is valid for $\mathcal{C}_{\Sigma_{k'}}$, with $k \leq k' < \frac{n}{2}$). It is an open question to determine if we have unicity model by model. As the result might change for tighter compliance measures, we leave this question for future work.

In the next section, we will see that the optimization of the sufficient RIP constant leads to very similar expressions.

3.3 Compliance measures based on sufficient RC conditions

When Σ is a union of subspaces and R is an arbitrary regularizer, an "explicit" RIP constant $\delta^{\mathtt{suff}}_\Sigma(R)$ is sufficient to guarantee reconstruction [36]. The expression of this constant [36][Eq. (5)] is recalled in the Appendix (Equation (111)) and can be used as a compliance measure. It gives rise to the following lemma, which is proved in Appendix A.5.

Lemma 10 *Assume that $\Sigma = \cup_{V \in \mathcal{V}} V$ is a union of subspaces and that $\Sigma \cap S(1)$ is compact. Consider R any regularizer such that $\Sigma \subseteq \mathtt{dom}(R)$. We have*

$$\delta^{\mathtt{suff}}_\Sigma(R) \geq \frac{1}{\sqrt{\sup\limits_{z \in \mathcal{T}_R(\Sigma) \setminus \{0\}} \frac{\|z - P_\Sigma(z)\|^2_\Sigma}{\|P_\Sigma(z)\|^2_2} + 1}} =: \delta^{\mathtt{suff2}}_\Sigma(R). \tag{53}$$

Further assume that $P_\Sigma(z) \subseteq \arg\min_{x \in \Sigma} \|x - z\|_\Sigma$ for every $z \in \mathcal{H}$ and that, for every $V \in \mathcal{V}$ and every $u \in \Sigma$, $P_{V^\perp}(u) \in \Sigma$. Then, there is equality in (53).

Proof See Appendix A.5. Note that the assumption $P_\Sigma(z) \subseteq \arg\min_{x \in \Sigma} \|x - z\|_\Sigma$ could be replaced by the slightly weaker $P_\Sigma(z) \cap \arg\min_{x \in \Sigma} \|x - z\|_\Sigma / \|x\|_2 \neq \emptyset$. □

We get an immediate corollary of the first claim in the above lemma.

Corollary 4 *Assume that $\Sigma = \cup_{V \in \mathcal{V}} V$ is a union of subspaces and that $\Sigma \cap S(1)$ is compact. For any class $\mathcal{C}'$ of regularizers such that $\Sigma \subseteq \mathtt{dom}(R)$ for every $R \in \mathcal{C}'$, the set of maximizers of $\delta^{\mathtt{suff2}}_\Sigma(\cdot)$ satisfies (whether or not this set of maximizers is empty)*

$$\arg\max_{R \in \mathcal{C}'} \delta^{\mathtt{suff2}}_\Sigma(R) = \arg\min_{R \in \mathcal{C}'} D_\Sigma(R) \qquad \textit{with} \qquad D_\Sigma(R) := \sup_{z \in \mathcal{T}_R(\Sigma) \setminus \{0\}} \frac{\|z - P_{\Sigma - \Sigma}(z)\|^2_\Sigma}{\|P_{\Sigma - \Sigma}(z)\|^2_{\mathcal{H}}}. \tag{54}$$

For any optimal regularizer $R^\star$ we have

$$\delta^{\mathtt{suff2}}_\Sigma(R^\star) = (1 + D_\Sigma(R^\star)^{-1/2}. \tag{55}$$

Note the subtle difference in the norm at the numerator in $B_\Sigma(R)$ and $D_\Sigma(R)$.

Optimal regularization for sparse recovery and low-rank recovery

When considering sparse recovery or low-rank recovery, there is indeed equality $\delta_\Sigma^{\texttt{suff}}(R) = \delta_\Sigma^{\texttt{suff2}}(R)$ thanks to the following Lemma.

Lemma 11 *The assumptions for the equality case of Lemma 10 hold for $\Sigma = \Sigma_k$ the set of k-sparse vectors in $\mathcal{H} = \mathbb{R}^n$, as well as for the set $\Sigma = \Sigma_r$ of symmetric matrices of rank at most r in $\mathcal{H}$ the set of symmetric $n \times n$ matrices.*

Proof See Appendix A.5.

Consider $\Sigma := \Sigma_k$, and regularizers in $\mathcal{C}_\Sigma$. Similarly to the necessary case, from Lemma 10, we have (when Σ is a union of subspace and $\Sigma \cap S(1)$ is closed)

$$D_\Sigma(R) = \sup_{z \in \mathcal{T}_R(\Sigma)\setminus\{0\}} \frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_2^2} \tag{56}$$

where T denotes the support of k largest coordinates of z.

We obtain similar results as in the necessary RIP constant case.

Theorem 5 *With k-sparse vectors, $\Sigma = \Sigma_k \subseteq \mathcal{H} = \mathbb{R}^n$, $k < \frac{n}{2}$, and $R^\star(\cdot) = \|\cdot\|_1$, we have*

$$\delta_\Sigma^{\texttt{suff}}(R^\star) = \sup_{R \in \mathcal{C}} \delta_\Sigma^{\texttt{suff}}(R) = \frac{1}{\sqrt{2}}. \tag{57}$$

Moreover, for $k = 1$, the unique functions $R \in \mathcal{C}_\Sigma$ maximizing $\delta_\Sigma^{\texttt{suff}}$ are of the form $R(\cdot) = \lambda\|\cdot\|_1$ with $\lambda > 0$.

Theorem 6 *With the set of rank-r matrices, $\Sigma = \Sigma_r$, in the space $\mathcal{H}$ of symmetric $n \times n$ matrices, $r < \frac{n}{2}$, and with $R^\star(\cdot) = \|\cdot\|_*$ (the nuclear norm), we have*

$$\delta_\Sigma^{\texttt{suff}}(R^\star) = \sup_{R \in \mathcal{C}} \delta_\Sigma^{\texttt{suff}}(R) = \frac{1}{\sqrt{2}}. \tag{58}$$

Proof We give the proof for the case of sparse recovery. To express it for low-rank recovery simply replace the notation k by r. For $1 \leq s \leq n$ and any regularizer R we define

$$D_\Sigma^s(R) := \sup_{z \in \mathcal{T}_R(\Sigma)\setminus\{0\}, z \in \Sigma_s} \frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_2^2}.$$

For $s \leq k$ and any $z \in \Sigma_s$ we have $z_{T^c} = 0$ hence $D_\Sigma^s(R) = 0$, thus $D_\Sigma(R) = \max_{1 \leq L \leq n-k} D_\Sigma^{k+L}(R)$.

First consider $R \in \mathcal{C}_\Sigma$. Since R is positively homogeneous and subadditive, by Lemma 24 for Σ_k / Lemma 26 for Σ_r,

$$D_\Sigma^{k+L}(R) \geq \min(1, \frac{L}{k}), \qquad \text{for each } 1 \leq L \leq n-k.$$

For $R^\star$ and $1 \leq L \leq n-k$ we also have (with Lemma 23 / Lemma 25) that

$$D_\Sigma^{k+L}(R^\star) = \min(1, \frac{L}{k}).$$

As a result,

$$D_\Sigma(R) \geq D_\Sigma(R^\star) = \max_{1 \leq L \leq n-k} \min(1, \frac{L}{k}) = 1.$$

Finally, remark that $D_\Sigma(R)$ is an increasing function of $\mathcal{T}_R(\Sigma)$. Using Lemma 4, for any $R \in \mathcal{C}$ there is $R' \in \mathcal{C}_\Sigma$ such that

$$D_\Sigma(R) \geq D_\Sigma(R') \geq 1.$$

□

3.4 Discussion

Even without an analytic expression of the sharp RIP constant, it would have been possible to show that $R^\star$ optimizes $\delta_\Sigma^{\mathtt{sharp}}$ if it were simultaneously optimizing its lower and upper bound, *i.e.,* if we had

$$\sup_{R\in\mathcal{C}} \delta_\Sigma^{\mathtt{suff}}(R) = \delta_\Sigma^{\mathtt{suff}}(R^\star) = \delta_\Sigma^{\mathtt{nec}}(R^\star) = \sup_{R\in\mathcal{C}} \delta_\Sigma^{\mathtt{nec}}(R). \tag{59}$$

Unfortunately this is not the case in the sparse and low rank examples. We observe that for a fixed k, n we have in both cases

$$\frac{1}{\sqrt{2}} = \delta_\Sigma^{\mathtt{suff}}(R^\star) < \delta_\Sigma^{\mathtt{nec}}(R^\star). \tag{60}$$

Because of this fact, we cannot conclude on the optimality of $R^\star$ for $\delta_\Sigma^{\mathtt{sharp}}$. However, indexing all objects of the problem by n the dimension of $\mathcal{H}$ (respectively the dimension of the diagonals): the set of regularizers $\mathcal{C}^{(n)}$, the models $\Sigma_k^{(n)}$ and the corresponding $R^{\star,(n)}$ (independent of k for $k < n/2$ as we saw previously). We have (see Remark 3)

$$\inf_{n\geq 3} \inf_{k\in\{1,\dots,\lfloor n/2\rfloor\}} \sup_{R\in\mathcal{C}^{(n)}} \delta_{\Sigma_k^{(n)}}^{\mathtt{nec}}(R) = \frac{1}{\sqrt{2}} = \delta_{\Sigma_k^{(n)}}^{\mathtt{suff}}(R^{\star,(n)}). \tag{61}$$

We deduce

$$\inf_{n\geq 3} \inf_{k\in\{1,\dots,\lfloor n/2\rfloor\}} \sup_{R\in\mathcal{C}^{(n)}} \delta_{\Sigma_k^{(n)}}^{\mathtt{sharp}}(R) = \frac{1}{\sqrt{2}}. \tag{62}$$

and

$$\inf_{n\geq 3} \inf_{k\in\{1,\dots,\lfloor n/2\rfloor\}} \left| \delta_{\Sigma_k^{(n)}}^{\mathtt{sharp}}(R^{\star,(n)}) - \left[\sup_{R\in\mathcal{C}^{(n)}} \delta_{\Sigma_k^{(n)}}^{\mathtt{sharp}}(R) \right] \right| = 0. \tag{63}$$

This shows that the functions $R^{\star,(n)}$ are optimal as a family with respect to a family of models $\Sigma_k^{(n)}$ and the worst case of their associated compliance measures $\delta_{\Sigma_k^{(n)}}^{\mathtt{sharp}}(R)$.

4 Towards the construction of optimal convex regularizers? The example of sparsity in levels.

In the previous Section, optimality was shown by exhibiting the optimal regularizer (ℓ^1-norm and nuclear norm). The only constructive part in these results is given by Theorem 1 that shows that we can look for optimal regularizers in the set of atomic norms $\mathcal{C}_\Sigma$ constructed using the model set Σ. Ideally, given a compliance measure, we would like to be able to construct for any model Σ, an optimal regularizer $R^\star \in \mathcal{C}_\Sigma$. As such an objective seems out of reach with the tools we have developed so far, we study on an example (the case of *sparsity in levels*) the simpler problem of looking for the optimal regularizer in a smaller set of regularizers. We consider a set of weighted ℓ^1-norms and explore the explicit construction of an optimal regularizer for the compliance measure $\delta_\Sigma^{\mathtt{nec}}$.

Sparsity *in levels* is a possible extension of plain sparsity, which is useful for the precise modeling of signals such as medical images [1,6]. It is also useful for simultaneous modeling of sparse signal and sparse noise [33,37].

Definition 7 (Sparsity in levels) In $\mathcal{H} = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \ldots \times \mathbb{R}^{n_L}$, given sparsity levels $k_1, \ldots, k_L$, we define the "sparsity in levels" model with

$$\Sigma_{k_1,\ldots,k_L} := \{x = (x_1, \ldots, x_L) : x_i \in \Sigma_{k_i}\} \tag{64}$$

where Σ_{k_i} is the k_i-sparse model in $\mathbb{R}^{n_i}$.

While the ℓ^1-norm was shown to be is a candidate to perform regularization for models that are sparse in levels [1]. It was also shown that it is possible to obtain better sufficient RIP recovery guarantees when weighting the ℓ^1 norm by $\sqrt{k_i}$ in each level [36]. The following theorem permits to show that this weighting scheme is close to optimal for the compliance measure $\delta_\Sigma^{\mathtt{nec}}$ by giving explicit expressions for the calculation of optimal weights.

Given weights $w = (w_1, \ldots, w_L) \in \mathbb{R}_+^L$, we define the ℓ^1-norm weighted by levels.

$$\|(x_1, \ldots, x_L)\|_w = \sum_{i=1}^{L} w_i \|x_i\|_1. \tag{65}$$

We have the following result.

Theorem 7 *Consider two integers $k_1, k_2 \geq 2$ and the model set $\Sigma = \Sigma_{k_1,k_2}$ in $\mathcal{H} = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ where we assume that $n_1 \geq 4k_1$, $n_2 \geq 4k_2$. Consider $\tilde{a} = 2\sqrt{3} - 3$ and (ν_1^*, ν_2^*) minimizing*

$$\min_{\substack{\nu_1 \in [\tilde{a}, 1-\tilde{a}] \\ \nu_2 = 1-\nu_1}} \max_{i \in \{1,2\}} \max_{x_i \in \{\lfloor k_i \sqrt{1+1/\nu_i} \rfloor; \lceil k_i \sqrt{1+1/\nu_i} \rceil\}} \frac{x_i/k_i}{\nu_i (x_i/k_i + 1)^2 + 1} \tag{66}$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the lower and upper integer part. Then

$$(w_1^*, w_2^*) \in \arg\max_{w_1, w_2 > 0} \delta_\Sigma^{\mathtt{nec}}(\|\cdot\|_{(w_1,w_2)}) \tag{67}$$

if and only if

$$\frac{w_2^*}{w_1^*} = \sqrt{\frac{k_1}{k_2}(1/\nu_1^* - 1)}. \tag{68}$$

Proof See Appendix A.6.

This theorem comes from the fact that (see proof) the quantity defined in (49) satisfies

$$B_{\Sigma_{k_1,k_2}}(\|\cdot\|_{(w_1,w_2)}) = \max_{L_1,L_2} B_{\Sigma_{k_1,k_2}}^{L_1,L_2}((w_1, w_2)) \tag{69}$$

where $B_{\Sigma_{k_1,k_2}}^{L_1,L_2}(\|\cdot\|_{(w_1,w_2)})$ can be computed explicitly (similarly to B_Σ^{2k+L} from (52) for sparsity).

Thanks to the expression from Theorem 7, it becomes tractable to evaluate numerically optimal weights. We simply perform the minimization over $\nu_1 \in [\tilde{a}, 1-\tilde{a}]$ over a regular grid (of 10^6 points in our experiment) and take the minimum. The value of w_1^*/w_2^* is returned according to (68). Let $w_0 = (1/\sqrt{k_1}, 1/\sqrt{k_2})$. In Figure 3, we show a representation of the two criteria $C_1(k_1, k_2) = |1 - \frac{\langle w^*, w_0 \rangle}{\|w^*\|_2 \|w_0\|_2}|$ and $C_2(k_1, k_2) = |\delta_\Sigma^{\mathtt{nec}}(\|\cdot\|_{w^*}) - \delta_\Sigma^{\mathtt{nec}}(\|\cdot\|_{w_0})|$ for different pairs (k_1, k_2). The case $C_1(k_1, k_2) = C_2(k_1, k_2) = 0$ occurs if and only if w_0 is optimal).

We observe numerically that for $2 \leq k_1, k_2 \leq 200$, $C_1(k_1, k_2) \leq 10^{-5}$ and $C_2(k_1, k_2) \leq 5 \cdot 10^{-3}$ and that the error has a tendency to decrease for greater k_1, k_2. This comes from the fact that the result of the discrete optimization over the integers L_i in B_Σ gets closer to the result of a

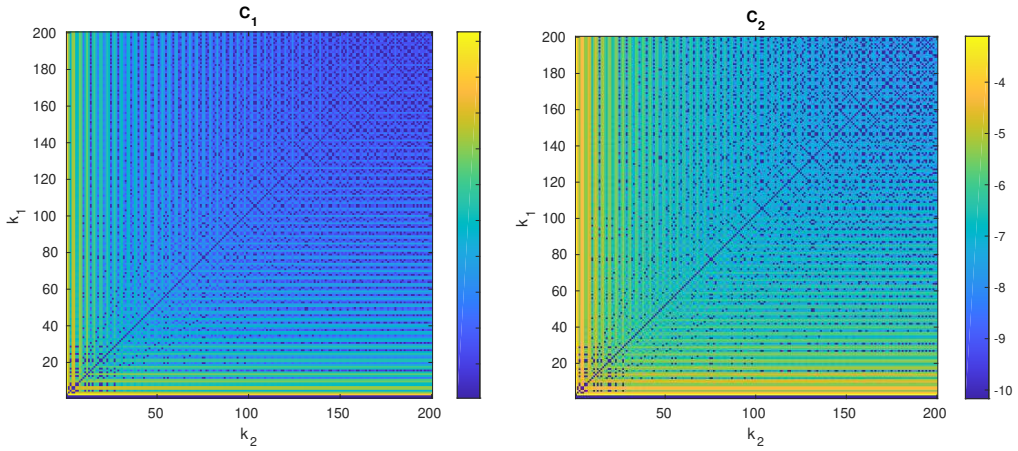

Fig. 3 Then quantities $\log_{10}(C_1(k_1,k_2)) := \log_{10}\left(|1-\frac{\langle w^*,w_0\rangle}{\|w^*\|_2\|w_0\|_2}|\right)$ (left) and $\log_{10}(C_2(k_1,k_2)) := \log_{10}(|\delta_\Sigma^{\texttt{nec}}(\|\cdot\|_{w^*}) - \delta_\Sigma^{\texttt{nec}}(\|\cdot\|_{w_0})|)$ (right) where $w^* = (w_1^*, w_2^*)$ is obtained from Theorem 7 and $w_0 = (1/\sqrt{k_1}, 1/\sqrt{k_2})$ for different $k_1, k_2 \geq 2$.

continuous optimization that yields $w_2^*/w_1^* = \sqrt{k_1}/\sqrt{k_2}$ (obtained by dropping the integer parts in Theorem 7).

This study confirms that the weighting scheme $\left(\frac{1}{\sqrt{k_1}}, \frac{1}{\sqrt{k_2}}\right)$ is close to the optimal choice when the sparsities are known (up to a multiplicative constant). This also shows that even with a simple model and parametrized family of functions, optimization might lead to complicated expressions. We also remark that we could perform the optimization because restricting to weighted atomic norms leads to an analytical description of atoms generating the regularizers. This in turn leads to an analytical description of descent cones. The approach seems difficult to extend to generic atomic norms.

5 Discussion and future work

We gave a general way of defining compliance measures between a regularizer R and a low dimensional model set Σ, and described some elementary properties of such measures. This opens questions on conditions on compliance measures that guarantee the existence of an optimal convex regularizer. Also, the question of manipulating compliance measures for transformations and combinations of models (intersections, unions, sums, ...) is a wide and challenging potential area of research.

We considered noiseless observations instead of the classical noisy model $y = Mx_0 + e$ where e is a measurement noise with finite energy $\|e\|_2$ because of the following remark. Suppose we define an optimal regularizer for a given noise level $\|e\|_2$. There are two possible cases, either the regularizer is also optimal for $\|e\|_2 = 0$ or it is not. In the second case, it means that we would need to trade exact recovery capability for improved stability to noise. This is a possible route to follow especially if there is some latitude on the design of the measurement operator M, *i.e.,* it is possible to increase measurements to improve stability to noise. The analysis of such trade-offs is out of the scope of this article and left for future work.

We have shown that the ℓ^1-norm is optimal among coercive continuous convex functions for sparse recovery for compliance measures based on necessary and sufficient RIP conditions. This result had to be expected due to the symmetries of the problem. The important point is that we

could explicitly quantify the notion of good regularizer. We obtained the same expected results with the nuclear norm for low-rank matrix recovery.

It must be noted that we did not use constructive proofs (we *exhibited* the candidate maximum of the compliance measure) for the sparsity and low-rank cases. A full constructive proof, *i.e.,* an exact calculation and optimization of the quantities $B_\Sigma(R)$ and $D_\Sigma(R)$ would be intellectually more satisfying as it would not require the prior knowledge of the candidate optimum, which is our ultimate objective. We saw in the case of sparsity in levels that we can construct the regularizer that achieved optimality among a simple parametrized family of convex functions (weighted ℓ^1-norms in levels). It is an open question to obtain more general constructions.

We used compliance measures based on (uniform) RIP recovery guarantees to give results for the sparse recovery case, it would be interesting to do such analysis using (non-uniform) recovery guarantees based on the statistical dimension or on the Gaussian width of the descent cones [14,2]. One would need to precisely lower and upper bound these quantities, similarly to our approach with the RIP, to get satisfying results.

Finally, while these compliance measures are designed to make sense with respect to known results in the area of sparse recovery, one might design other compliance measures tailored for particular needs, in this search for optimal regularizers.

Acknowledgements

This work was partly supported by the ANR, project EFFIREG number ANR-20-CE40-0001, project AllegroAssai number ANR-19-CHIA-0009 and project number GraVa ANR-18-CE40-0005.

References

1. B. Adcock, A. Hansen, B. Roman, and G. Teschke. Generalized sampling: stable reconstructions, inverse problems and compressed sensing over the continuum. *Advances in Imaging and Electron Physics*, 182(1):187–279, 2014.
2. D. Amelunxen, M. Lotz, M. B. McCoy, and J. A. Tropp. Living on the edge: phase transitions in convex programs with random data. *Information and Inference*, 3(3):224–294, 2014.
3. D. Amelunxen, M. Lotz, and J. Walvin. Effective condition number bounds for convex regularization. *Information Theory, IEEE Transactions on*, 66(4):2501–2516, 2020.
4. A. Argyriou, R. Foygel, and N. Srebro. Sparse Prediction with the k-Support Norm. *Advances in Neural Information Processing Systems*, 25:1457–1465, 2012.
5. F. Bach. *Learning with Submodular Functions: A Convex Optimization Perspective*. Foundations and Trends in Machine Learning. Now Publishers, 2013.
6. A. Bastounis and A. C. Hansen. On random and deterministic compressed sensing and the restricted isometry property in levels. In *Sampling Theory and Applications (SampTA), 2015 International Conference on*, pages 297–301, 2015.
7. D. P. Bertsekas. Convexification procedures and decomposition methods for nonconvex optimization problems. *Journal of Optimization Theory and Applications*, 29(2):169–197, 1979.
8. M. Bougeard, J.-P. Penot, and A. Pommellet. Towards minimal assumptions for the infimal convolution regularization. *Journal of approximation theory*, 64(3):245–270, 1991.
9. A. Bourrier, M. Davies, T. Peleg, P. Perez, and R. Gribonval. Fundamental performance limits for ideal decoders in high-dimensional linear inverse problems. *Information Theory, IEEE Transactions on*, 60(12):7928–7946, 2014.
10. E. Candès and B. Recht. Simple bounds for recovering low-complexity models. *Mathematical Programming*, 141(1-2):577–589, 2013.
11. E. J. Candes and Y. Plan. Matrix completion with noise. *Proceedings of the IEEE*, 98(6):925–936, 2010.
12. E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *Information Theory, IEEE Transactions on*, 52(2):489–509, 2006.
13. A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision*, 40(1):120–145, 2011.

14. V. Chandrasekaran, B. Recht, P. Parrilo, and A. Willsky. The convex geometry of linear inverse problems. *Foundations of Computational Mathematics*, 12(6):805–849, 2012.
15. S. Chen, D. Donoho, and M. Saunders. Atomic Decomposition by Basis Pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1998.
16. Y. Chi, Y. M. Lu, and Y. Chen. Nonconvex optimization meets low-rank matrix factorization: An overview. *Signal Processing, IEEE Transactions on*, 67(20):5239–5269, 2019.
17. M. E. Davies and R. Gribonval. Restricted isometry constants where ℓ^p sparse recovery can fail for $0 < p \leq 1$. *Information Theory, IEEE Transactions on*, 55(5):2203–2214, 2009.
18. D. L. Donoho. For most large underdetermined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics*, 59(6):797–829, 2006.
19. J. Fan and R. Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, 96(456):1348–1360, 2001.
20. M. Fazel, H. Hindi, and S. P. Boyd. A rank minimization heuristic with application to minimum order system approximation. In *Proceedings of the 2001 American Control Conference*, volume 6, pages 4734–4739 vol.6, 2001.
21. S. Foucart and M.-J. Lai. Sparsest solutions of underdetermined linear systems via ℓ^q-minimization for $0 < q <= 1$. *Applied and Computational Harmonic Analysis*, 26(3):395–407, 2009.
22. S. Foucart and H. Rauhut. *A mathematical introduction to compressive sensing*. Springer, 2013.
23. S. Friedland and L.-H. Lim. Nuclear norm of higher-order tensors. *Mathematics of Computation*, 87(311):1255–1281, 2018.
24. M. Jach, D. Michaels, and R. Weismantel. The Convex Envelope of (n–1)-Convex Functions. *SIAM Journal on Optimization*, 19(3):1451–1466, 2008.
25. J. B. Lasserre. The moment-sos hierarchy. *Proceedings of the International Congress of Mathematics*, pages 3761–3784, 2018.
26. M. März, C. Boyer, J. Kahn, and P. Weiss. Sampling rates for ℓ^1-synthesis. *arXiv preprint arXiv:2004.07175*, 2020.
27. S. N. Negahban, P. Ravikumar, M. J. Wainwright, and B. Yu. A Unified Framework for High-Dimensional Analysis of M-Estimators with Decomposable Regularizers. *Statistical Science*, 27(4):538–557, 2012.
28. T. Pock, D. Cremers, H. Bischof, and A. Chambolle. Global solutions of variational models with convex regularization. *SIAM Journal on Imaging Sciences*, 3(4):1122–1145, 2010.
29. G. Puy, M. E. Davies, and R. Gribonval. Recipes for stable linear embeddings from hilbert spaces to $\mathbb{R}^m$. *Information Theory, IEEE Transactions on*, 63(4):2171–2187, 2017.
30. B. Recht, M. Fazel, and P. Parrilo. Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization. *SIAM Review*, 52(3):471–501, 2010.
31. R. T. Rockafellar. *Convex analysis*. Number 28. Princeton university press, 1970.
32. E. Soubies, L. Blanc-Féraud, and G. Aubert. A continuous exact ℓ_0 penalty (cel0) for least squares regularized problem. *SIAM Journal on Imaging Sciences*, 8(3):1607–1639, 2015.
33. C. Studer and R. G. Baraniuk. Stable restoration and separation of approximately sparse signals. *Applied and Computational Harmonic Analysis*, 37(1):12–35, 2014.
34. Y. Traonmilin and J.-F. Aujol. The basins of attraction of the global minimizers of the non-convex sparse spike estimation problem. *Inverse Problems*, 36(4):045003, 2020.
35. Y. Traonmilin, J.-F. Aujol, and A. Leclaire. The basins of attraction of the global minimizers of non-convex inverse problems with low-dimensional models in infinite dimension. *arXiv preprint arXiv:2009.08670*, 2020.
36. Y. Traonmilin and R. Gribonval. Stable recovery of low-dimensional cones in hilbert spaces: One rip to rule them all. *Applied and Computational Harmonic Analysis*, 45(1):170–205, 2018.
37. Y. Traonmilin, S. Ladjal, and A. Almansa. Robust multi-image processing with optimal sparse regularization. *Journal of Mathematical Imaging and Vision*, 51(3):413–429, 2015.
38. Y. Traonmilin and S. Vaiter. Optimality of 1-norm regularization among weighted 1-norms for sparse recovery: a case study on how to find optimal regularizations. *Journal of Physics: Conference Series*, 1131:012009, 2018.
39. S. Vaiter, M. Golbabaee, J. Fadili, and G. Peyré. Model selection with low complexity priors. *Information and Inference: A Journal of the IMA*, 4(3):230–287, 2015.
40. S. Vaiter, G. Peyre, and J. Fadili. Model Consistency of Partly Smooth Regularizers. *Information Theory, IEEE Transactions on*, PP(99):1–1, 2017.
41. R. Vershynin. Estimation in high dimensions: a geometric perspective. In *Sampling theory, a renaissance*, pages 3–66. Springer, 2015.
42. M. Yuan and Y. Lin. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(1):49–67, 2006.

A Appendices

This section describes the tools and proofs used to obtain our results.

A.1 Summary of properties used in proofs

From [36, Table 1] (which summarizes results from [31]), the function $x \in \mathcal{E}(\mathcal{A}) \mapsto \|x\|_{\mathcal{A}}$ is always non-negative, lower semi-continuous and subadditive (*i.e.,* it satisfies the triangle inequality). It is furthermore positively homogeneous as soon as $0 \in \overline{\text{conv}}(\mathcal{A})$, continuous as soon as 0 is in the interior of $\overline{\text{conv}}(\mathcal{A})$, and coercive as soon as $\overline{\text{conv}}(\mathcal{A})$ is bounded. Finally, it is indeed a norm if $\overline{\text{conv}}(\mathcal{A}) = -\overline{\text{conv}}(\mathcal{A})$.

We refer the reader to [36][Section 2.2] and [4] for properties of the atomic norm $\|\cdot\|_{\Sigma}$ (cf Definition 3). We will use the following two properties of $\|\cdot\|_{\Sigma}$ (defined in Section 2.5).

Fact A1 (From e.g. [36]) *For all $x \in \Sigma$, $\|x\|_{\Sigma} = \|x\|_{\mathcal{H}}$.*

Fact A2 (From [36][Fact 2.1] applied to $\|\cdot\|_{\Sigma}$) *For all $z \in \mathcal{H}$*

$$\|z\|_{\Sigma} = \inf\left\{\sqrt{\sum \lambda_i \|u_i\|_{\mathcal{H}}^2} : \lambda_i \in \mathbb{R}_+, \sum \lambda_i = 1, u_i \in \Sigma, z = \sum \lambda_i u_i\right\}. \tag{70}$$

A.2 Proofs for Section 2

A.2.1 Proof of Lemma 1

Consider $x \in \Sigma$, and $z \in \mathcal{H}$. We have $\iota_{\Sigma}(x+z) \leq \iota_{\Sigma}(x) = 0$ if and only if $x + z \in \Sigma$, *i.e.,* if there is $x' \in \Sigma$ such that $z = x' - x$. Hence, $\mathcal{T}_{\iota_{\Sigma}}(x) = \{\gamma(x'-x) : \gamma \in \mathbb{R}, x' \in \Sigma\}$. It follows that $\mathcal{T}_{\iota_{\Sigma}}(\Sigma) = \{\gamma z : \gamma \in \mathbb{R}, z \in \Sigma - \Sigma\} \supseteq \Sigma - \Sigma$. When Σ is positively homogeneous, for any $z = x' - x \in \Sigma - \Sigma$ and $\gamma \in \mathbb{R}$ we have: if $\gamma > 0$ then $\gamma z = \gamma x' - \gamma x \in \Sigma - \Sigma$; if $\gamma < 0$ then $\gamma z = (-\gamma)x - (-\gamma)x' \in \Sigma - \Sigma$; if $\gamma = 0$ then $\gamma z = 0 = x - x \in \Sigma - \Sigma$, hence indeed $\mathcal{T}_{\iota_{\Sigma}}(\Sigma) \subseteq \Sigma - \Sigma$.

Now consider $y \in \mathcal{T}_{\iota_{\Sigma}}(\Sigma)$ and write it as $y = \gamma(x_1 - x_2)$ where $x_1, x_2 \in \Sigma$ and $\gamma \in \mathbb{R}$. Since $\Sigma \subseteq \texttt{dom}(R)$ we have $\max(R(x_1), R(x_2)) < \infty$. We will prove that $y \in \mathcal{T}_R(\Sigma)$. We distinguish two cases: if $R(x_1) \leq R(x_2)$ then $R(x_2 + (x_1 - x_2)) = R(x_1) \leq R(x_2)$ hence $y = \gamma(x_1 - x_2) \in \mathcal{T}_R(x_2)$, and as $x_2 \in \Sigma$ it follows that $y \in \mathcal{T}_R(\Sigma)$; otherwise $R(x_2) < R(x_1)$ hence $R(x_1 + (x_2 - x_1)) = R(x_2) < R(x_1)$ hence $y = (-\gamma)(x_2 - x_1) \in \mathcal{T}_R(x_1)$ and therefore $y \in \mathcal{T}_R(\Sigma)$. □

A.2.2 Proof of Lemma 4

Given $t > R(0)$, the level set $\mathcal{L}(R,t) = \{y \in \mathcal{H} : R(y) \leq t\}$ is non empty, convex and closed (by convexity and lower semi-continuity of R), and bounded (by coercivity of R). We define $\mathcal{A} := \mathcal{L}(R,t) \cap \Sigma = \{x \in \Sigma : R(x) \leq t\}$.

Consider $z \in \mathcal{T}_{\|\cdot\|_{\mathcal{A}}}(\Sigma)$. If $z = 0$ then clearly $z \in \mathcal{T}_R(\Sigma)$. Let us prove that the same holds when $z \neq 0$. By definition, there exists $\gamma \in \mathbb{R} \setminus \{0\}$ and $x \in \Sigma$ such that

$$\|x + z/\gamma\|_{\mathcal{A}} \leq \|x\|_{\mathcal{A}}.$$

On the one hand we have $R(0 \cdot x) = R(0) < t$. On the other hand, since R is coercive, we have $R(\lambda x) \underset{\lambda \to +\infty}{\to} +\infty$. Since R is continuous, by the mean value theorem, there is $\lambda_0 > 0$ such that

$$R(\lambda_0 x) = t.$$

Since Σ is a cone, the vector $x' = \lambda_0 x$ belongs to Σ and, since $R(x') = t$, by definition of $\mathcal{A}$ we have indeed $x' \in \mathcal{A}$, hence $\|x'\|_{\mathcal{A}} \leq 1$. Furthermore, since $\|\cdot\|_{\mathcal{A}}$ is positively homogeneous (because $0 \in \overline{\text{conv}}(\mathcal{A})$), we have

$$\|x' + \lambda_0 z/\gamma\|_{\mathcal{A}} = \lambda_0 \|x + z/\gamma\|_{\mathcal{A}} \leq \lambda_0 \|x\|_{\mathcal{A}} = \|x'\|_{\mathcal{A}}.$$

We now observe that, on the one hand, the level set $\mathcal{L}(\|\cdot\|_{\mathcal{A}}, 1) = \overline{\text{conv}}(\mathcal{A})$ is the smallest closed convex set containing $\mathcal{A}$; on the other hand $\mathcal{A} \subset \mathcal{L}(R,t)$ and $\mathcal{L}(R,t)$ is convex and closed. Thus $\mathcal{L}(\|\cdot\|_{\mathcal{A}}, 1) \subset \mathcal{L}(R,t)$ and the fact that $\|x' + \lambda_0 z/\gamma\|_{\mathcal{A}} \leq \|x'\|_{\mathcal{A}} \leq 1$ therefore implies

$$R(x' + \lambda_0 z/\gamma) \leq t = R(x'). \tag{71}$$

This shows that $z \in \mathcal{T}_R(\Sigma)$ and establishes that $\mathcal{T}_{\|\cdot\|_{\mathcal{A}}}(\Sigma) \subseteq \mathcal{T}_R(\Sigma)$.

Let us now prove that $\|\cdot\|_{\mathcal{A}}$ is continuous, convex, coercive and positively homogeneous. First, from the property of gauges (see Appendix A.1), $\|\cdot\|_{\mathcal{A}}$ is always convex and lower semi-continuous. Second, since R is coercive, its level sets are bounded, hence $\overline{\text{conv}}(\mathcal{A})$ is bounded and $\|\cdot\|_{\mathcal{A}}$ is coercive. Finally, as $R(0) < t$ and R is

continuous, 0 is in the interior of $\mathcal{L}(R,t)$. There exists $\epsilon > 0$ such that an open ball O of radius ϵ centered on 0 is included in $\mathcal{L}(R,t)$. This implies $O \cap \Sigma \subset \mathcal{L}(R,t) \cap \Sigma = \mathcal{A}$ which in turns imply $\text{conv}(O \cap \Sigma) \subset \text{conv}(\mathcal{A}) \subset \overline{\text{conv}(\mathcal{A})}$. Remark that $\mathcal{E}(O \cap \Sigma) = \mathcal{E}(\Sigma) = \mathcal{H}$. Now we need to find O' an open ball of radius ϵ' such that $O' \subset \text{conv}(O \cap \Sigma)$. In each orthant Ω_r, we can find a normalised basis $E = (e_i) \in \Sigma$ such that $\Omega_r \subset \mathcal{E}(E)$. We define the norm $\| \sum_i \mu_i e_i \|_E = \sum \mu_i$. This norm is equivalent to $\| \cdot \|_{\mathcal{H}}$. This implies there is a constant c_r depending on the orthant Ω_r, such that for $x = \sum_i \mu_i e_i \in O' \cap \Omega_r$, $\max_i \mu_i < c_r \epsilon'$. This implies

$$x = t \sum_i \frac{\mu_i}{\sum_j \mu_j} \epsilon e_i \tag{72}$$

with $t = \frac{\sum_j \mu_j}{\epsilon} \leq n c_r \frac{\epsilon'}{\epsilon}$. Taking $\epsilon' < \epsilon/(n c_r)$ implies $t < 1$ and $x \in \text{conv}(O \cap \Sigma)$. As there is a finite number of orthants we can chose ϵ' such that we always have $x \in O'$ implies $x \in \text{conv}(O \cap \Sigma)$. □

A.3 Proofs for Section 3.1

Proof (Proof of Lemma 6)
Denote $\alpha = \inf_{x \in (\Sigma - \Sigma) \cap S(1)} \|Mx\|_2^2$ and $\beta = \sup_{x \in (\Sigma - \Sigma) \cap S(1)} \|Mx\|_2^2$, so that $\gamma(M) = \beta/\alpha$. Since Σ is a cone, we have for every $x \in \Sigma - \Sigma$,

$$\alpha \|x\|_{\mathcal{H}}^2 \leq \|Mx\|_2^2 \leq \beta \|x\|_{\mathcal{H}}^2 = \gamma(M) \alpha \|x\|_{\mathcal{H}}^2, \tag{73}$$

Multiplying x in (73) by any $\lambda > 0$, we have

$$\lambda^2 \alpha \|x\|_{\mathcal{H}}^2 \leq \|\lambda M x\|_2^2 \leq \lambda^2 \gamma(M) \alpha \|x\|_{\mathcal{H}}^2.$$

We look for $\lambda > 0$, $\delta \neq 1$ such that λM satisfies a symmetric RIP with constant δ, *i.e.*,

$$\lambda^2 \alpha = 1 - \delta \quad \text{and} \quad \lambda^2 \gamma(M) \alpha = 1 + \delta.$$

Adding these two equalities yields $\lambda^2 \alpha (1 + \gamma(M)) = 1$, hence $\lambda^2 = \frac{1}{\alpha(1+\gamma(M))}$. Dividing them yields

$$\frac{1-\delta}{1+\delta} = \gamma(M) \iff \delta = \frac{\gamma(M) - 1}{\gamma(M) + 1}.$$

We have shown that for any M, there exists $\lambda > 0$ such that

$$\delta(\lambda M) \leq \frac{\gamma(M) - 1}{\gamma(M) + 1}.$$

Remark that the value of λ that makes the RIP bounds symmetrical is unique, and that no better symmetrical RIP bound can be obtained, otherwise we could construct a better restricted conditioning (which is impossible by definition of $\gamma(M)$). We deduce

$$\delta(\lambda M) = \frac{\gamma(M) - 1}{\gamma(M) + 1}.$$

□

Lemma 12 *Consider a cone $\Sigma \subseteq \mathcal{H}$ and $\mathcal{T} \subseteq \mathcal{H}$ a non-empty set, and denote $\mathcal{P}$ the set of symmetric positive semi-definite linear operators on $\mathcal{H}$,* i.e., *$N \in \mathcal{P}$ if and only if $N^H = N$ and $N \succeq 0$. Then*

$$\inf_{M : \ker M \cap \mathcal{T} \neq \{0\}} \gamma_\Sigma(M) = \inf_{N \in \mathcal{P} : \dim \ker N = 1, \ker N \cap \mathcal{T} \neq \{0\}} \gamma_\Sigma(N). \tag{74}$$

Proof The infimum on the r.h.s. of (74) is over a more constrained set than on the l.h.s., hence

$$\inf_{M : \ker M \cap \mathcal{T} \neq \{0\}} \gamma_\Sigma(M) \leq \inf_{N \in \mathcal{P} : \dim \ker N = 1, \ker M \cap \mathcal{T} \neq \{0\}} \gamma_\Sigma(N).$$

If the l.h.s. is infinite, then the right hand side must also be infinite and we are done.

Assume that the l.h.s. is finite. We now prove the reverse inequality. For this, consider M a linear operator with $\ker M \cap \mathcal{T} \neq \{0\}$ and $\gamma_\Sigma(M) < \infty$. There exists a nonzero vector $t \in \ker M \cap \mathcal{T}$. We build an operator $N \in \mathcal{P}$ such that $\ker N = \text{span}(t)$ and with $\gamma_\Sigma(N)$ arbitrarily close to $\gamma_\Sigma(M)$.

Since $\gamma_\Sigma(M) < \infty$, M is nonzero hence a singular value decomposition allows to write $M = \sum_{i=1}^r \sigma_i u_i v_i^H$ where $(u_i)_{i=1}^r$ and $(v_i)_{i=1}^r$ are orthonormal families and $\min_{1\leq i\leq r} \sigma_i > 0$. First we deal with the case where $\dim\ker M = 1$. We set $N = \sum_{i=1}^r \sigma_i v_i v_i^H$ so that $N \in \mathcal{P}$ and $\dim\ker N = 1$ too. Since $\|Nx\|_2^2 = \sum_{i=1}^r \sigma_i^2 \langle v_i, x\rangle^2 = \|Mx\|_2^2$ for any vector x we have $\gamma(N) = \gamma(M)$ and we are done. Assume now that $k := \dim\ker M \geq 2$. Observe that $\mathrm{span}(t) \subset \ker M$ and let $(e_1, \dots, e_{k-1})$ be an orthonormal basis of the orthogonal complement of $\mathrm{span}(t)$ in $\ker M$, so that $(v_1, \dots, v_r, e_1, \dots, e_{k-1})$ is an orthonormal family. For each $\epsilon > 0$, define $N_\epsilon = \sum_{i=1}^r \sigma_i v_i v_i^H + \epsilon \sum_{j=1}^{k-1} e_j e_j^H$. Again, $N_\epsilon \in \mathcal{P}$ and $\mathrm{span}(t) = \ker N_\epsilon$ so that $\dim\ker N_\epsilon = 1$, and for each $x \in \mathcal{H}$ we have

$$\|N_\epsilon x\|_2^2 = \sum_{i=1}^r \sigma_i^2 \langle v_i, x\rangle^2 + \epsilon^2 \sum_{j=1}^{k-1} \langle e_j, x\rangle^2 = \|Mx\|_2^2 + \epsilon^2 \sum_{j=1}^{k-1} \langle e_j, x\rangle^2,$$

hence $\|Mx\|_2^2 \leq \|N_\epsilon x\|_2^2 \leq \|Mx\|_2^2 + \epsilon^2 \|x\|_2^2$. Since $\gamma_\Sigma(M) < \infty$, we get

$$0 < \inf_{x\in(\Sigma-\Sigma)\cap S(1)} \|Mx\|_2^2 \leq \inf_{x\in(\Sigma-\Sigma)\cap S(1)} \|N_\epsilon x\|_2^2 \leq \sup_{x\in(\Sigma-\Sigma)\cap S(1)} \|N_\epsilon x\|_2^2 \leq \sup_{x\in(\Sigma-\Sigma)\cap S(1)} \|Mx\|_2^2 + \epsilon^2$$

which implies

$$\gamma_\Sigma(N_\epsilon) \leq \frac{\sup_{x\in(\Sigma-\Sigma)\cap S(1)} \|Mx\|_2^2 + \epsilon^2}{\inf_{x\in(\Sigma-\Sigma)\cap S(1)} \|Mx\|_2^2} = \gamma_\Sigma(M) + \frac{\epsilon^2}{\inf_{x\in(\Sigma-\Sigma)\cap S(1)} \|Mx\|_2^2}.$$

This implies that $\inf_{\epsilon>0} \gamma_\Sigma(N_\epsilon) \leq \gamma_\Sigma(M)$ as claimed. □

Proof (Proof of Lemma 7) We define

$$G(\Sigma, E, M) := \frac{\sup_{y\in(\Sigma-\Sigma)\cap E} \|My\|_2^2}{\inf_{y\in(\Sigma-\Sigma)\cap E} \|My\|_2^2}. \tag{75}$$

For any nonzero M, we have

$$\gamma_{F\Sigma}(M) = \frac{\sup_{x\in(F\Sigma-F\Sigma)\cap S(1)} \|Mx\|_2^2}{\inf_{x\in(F\Sigma-F\Sigma)\cap S(1)} \|Mx\|_2^2} = \frac{\sup_{y\in(\Sigma-\Sigma)\cap F^{-1}S(1)} \|MFy\|_2^2}{\inf_{y\in(\Sigma-\Sigma)\cap F^{-1}S(1)} \|MFy\|_2^2}. \tag{76}$$

Hence

$$\begin{aligned} A_{F\Sigma}^{RC}(R\circ F^{-1}) &= \inf_{M:\ker M\cap\mathcal{T}_{R\circ F^{-1}}(F\Sigma)\neq\{0\}} \gamma_{F\Sigma}(M) \\ &= \inf_{M:\ker M\cap\mathcal{T}_{R\circ F^{-1}}(F\Sigma)\neq\{0\}} G(\Sigma, F^{-1}S(1), MF). \end{aligned}$$

By Lemma 5 with $R' = R\circ F^{-1}$, $\mathcal{T}_{R\circ F^{-1}}(F\Sigma) = \mathcal{T}_{R'}(F\Sigma) = F(\mathcal{T}_{R'\circ F}(\Sigma)) = F(\mathcal{T}_R(\Sigma))$. Also, $\ker M \cap \mathcal{T}_{R\circ F^{-1}}(F\Sigma) \neq \{0\}$ is equivalent to the existence of $z \in \ker M$ such that $z' := F^{-1}z \in \mathcal{T}_R(\Sigma)$, *i.e.,* of $z' \in \mathcal{T}_R(\Sigma)$ such that $z := Fz' \in \ker M$. As a result,

$$\inf_{M:\ker M\cap\mathcal{T}_{R\circ F^{-1}}(F\Sigma)\neq\{0\}} \gamma_{F\Sigma}(M) = \inf_{M:F^{-1}\ker M\cap\mathcal{T}_R(\Sigma)\neq\{0\}} G(\Sigma, F^{-1}S(1), MF). \tag{77}$$

Rewriting $M' = MF$, we have $\ker M' = F^{-1}\ker M$ and

$$\inf_{M:\ker M\cap\mathcal{T}_{R\circ F^{-1}}(F\Sigma)\neq\{0\}} \gamma_{F\Sigma}(M) = \inf_{M':\ker M'\cap\mathcal{T}_R(\Sigma)\neq\{0\}} G(\Sigma, F^{-1}S(1), M') \tag{78}$$

which gives the desired result using the fact that $F^{-1}S(1) = S(1)$ since F is a linear isometry. □

A.4 Proofs for Section 3.2

Proof (Proof of Lemma 8)
Consider $z \in \mathcal{H}\setminus\{0\}$ and $M = I - \Pi_z$. For every $x \in S(1)$, we have

$$\|Mx\|_2^2 = 1 - \frac{\langle x, z\rangle^2}{\|z\|_{\mathcal{H}}^2} \tag{79}$$

hence

$$\gamma_\Sigma(M) = \frac{\sup_{x\in(\Sigma-\Sigma)\cap S(1)} \|Mx\|_2^2}{\inf_{x\in(\Sigma-\Sigma)\cap S(1)} \|Mx\|_2^2} = \frac{1-\inf_{x\in(\Sigma-\Sigma)\cap S(1)} \frac{\langle x,z\rangle^2}{\|z\|_{\mathcal{H}}^2}}{1-\sup_{x\in(\Sigma-\Sigma)\cap S(1)} \frac{\langle x,z\rangle^2}{\|z\|_{\mathcal{H}}^2}}$$

Case 1: By assumption there is x_0 such that $\|x_0\|_{\mathcal{H}} = 1$ and $\Sigma \subseteq \mathrm{span}(x_0)$. Since $\Sigma \neq \{0\}$ is a cone, it follows that $(\Sigma-\Sigma)\cap S(1) = \mathrm{span}(x_0)\cap S(1) = \{-x_0, +x_0\}$ and

$$\inf_{x\in(\Sigma-\Sigma)\cap S(1)} \frac{\langle x,z\rangle^2}{\|z\|_{\mathcal{H}}^2} = \sup_{x\in(\Sigma-\Sigma)\cap S(1)} \frac{\langle x,z\rangle^2}{\|z\|_{\mathcal{H}}^2} = \frac{\langle x_0,z\rangle^2}{\|z\|_{\mathcal{H}}^2}. \tag{80}$$

Hence if $z \in \Sigma = \mathrm{span}(x_0)$ we have $\gamma_\Sigma(M) = +\infty$, otherwise $\frac{\langle x_0,z\rangle^2}{\|z\|_{\mathcal{H}}^2} < 1$ and $\gamma_\Sigma(M) = 1$. Thus, if $\mathcal{T}_R(\Sigma) \subseteq \Sigma$ we have $A_\Sigma^{RIP,\mathtt{nec}}(R) = +\infty$, otherwise there is $z \in \mathcal{T}_R(\Sigma)\setminus\Sigma$, and $A_\Sigma^{RIP,\mathtt{nec}}(R) = 1$.
Case 2: Let us show that for any $z \neq 0$ there is some $x \in (\Sigma-\Sigma)\setminus\{0\}$ such that $\langle x,z\rangle = 0$. This implies $\inf_{x\in(\Sigma-\Sigma)\cap S(1)} \frac{\langle x,z\rangle^2}{\|z\|_{\mathcal{H}}^2} = 0$ and yields the result. Indeed, by assumption, given any $x_1 \in \Sigma\setminus\{0\}$ there is $x_2 \in \Sigma$ such that $x_2 \notin \mathrm{span}(x_1)$ (hence $x_2 \neq 0$). If $\langle x_1,z\rangle = 0$ we take $x = x_1 = x_1 - \lambda x_2$ with $\lambda = 0$. Otherwise, with $\lambda = \frac{\langle x_2,z\rangle}{\langle x_1,z\rangle}$ we set $x = \lambda x_1 - x_2$. In both cases we have $x \neq 0$ and, since Σ is a cone, $x \in \Sigma-\Sigma$ and $\langle \lambda x_1 - x_2, z\rangle = 0$. □

Proof (Proof of Lemma 9) Since $E\cap S(1)$ is compact, for any z there exists $\tilde{x} \in E\cap S(1)$ such that

$$|\langle \tilde{x},z\rangle|^2 = \max_{\tilde{y}\in E\cap S(1)} |\langle \tilde{y},z\rangle|^2. \tag{81}$$

Since E is a union of subspaces, it is homogeneous. Thus, as $\tilde{x} \in E$, we have $x := \langle \tilde{x},z\rangle\tilde{x} \in E$. If $y \in E\setminus\{0\}$, we have $\tilde{y} := y/\|y\|_{\mathcal{H}} \in E\cap S(1)$, $\langle z,\tilde{y}\rangle\tilde{y}$ is the orthogonal projection of z on $\tilde{y}$ and

$$\begin{aligned}\|z-y\|_{\mathcal{H}}^2 = \left\|z-\|y\|_{\mathcal{H}}\cdot\tilde{y}\right\|_{\mathcal{H}}^2 &\geq \|z-\langle z,\tilde{y}\rangle\tilde{y}\|_{\mathcal{H}}^2 = \|z\|_{\mathcal{H}}^2 - |\langle z,\tilde{y}\rangle|^2 \\ &\overset{(81)}{\geq} \|z\|_{\mathcal{H}}^2 - |\langle z,\tilde{x}\rangle|^2\end{aligned} \tag{82}$$

Since $\|z-x\|_{\mathcal{H}}^2 = \|z\|_{\mathcal{H}}^2 - 2\mathcal{R}e\langle z,x\rangle + \|x\|_{\mathcal{H}}^2 = \|z\|_{\mathcal{H}}^2 - |\langle z,\tilde{x}\rangle|^2$, we conclude

$$\|z-y\|_{\mathcal{H}}^2 \geq \|z-x\|_{\mathcal{H}}^2 \tag{83}$$

and $x \in P_E(z)$ by definition of P_E.

If $x' \in P_E(z)$, we have $\|z-x'\|_{\mathcal{H}}^2 = \|z-x\|_{\mathcal{H}}^2 = \min_{y\in E}\|z-y\|_{\mathcal{H}}^2$ hence the notation $\|z-P_E(z)\|_{\mathcal{H}}^2$ is unambiguous. Since $x' \in P_E(z)$, there is equality in the above equation with $y = x'$, hence $\|y\|_{\mathcal{H}} = \langle z,\tilde{y}\rangle$ and $|\langle z,\tilde{y}\rangle|^2 = |\langle z,\tilde{x}\rangle|^2$, therefore $\langle z,y\rangle = \langle z,\|y\|_{\mathcal{H}}\tilde{y}\rangle = \|y\|_{\mathcal{H}}\langle z,\tilde{y}\rangle = \|y\|_{\mathcal{H}}^2 = \langle z,\tilde{y}\rangle^2 = \langle z,\tilde{x}\rangle^2 = \|x\|_{\mathcal{H}}^2$. This shows that the notations $\|P_E(z)\|_{\mathcal{H}}^2$ and $\langle z,P_E(z)\rangle$ are unambiguous and that $\|P_E(z)\|_{\mathcal{H}}^2 = \langle z,P_E(z)\rangle$.

We also have $\|z\|_{\mathcal{H}}^2 = \|x\|_{\mathcal{H}}^2 + \|z-x\|_{\mathcal{H}}^2 = \|x'\|^2$, and $\langle z,y\rangle = \|y\|_{\mathcal{H}}$ hence the notations $\|z-P_E(z)\|_{\mathcal{H}}^2$ and $\|P_E(z)\|_{\mathcal{H}}^2$ are unambiguous. □

Proof (Proof of Corollary 3) Since $\Sigma-\Sigma$ is a union of subspaces and $(\Sigma-\Sigma)\cap S(1)$ is compact, by Lemma 9, $\sup_{x\in(\Sigma-\Sigma)\cap S(1)} \frac{\langle x,z\rangle^2}{\|z\|_{\mathcal{H}}^2} = \frac{\|P_{\Sigma-\Sigma}(z)\|_{\mathcal{H}}^2}{\|z\|_{\mathcal{H}}^2}$, hence we have

$$\begin{aligned}(B_\Sigma(R)+1)^{-1} &= \left(\sup_{z\in\mathcal{T}_R(\Sigma)\setminus\{0\}} \frac{\|z-P_{\Sigma-\Sigma}(z)\|_{\mathcal{H}}^2}{\|P_{\Sigma-\Sigma}(z)\|_{\mathcal{H}}^2} + 1\right)^{-1} = \inf_{z\in\mathcal{T}_R(\Sigma)\setminus\{0\}} \frac{\|P_{\Sigma-\Sigma}(z)\|_{\mathcal{H}}^2}{\|z-P_{\Sigma-\Sigma}(z)\|_{\mathcal{H}}^2 + \|P_{\Sigma-\Sigma}(z)\|_{\mathcal{H}}^2} \\ &= \inf_{z\in\mathcal{T}_R(\Sigma)\setminus\{0\}} \frac{\|P_{\Sigma-\Sigma}(z)\|_{\mathcal{H}}^2}{\|z\|_{\mathcal{H}}^2}.\end{aligned}$$

Since Σ is a cone and $\Sigma \neq \mathrm{span}(x)$ for each $x \in \Sigma$, by Lemma 8, using (32) we have $\gamma_\Sigma^{\mathtt{nec}}(R) = \frac{1}{1-(1+B_\Sigma(R))^{-1}} = 1 + 1/B_\Sigma(R)$ hence $\delta_\Sigma^{\mathtt{nec}}(R) = \frac{\gamma_\Sigma^{\mathtt{nec}}(R)-1}{\gamma_\Sigma^{\mathtt{nec}}(R)+1} = (2B_\Sigma(R)+1)^{-1}$.

We conclude using that $b \mapsto= 1/(1+2b)$ is decreasing. □

A.4.1 Lemmas for the proof of Theorem 3 (sparse recovery)

We begin by some technical lemmas. We recall that $T_2 = T_2(z) \subseteq \{1,\dots,n\}$ denotes a set indexing any $2k$ largest components (in magnitude) of vector z , while $T = T(z) \subseteq \{1,\dots,n\}$ will denote a set indexing k largest components (in magnitude). Given an index set $\emptyset \neq H \subseteq \{1,\dots,n\}$, Q_H is the "cube" of all vectors $v \in \mathbb{R}^n$ such that $\texttt{supp}(v) = H$ and $|v_i| = 1$ for every $i \in H$. The restriction of v to H, $v_H \in \mathbb{R}^n$, is such that $(v_H)_i = v_i$, $i \in H$ and $\texttt{supp}(v_H) \subseteq H$.

Lemma 13 *Let $\Sigma = \Sigma_k$. Let $\|\cdot\|_w$ be a weighted ℓ^1-norm (for $w = (w_i)_{i=1}^n$ with $w_i > 0$, $\|x\|_w = \sum w_i \|x\|_1$). Let $z \in \mathcal{T}_{\|\cdot\|_w}(\Sigma)$. There is a support H of size $\leq k$ such that*

$$\|z_{H^c}\|_w - \|z_H\|_w = \inf_{x\in\Sigma} \{\|x+z\|_w - \|x\|_w\} \leq 0, \tag{84}$$

i.e., *the infimum is achieved at $x^* = -z_H$.*

Moreover, if $\|\cdot\|_w = \|\cdot\|_1$, $H = T(z)$.

Proof The result is trivial for $z = 0$, so we prove it for $z \in \mathcal{T}_{\|\cdot\|_w}(\Sigma)\backslash\{0\}$. Consider $H \in \arg\min_{T;|T|\leq k} \{\|z_{T^c}\|_w - \|z_T\|_w\}$. By definition of $\mathcal{T}_{\|\cdot\|_w}(\Sigma)$, since $z \in \mathcal{T}_{\|\cdot\|_w}(\Sigma) \setminus \{0\}$, there are $x' \in \Sigma$, $\lambda \in \mathbb{R} \setminus \{0\}$ such that $\|x' + \lambda z\|_w \leq \|x'\|_w$. By homogeneity of Σ, $x := x'/\lambda \in \Sigma$ and $\|x+z\|_w \leq \|x\|_w$. This shows that $\inf_{x\in\Sigma}\{\|x+z\|_w - \|x\|_w\} \leq 0$ as claimed. For any such $x \in \Sigma$, consider $T = \texttt{supp}(x)$.

By the reverse triangle inequality $|x_i + z_i| - |x_i| \geq -|z_i|$, we have

$$\|x+z_T\|_w - \|x\|_w = \sum_{i\in T} w_i(|x_i+z_i| - |x_i|) \geq -\sum_{i\in T} w_i|z_i| = -\|z_T\|_w \tag{85}$$

Hence $\|x+z\|_w - \|x\|_w = \|x+z_T\|_w + \|z_{T^c}\|_w - \|x\|_w \geq \|z_{T^c}\|_w - \|z\|_w \geq \|z_{H^c}\|_w - \|z_H\|_w$.

If $\|\cdot\|_w = \|\cdot\|_1$, let $T = T(z)$ and remark that $\|z_{H^c}\|_1 - \|z_H\|_1 \geq \|z_{T^c}\|_1 - \|z_t\|_1$

□

The following Lemma permits to construct and characterize elements of descent cones.

Lemma 14 *Assume that R and Σ are positively homogeneous. For every $v_0 \in \Sigma$ such that $R(v_0) > 0$ and any $v_1 \in \mathcal{H}$, we have that $z := v_1 - \alpha v_0 \in \mathcal{T}_R(\Sigma)$ where $\alpha = \max(R(v_1)/R(v_0), 1)$. If, in addition, Σ is homogeneous and R is even, we have conversely that any $z \in \mathcal{T}_R(\Sigma)$ can be written as $z = v_1 - v_0$ where $v_0 \in \Sigma$, $v_1 \in \mathcal{H}$, and $R(v_1) \leq R(v_0)$.*

Proof Since Σ is positively homogeneous, $x := \alpha v_0 \in \Sigma$, and $R(x+z) = R(\alpha v_0 + z) = R(v_1)$. If $R(v_1) > R(v_0)$ then $\alpha > 1$ and $R(x+z) = R(v_1) = \alpha R(v_0) = R(\alpha v_0) = R(x)$. Otherwise $\alpha = 1$ and $R(x+z) = R(v_1) \leq R(v_0) = R(x)$. In both cases we obtain that $z \in \mathcal{T}_R(x) \subseteq \mathcal{T}_R(\Sigma)$.

Regarding the second claim, when $z \in \mathcal{T}_R(\Sigma)$, by definition there exists $x \in \Sigma$, $u \in \mathcal{H}$ and $\gamma \in \mathbb{R}$ such that $z = \gamma u$ where $R(x+u) \leq R(x)$. Denote $v_0 := \gamma x$ and $v_1 := v_0 + z$. Since Σ is homogeneous, we have $v_0 \in \Sigma$. Since R is even and positively homogeneous, $R(v_1) = R(\gamma x + \gamma u) = |\gamma| R(x+u) \leq |\gamma| R(x) = R(\gamma x) = R(v_0)$. □

The next lemma permits to compare $B^s_\Sigma(R)$ with $B^s_\Sigma(\|\cdot\|_1)$ (see definition in (52)) which was calculated in [17] to characterize the necessary RIP condition for sparse recovery.

Lemma 15 *Let $\Sigma = \Sigma_k$ be the set of k-sparse vectors in $\mathbb{R}^n$ with $k < n/2$ and $1 \leq L \leq n - 2k$. Assume that R is positively homogeneous, subadditive, and nonzero.*

Consider

$$(H_0, v_0) \in \arg \max_{\substack{H\subseteq\{1,\dots,n\}:\ |H|=k \\ v\in Q_H}} R(v) \tag{86}$$

$$(H_1, v_1) \in \arg \min_{\substack{H\subseteq\{1,\dots,n\}\backslash H_0, |H|=k+L \\ v\in Q_H}} R(v). \tag{87}$$

1. *We have $R(v_0) > 0$, and for any H of size $k' \geq k$ and any $v \in Q_H$, we have*

$$R(v) \leq \frac{k'}{k} R(v_0). \tag{88}$$

If $R = R^\star = \|\cdot\|_1$ then we have indeed equality $R^\star(v) = \frac{k'}{k} R^\star(v_0)$.

2. We have

$$B_{\Sigma}^{2k+L}(R) := \sup_{z\in\mathcal{T}_R(\Sigma)\setminus\{0\}:|\mathtt{supp}(z)|=2k+L} \frac{\|z_{T_2^c}\|_2^2}{\|z_{T_2}\|_2^2} \geq \frac{\frac{L}{k}}{\max\left(\left(\frac{R(v_1)}{R(v_0)}\right)^2,1\right)+1} \geq \frac{\frac{L}{k}}{\left(\frac{L}{k}+1\right)^2+1}. \tag{89}$$

Proof As a preliminary observe that if $R^\star = \|\cdot\|_1$ then $R^\star(v) = |H|$ for any $H, v \in Q_H$, hence H_0, H_1 can be any pair of disjoint sets of respective sizes $k, k+L$, and $v_i \in Q_{H_i}$ can be arbitrary, for example $v_i = 1_{H_i}$. This yields $R^\star(v_0) = k$, $R^\star(v_1) = k+L$, hence $R^\star(v_1) = (1+L/k)R^\star(v_0)$.

To prove the first claim, consider $\{G_i\}_{1\leq i\leq\binom{k'}{k}}$ the collection of all subsets $G_i \subseteq H$ of size exactly k. Since $v \in Q_H$, we have $v_{G_i} \in Q_{G_i}$ for each i. Also, since $|G_i| = k$ for every i, by definition of H_0, v_0 we obtain $\max_i R(v_{G_i}) \leq R(v_0)$. Notice that given a coordinate $j \in H$, there are $\binom{k'-1}{k-1}$ sets G_i such that $j \in G_i$. With $\lambda := \frac{1}{\binom{k'-1}{k-1}}$ we get $v = \lambda\sum_i v_{G_i}$ hence by positive homogeneity and subadditivity of R (which imply convexity)

$$R(v) = R(\lambda\sum_{i=1}^{\binom{k'}{k}} v_{G_i}) \leq \sum_{i=1}^{\binom{k'}{k}} R(\lambda v_{G_i}) = \lambda\sum_{i}^{\binom{k'}{k}} R(v_{G_i}) \leq \frac{\binom{k'}{k}}{\binom{k'-1}{k-1}} R(v_0) = \frac{k'}{k}R(v_0). \tag{90}$$

This establishes (88). With $R = R^\star$, we have $R^\star(v) = \|v\|_1 = k'$ for $v \in Q_H$, hence $R^\star(v) = (k'/k)R^\star(v_0)$ as claimed.

For the sake of contradiction, assume that $R(v_0) \leq 0$. As we have just proved, this implies $R(v) \leq (n/k)R(v_0) \leq 0$ for every $v \in \{-1,+1\}^n = Q_H$ with $H = \{1,\ldots,n\}$. By convexity of R it follows that $R(v) \leq 0$ for each $v \in [-1,1]^n = \mathrm{conv}(Q_H)$, and by positive homogeneity,

$$R(v) \leq 0, \ \forall v \in \mathcal{H}. \tag{91}$$

Positive homogeneity and subadditivity also imply

$$0 = 0\cdot R(v_0) = R(0\cdot v_0) = R(0) = R(-v+v) \leq R(-v) + R(v) \overset{(91)}{\leq} R(-v)$$

for every $v \in \mathcal{H}$, hence $R(v) = 0$ on $\mathcal{H}$, which yields the desired contradiction since we assume that R is nonzero.

Regarding the second claim, since $2k+L \leq n$ there is indeed some H of size $k+L$ such that $H \cap H_0 = \emptyset$, hence H_1 is well defined. By construction, $H_1 \cap H_0 = \emptyset$. Since $R(v_0) > 0$, R is positively homogeneous and Σ is homogeneous, by Lemma 14, $z = -\alpha v_0 + v_1 \in \mathcal{T}_R(\Sigma)$ with $\alpha := \max(R(v_1)/R(v_0), 1)$. Observe that $|\mathtt{supp}(z)| = |H_0| + |H_1| = 2k+L$. Since $\alpha \geq 1$ and all nonzero entries of v_0, v_1 have magnitude one, a set of $2k$ largest components of z is $T_2 = H_0 \cup T_1'$ with T_1' any subset of H_1 with k components, and we obtain (89). once we observe that

$$\frac{\|z_{T_2^c}\|_2^2}{\|z_{T_2}\|_2^2} = \frac{L}{k\alpha^2+k} = \frac{L/k}{\alpha^2+1}.$$

□

Lemma 16 *Consider $c_\infty, c_1 > 0$, an integer $n \geq 2$, and the optimization problem*

$$\sup_{x\in\mathbb{R}_+^n:\|x\|_\infty\leq c_\infty;\|x\|_1\leq c_1} \|x\|_2^2. \tag{92}$$

If $c_1 \geq c_\infty$ then there exists $1 \leq L \leq n-1$ and $0 \leq \theta \leq 1$ such that

$$x^* := c_\infty(\underbrace{1,\ldots,1}_{L\geq 1},\theta,\ \underbrace{0,\ldots,0}_{n-(L+1)\geq 0})$$

is a maximizer. Otherwise a maximizer is $x^* = (c_1, 0, \ldots, 0)$.

Proof Standard compactness arguments show the existence of a maximizer x^*. We distinguish two cases:

- If $\|x^*\|_\infty < c_\infty$ then x^* is indeed a maximizer of the Euclidean norm under an ℓ^1 constraint, hence x^* is a Dirac: without loss of generality, $x^* = (c_1, 0, \ldots, 0)$ so that $c_1 = \|x^*\|_\infty < c_\infty$.

- Otherwise $\|x^*\|_\infty = c_\infty$, in which case we show that all entries of x^*, except at most one, are either zero or equal to c_∞. For the sake of contradiction, assume that x^* contains two distinct entries with values $0 < a < b < c_\infty$, then for small enough $t > 0$, replacing these entries with $0 < a - t < b + t < c_\infty$ and keeping all other entries unchanged would lead to a vector x satisfying $\|x\|_\infty = \|x^*\|_\infty = c_\infty$, $\|x\|_1 = \|x^*\|_1$. However, since $\|x\|_2^2 - \|x^*\|_2^2 = (a-t)^2 + (b+t)^2 - (a^2+b^2) = 2t^2 + 2(b-a)t > 0$. Since x^* has optimal objective value, this yields the desired contradiction. Since the objective value and the constraints are invariant to index permutations, there is thus a maximizer with the claimed shape, and we have $c_1 \geq \|x^*\|_1 \geq \|x^*\|_\infty = c_\infty$.

The two cases respectively correspond to $c_1 < c_\infty$ or $c_1 \geq c_\infty$, which are mutually exclusive, hence the conclusion. □

Lemma 17 ([17]) *Consider $\Sigma = \Sigma_k \subseteq \mathbb{R}^n$. We have*

$$B_\Sigma(\|\cdot\|_1) = \max_{1 \leq L \leq n-2k} \frac{\frac{L}{k}}{\left(\frac{L}{k}+1\right)^2 + 1}. \tag{93}$$

Proof With $B_\Sigma^s(R)$ defined in (52), and recalling the expression (49) of $B_\Sigma(R)$, we have

$$B_\Sigma(\|\cdot\|_1) = \max_{1 \leq L \leq n-2k} B_\Sigma^{2k+L}(\|\cdot\|_1)$$

By Lemma 15, $\frac{R^\star(v_1)}{R^\star(v_0)} = (L/k) + 1 > 1$ and $B_\Sigma^{2k+L}(\|\cdot\|_1) \geq \frac{\frac{L}{k}}{\left(\frac{R^\star(v_1)}{R^\star(v_0)}\right)^2+1} = \frac{\frac{L}{k}}{\left(\frac{L}{k}+1\right)^2+1}$. This implies

$$B_\Sigma(\|\cdot\|_1) \geq \max_{1 \leq L \leq n-2k} \frac{\frac{L}{k}}{\left(\frac{L}{k}+1\right)^2 + 1}. \tag{94}$$

and there only remains to show there is indeed equality. We isolate this result from [17] for completeness. This will also help understand the case of sparsity in levels in Appendix A.6.

First, we show we can restrict the maximization used to express $B_\Sigma(\|\cdot\|_1)$ (cf (49)) over vectors z having constant amplitude $\alpha > 0$ on $T(z)$.

Indeed, consider $z \neq 0$ such that $z \in \mathcal{T}_{\Sigma_k}(\|\cdot\|_1)$. By Lemma 13, we have $\|z_{T^c}\|_1 \leq \|z_T\|_1$ with $T = T(z)$ a set of k indices of components of largest magnitude of z. Assume that there are $i \neq j$ in T such that $|z_i| \neq |z_j|$. Let y such that $y_l = z_l$ for $l \notin \{i, j\}$ and $y_i = y_j = (|z_i| + |z_j|)/2$. The set T remains a support of k largest amplitudes in y, and $T_2 = T_2(z)$ remains a support of $2k$ largest amplitudes in y. Moreover, we have $\|y_T\|_1 = \|z_T\|_1 \geq \|z_{T^c}\|_1 = \|y_{T^c}\|_1 = \|-y_T + y\|_1$ hence we have $y \in \mathcal{T}_{\Sigma_k}(\|\cdot\|_1)$. Since $\|y_{T_2}\|_2^2 - \|z_{T_2}\|_2^2 = \|y_T\|_2^2 - \|z_T\|_2^2 = 2[(|z_i| + |z_j|)/2]^2 - |z_i|^2 - |z_j|^2 = -(|z_i| - |z_j|)^2/2 < 0$ and $\|y_{T_2^c}\|_2^2 = \|z_{T_2^c}\|_2^2$ we have $\|y_{T_2^c}\|_2^2 / \|y_{T_2}\|_2^2 > \|z_{T_2^c}\|_2^2 / \|z_{T_2}\|_2^2$.

Second, the same reasoning on $T' = T_2 \setminus T$, shows that we can further restrict the maximization used to define $B_\Sigma(\|\cdot\|_1)$ to vectors having constant amplitude $0 \leq \beta \leq \alpha$ over T'. This leads to

$$B_\Sigma(\|\cdot\|_1) = \sup_{z \neq 0 : \|z_{T^c}\|_1 \leq \|z_T\|_1} \frac{\|z_{T_2^c}\|_2^2}{\|z_{T_2}\|_2^2} = \sup_{\alpha,\beta:\alpha \geq \beta > 0} \sup_{x \in \mathbb{R}^{n-2k} : \|x\|_\infty \leq \beta, \|x\|_1 \leq k(\alpha-\beta)} \frac{\|x\|_2^2}{k(\alpha^2+\beta^2)}. \tag{95}$$

Using Lemma 16, the supremum with respect to x is reached with vectors with the shape

$$(\underbrace{\beta, \dots, \beta}_{L}, \theta, \quad \underbrace{0, \dots, 0}_{n-2k-(L+1) \geq 0})$$

with $0 \leq \theta \leq \beta$ and $0 \leq L \leq n - 2k - 1$. We deduce

$$\begin{aligned} B_\Sigma(\|\cdot\|_1) &= \sup_{\alpha,\beta:\alpha \geq \beta > 0} \sup_{\substack{L,\theta: 0 \leq L \leq n-2k-1, 0 \leq \theta \leq \beta \\ \theta \leq k\alpha - (k+L)\beta}} \frac{L\beta^2 + \theta^2}{k(\alpha^2+\beta^2)} \\ &= \max_{0 \leq L \leq n-2k-1} \sup_{\alpha,\beta:\alpha \geq \beta > 0} \sup_{\substack{\theta: 0 \leq \theta \leq \beta \\ \theta \leq k\alpha - (k+L)\beta}} \frac{L\beta^2 + \theta^2}{k(\alpha^2+\beta^2)} \end{aligned} \tag{96}$$

When $0 \leq \beta \leq k\alpha - (k+L)\beta$ we have

$$\sup_{\theta: 0 \leq \theta \leq \beta, \theta \leq k\alpha - (k+L)\beta} \frac{L\beta^2 + \theta^2}{k(\alpha^2+\beta^2)} = \frac{(L+1)\beta^2}{k(\alpha^2+\beta^2)} \tag{97}$$

while when $\beta \geq k\alpha - (k+L)\beta \geq 0$ we have

$$\sup_{\theta:0\leq\theta\leq\beta,\theta\leq k\alpha-(k+L)\beta} \frac{L\beta^2+\theta^2}{k(\alpha^2+\beta^2)} = \frac{L\beta^2+(k\alpha-(k+L)\beta)^2}{k(\alpha^2+\beta^2)}. \tag{98}$$

On the one hand, when $0 < \beta \leq \alpha$ satisfies $\beta \leq k\alpha - (k+L)\beta$ we have $\alpha \geq (1+(L+1)/k)\beta$ hence

$$\sup_{\theta:0\leq\theta\leq\beta,\theta\leq k\alpha-(k+L)\beta} \frac{L\beta^2+\theta^2}{k(\alpha^2+\beta^2)} = \frac{(L+1)\beta^2}{k(\alpha^2+\beta^2)} = \frac{(L+1)/k}{(\alpha/\beta)^2+1} \leq \frac{(L+1)/k}{[1+(L+1)/k]^2+1}. \tag{99}$$

On the other hand, when $0 < \beta \leq \alpha$ satisfies $\beta \geq k\alpha - (k+L)\beta \geq 0$ we have $(1+(L+1)/k))\beta \geq \alpha \geq (1+L/k)\beta$ and, denoting $g(t) := \frac{L/k+kt^2}{(1+L/k+t)^2+1}$ for $t \geq 0$, we get

$$\sup_{\theta:0\leq\theta\leq\beta,\theta\leq k\alpha-(k+L)\beta} \frac{L\beta^2+\theta^2}{k(\alpha^2+\beta^2)} = \frac{L\beta^2+(k\alpha-(k+L)\beta)^2}{k(\alpha^2+\beta^2)} = \frac{L/k+k[\alpha/\beta-(1+L/k)]^2}{(\alpha/\beta)^2+1} = g(\alpha/\beta-(1+L/k)). \tag{100}$$

A simple function study shows that $g'(t)$ is positively proportional to a second degree polynomial $P(t)$ with positive leading coefficient and such that $P(0) < 0$. It follows that there is $t_0 > 0$ such that $g'(t) \leq 0$ for $0 \leq t \leq t_0$ and $g'(t) \geq 0$ for $t \geq t_0$. Hence, g is decreasing on $[0, t_0]$ and increasing on $[t_0, +\infty)$, so that

$$g(\alpha/\beta-(1+L/k)) \leq \sup_{0\leq t\leq 1/k} g(t) = \max\left(g(0), g(1/k)\right) = \max\left(\frac{L/k}{(1+L/k)^2+1}, \frac{(L+1)/k}{(1+(L+1)/k)^2+1}\right).$$

As all of the above bounds also hold if $\beta = 0$, we obtain the claimed result.

□

Remark 3 The maximum value of $\frac{L/k}{((k+L)/k)^2+1}$ (with respect to L) is reached for L/k maximizing $f(u) = u/((u+1)^2+1)$ (which is maximized at $\sqrt{2}$ over $\mathbb{R}$). We verify that it matches the necessary RIP condition $\frac{1}{\sqrt{2}}$ from [17], $f(\sqrt{2}) = 2\sqrt{2}/(2+\sqrt{2})$ which gives $\gamma_\Sigma(\|\cdot\|_1) = (4+3\sqrt{2})/\sqrt{2} = \frac{\sqrt{2}+1}{\sqrt{2}-1}$.

A.4.2 Lemmas for the proof of Theorem 4

Given a matrix U, we denote $U_{k:l}$ the restriction of U to its rows $k, \dots, l$. We denote $O(n)$ the orthogonal group. Given a symmetric matrix z, we write $\texttt{eig}(z)$ the vector of eigenvalues ordered decreasingly with respect to their absolute value. Given a vector x of size n, we write $\texttt{diag}(x)$ the diagonal matrix with diagonal equal to x. To match the notations for the case of sparsity, given a matrix $z = U^T\texttt{diag}(w)U$, we write $z_H = U^T\texttt{diag}(w_H)U$ and Q_H as in the previous section. We denote $T = \{1, .., r\}$ and $T_2 = \{1, .., 2r\}$. We denote $\|\cdot\|_F$ the Frobenius norm.

Using the same demonstration as Lemma 13 we characterize the descent cones of of the nuclear norm.

Lemma 18 *Let $\Sigma = \Sigma_r$. Let $\|\cdot\|_w$ be a weighted nuclear-norm. Let $z \in \mathcal{T}_{\|\cdot\|_w}(\Sigma)$. There is a support H of size $\leq r$ such that*

$$\|z_{H^c}\|_w - \|z_H\|_w = \inf_{x\in\Sigma}\{\|x+z\|_w - \|x\|_w\} \leq 0, \tag{101}$$

i.e., *the infimum is achieved at $x^* = -z_H$. Moreover, if $\|\cdot\|_w = \|\cdot\|_*$, $H = T(z)$.*

Lemma 19 *Let $\Sigma = \Sigma_r$ be the set of $n \times n$ symmetric matrices with rank at most r with $r < n/2$, and $1 \leq L \leq n - 2r$. Assume R is positively homogeneous, subadditive and nonzero. Consider the supports $H_0 = \{1, 2, .., r\}$ and $H_1 = \{r+1, \dots, 2r+L\}$.*

$$(U_0, v_0) \in \arg\max_{U\in O(n), v\in Q_{H_0}} \|U^T\texttt{diag}(v)U\|_{\mathcal{A}}, \tag{102}$$

$$(U_1, v_1) \in \arg\min_{U\in O(n), v\in Q_{H_1}:\, U_{0,1:r}U^T_{r+1:2r+L}=0} \|U^T\texttt{diag}(v)U\|_{\mathcal{A}}. \tag{103}$$

1. *We have $R(U_0^T v_0 U_0) > 0$, and for any H of size $r' \geq r$, $V \in O(n)$ and $w \in Q_H$, we have*

$$R(V^T\texttt{diag}(w)V) \leq \frac{r'}{r}R(U_0^T v_0 U_0). \tag{104}$$

If $R = R^\star = \|\cdot\|_*$ then we have indeed equality $R(V^T\texttt{diag}(w)V) = \frac{r'}{r}R(U_0^T v_0 U_0)$.

2. We have

$$B_\Sigma^{L+2r}(R) := \sup_{z\in\mathcal{T}_{\|\cdot\|_\mathcal{A}}(\Sigma)\setminus\{0\}:|\mathtt{supp}(\mathtt{eig}(z))|=2r+L} \frac{\|z_{T_2^c}\|_F^2}{\|z_{T_2}\|_F^2} \geq \frac{\frac{L}{r}}{\left(\max\left(\frac{R(U_1^T\mathtt{diag}(v_1)U_1)}{R(U_0^T\mathtt{diag}(v_0)U_0)},1\right)\right)^2+1} \geq \frac{\frac{L}{r}}{\left(\frac{L}{r}+1\right)^2+1}. \tag{105}$$

Proof As a preliminary observe that if $R^\star = \|\cdot\|_*$ then $R^\star(V^TwV) = |H|$ for any $H, w \in Q_H, V \in O(n)$, hence $w_i \in Q_{H_i}$ can be arbitrary, for example $w_i = 1_{H_i}$. This yields $R^\star(U_0^T\mathtt{diag}(v_0)U_0) = r$, $R^\star(U_1^T\mathtt{diag}(v_1)U_1) = r+L$, hence $R^\star(U_1^T\mathtt{diag}(v_1)U_1) = (1+L/r)R^\star((U_0^T\mathtt{diag}(v_0)U_0)$.

To prove the first claim, consider $\{G_i\}_{1\leq i\leq\binom{r'}{r}}$ the collection of all subsets $G_i \subseteq H$ of size exactly r. Since $w \in Q_H$, we have $w_{G_i} \in Q_{G_i}$ for each i. Also, since $|G_i| = r$ for every i, by definition of H_0, v_0 and remarking that the maximization over $O(n)$ permits to consider any permutation of the support, we obtain $\max_i R(V^T\mathtt{diag}(v_{G_i})V) \leq R(U_0^T\mathtt{diag}(v_0)U_0)$.

Notice that given a coordinate $j \in H$, there are $\binom{r'-1}{r-1}$ sets G_i such that $j \in G_i$. With $\lambda := \frac{1}{\binom{r'-1}{r-1}}$, we get $V^T\mathtt{diag}(w)V = V^T\lambda\sum_i \mathtt{diag}(w_{G_i})V$ hence by positive homogeneity and subadditivity of R (which imply convexity)

$$\begin{aligned} R(V^TwV) = R(\lambda V^T\sum_{i=1}^{\binom{r'}{r}}\mathtt{diag}(w_{G_i})V) &\leq \sum_{i=1}^{\binom{r'}{r}} R(V^T\lambda\mathtt{diag}(w_{G_i})V) = \lambda\sum_i^{\binom{r'}{r}} R(V^T\mathtt{diag}(w_{G_i})V) \\ &\leq \frac{\binom{r'}{r}}{\binom{r'-1}{r-1}}R(U_0^T\mathtt{diag}(v_0)U_0) = \frac{r'}{r}R(U_0^T\mathtt{diag}(v_0)U_0). \end{aligned} \tag{106}$$

This establishes (104). With $R = R^\star$, we have $R^\star(V^T\mathtt{diag}(w)V) = \|w\|_1 = r'$ for $w \in Q_H$, hence $R^\star(V^T\mathtt{diag}(w)V) = (r'/r)R^\star(U_0^T\mathtt{diag}(v_0)U_0)$ as claimed.

For the sake of contradiction, assume that $R(U_0^T\mathtt{diag}(v_0)U_0) \leq 0$. As we have just proved, this implies $R(V^T\mathtt{diag}(w)V) \leq (n/k)R(U_0^T\mathtt{diag}(v_0)U_0) \leq 0$ for every $w \in \{-1,+1\}^n = Q_H$ with $H = \{1,\dots,n\}$ and $V \in O(n)$. By convexity of R it follows that $R(V^T\mathtt{diag}(w)V) \leq 0$ for each $w \in [-1,1]^n = \mathrm{conv}(Q_H)$, and by positive homogeneity,

$$R(V^T\mathtt{diag}(w)V) \leq 0, \quad \forall w \in \mathbb{R}^n. \tag{107}$$

Positive homogeneity and subadditivity also imply

$$\begin{aligned} 0 = 0\cdot R(U_0^T\mathtt{diag}(v_0)U_0) = R(0\cdot U_0^T\mathtt{diag}(v_0)U_0) = R(0) &= R(-V^T\mathtt{diag}(w)V + V^T\mathtt{diag}(w)V) \\ &\leq R(-V^T\mathtt{diag}(w)V) + R(V^T\mathtt{diag}(w)V) \\ &\overset{(107)}{\leq} R(-V^T\mathtt{diag}(w)V) \end{aligned}$$

for every $V^T\mathtt{diag}(w)V \in \mathcal{H}$, hence $R(V^T\mathtt{diag}(w)V) = 0$ on $\mathcal{H}$, which yields the desired contradiction since we assume that R is nonzero.

Regarding the second claim, since $2r+L \leq n$, by construction, $H_1 \cap H_0 = \emptyset$. Since $R(U_0^T\mathtt{diag}(v_0)U_0) > 0$, R is positively homogeneous and Σ is homogeneous, by Lemma 14, $z = -\alpha U_0^T\mathtt{diag}(v_0)U_0 + U_1^T\mathtt{diag}(v_1)U_1 \in \mathcal{T}_R(\Sigma)$ with $\alpha := \max(R(U_1^T\mathtt{diag}(v_1)U_1)/R(U_0^T\mathtt{diag}(v_0)U_0), 1)$. Observe that $|\mathtt{supp}(\mathtt{eig}(z))| = |H_0| + |H_1| = 2r+L$. Since $\alpha \geq 1$ and all nonzero entries of v_0, v_1 have magnitude one, a set of $2r$ largest components of $eig(z)$ is $T_2 = H_0 \cup T_1'$ with T_1' any subset of H_1 with k components, and we obtain (105). once we observe that

$$\frac{\|z_{T_2^c}\|_2^2}{\|z_{T_2}\|_2^2} = \frac{L}{r\alpha^2+r} = \frac{L/r}{\alpha^2+1}. \tag{108}$$

□

Lemma 20 *Let $\Sigma = \Sigma_r$. Then*

$$B_\Sigma(\|\cdot\|_*) = \max_{0\leq L\leq n-2r} \frac{\frac{L}{r}}{\left(\frac{L}{r}+1\right)^2+1}. \tag{109}$$

Proof We have $z \in \mathcal{T}_{\|\cdot\|_*}(\Sigma_r)$ is equivalent to $\|z_{T_2^c}\|_* + \|z_{T'}\|_* \leq \|z_T\|_*$ where $T' = \mathtt{supp}(z) \backslash (T_2^c \cup T)$ (Lemma 18). Hence

$$B_\Sigma^{L+2r}(\|\cdot\|_*) = \sup_{z : \|z_{T_2^c}\|_* + \|z_{T'}\|_* \leq \|z_T\|_*} \frac{\|z_{T_2^c}\|_F^2}{\|z_{T_2}\|_F^2}. \tag{110}$$

Using the fact that $\|z\|_* = \|\mathtt{eig}(z)\|_1$ and $\|z\|_F = \|\mathtt{eig}(z)\|_2$, we fall on the expression of $B_\Sigma^{L+2r}(\|\cdot\|_1)$ and get the result using Lemma 17. □

A.5 Proofs for Section 3.3

Proof (Proof of Lemma 10) The constant $\delta_\Sigma^{\mathtt{suff}}(R)$ [36][Eq. (5)] has the following expression:

$$\delta_\Sigma^{\mathtt{suff}}(R) = \inf_{z \in \mathcal{T}_R(\Sigma) \backslash \{0\}} \sup_{x \in \Sigma} \frac{-\mathcal{R}e\langle x, z\rangle}{\|x\|_{\mathcal{H}} \sqrt{\|x+z\|_\Sigma^2 - \|x\|_{\mathcal{H}}^2 - 2\mathcal{R}e\langle x, z\rangle}}. \tag{111}$$

Considering any nonzero $z \in \mathcal{H}$, since Σ is a union of subspaces and $\Sigma \cap S(1)$ is compact, by Lemma 9 the set $P_\Sigma(z)$ is not empty and $\langle P_\Sigma(z), z\rangle = \|P_\Sigma(z)\|_{\mathcal{H}}^2$ is unambiguous. Choosing an arbitrary $y \in P_\Sigma(z)$ and setting $x = -y$, we obtain

$$\begin{aligned}
\sup_{x \in \Sigma} \frac{-\mathcal{R}e\langle x, z\rangle}{\|x\|_{\mathcal{H}} \sqrt{\|x+z\|_\Sigma^2 - \|x\|_{\mathcal{H}}^2 - 2\mathcal{R}e\langle x, z\rangle}} &\geq \frac{\|P_\Sigma(z)\|_{\mathcal{H}}^2}{\|P_\Sigma(z)\|_{\mathcal{H}} \sqrt{\|z - P_\Sigma(z)\|_\Sigma^2 - \|P_\Sigma(z)\|_{\mathcal{H}}^2 + 2\|P_\Sigma(z)\|_{\mathcal{H}}^2}} \\
&= \frac{1}{\sqrt{\sup_{z \in \mathcal{T}_R(\Sigma) \backslash \{0\}} \frac{\|z - P_\Sigma(z)\|_\Sigma^2}{\|P_\Sigma(z)\|_{\mathcal{H}}^2} + 1}}.
\end{aligned}$$

Considering the infimum over $z \in \mathcal{T}_R(\Sigma) \setminus \{0\}$ yields the first claim. Let us now proceed to the second claim.

Given $z \in \mathcal{T}_R(\Sigma) \setminus \{0\}$, consider an arbitrary $x \in \Sigma$, and $V \in \mathcal{V}$ such that $x \in V$. With Fact A2, for every $v \in \mathcal{H}$, $\|v\|_\Sigma^2$ is the infimum of $\sum_i \lambda_i \|u_i\|_{\mathcal{H}}^2$ over convex decompositions $v = \sum_i \lambda_i u_i$ over Σ, hence there exists $u_i \in \Sigma$, $\lambda_i \geq 0$ such that $\sum_i \lambda_i = 1$, $\sum_i \lambda_i u_i = x + z$ and

$$\|x+z\|_\Sigma^2 = \sum_i \lambda_i \|u_i\|_{\mathcal{H}}^2.$$

Since $V \subset \Sigma$, $u_{i,V} := P_V u_i \in \Sigma$. By the additional assumption, since $u_i \in \Sigma$ we also have and $u_{i,V^\perp} := P_{V^\perp} u_i \in \Sigma$ for each i. Observe also that $P_{V^\perp} x = 0$. Hence, with the notations $z_V = P_V z$, $z_{V^\perp} = P_{V^\perp} z$, we have the convex decompositions

$$\begin{aligned}
z_{V^\perp} = P_{V^\perp}(x+z) &= \sum \lambda_i u_{i,V^\perp} \\
x + z_V = P_V(x+z) &= \sum \lambda_i u_{i,V}.
\end{aligned}$$

Using Jensen's inequality for the convex functions $\|\cdot\|_\Sigma^2$ and $\|\cdot\|_{\mathcal{H}}^2$ and the identity $\|v\|_\Sigma^2 = \|v\|_{\mathcal{H}}^2$ for $v \in \Sigma$ (Fact A1), we have

$$\begin{aligned}
\|z_{V^\perp}\|_\Sigma^2 + \|x + z_V\|_{\mathcal{H}}^2 \leq \sum_i \lambda_i \|u_{i,V^\perp}\|_\Sigma^2 + \sum_i \lambda_i \|u_{i,V}\|_{\mathcal{H}}^2 &= \sum_i \lambda_i \|u_{i,V^\perp}\|_{\mathcal{H}}^2 + \sum_i \lambda_i \|u_{i,V}\|_{\mathcal{H}}^2 \\
&= \sum_i \lambda_i \|u_i\|_{\mathcal{H}}^2 = \|x+z\|_\Sigma^2.
\end{aligned}$$

Since P_V is the (linear and self-adjoint) orthogonal projection onto V, we have $\mathcal{R}e\langle x, z_V\rangle = \mathcal{R}e\langle x, P_V z\rangle = \mathcal{R}e\langle P_V x, z\rangle = \mathcal{R}e\langle x, z\rangle$, and we obtain

$$\begin{aligned}
\|z_{V^\perp}\|_\Sigma^2 + \|z_V\|_{\mathcal{H}}^2 &\leq \|x+z\|_\Sigma^2 - \|x+z_V\|_{\mathcal{H}}^2 + \|z_V\|_{\mathcal{H}}^2 \\
\|z_{V^\perp}\|_\Sigma^2 + \|z_V\|_{\mathcal{H}}^2 &\leq \|x+z\|_\Sigma^2 - \|x\|_{\mathcal{H}}^2 - 2\mathcal{R}e\langle x, z_V\rangle \\
\|z_{V^\perp}\|_\Sigma^2 + \|z_V\|_{\mathcal{H}}^2 &\leq \|x+z\|_\Sigma^2 - \|x\|_{\mathcal{H}}^2 - 2\mathcal{R}e\langle x, z\rangle.
\end{aligned} \tag{112}$$

Using Cauchy-Schwarz inequality, we have $(\mathcal{R}e(\langle x,z\rangle))^2 = (\mathcal{R}e(\langle x,z_V\rangle))^2 \leq \|x\|_{\mathcal{H}}^2\|z_V\|_{\mathcal{H}}^2$. Denoting V_0 such that $P_{V_0}(z) \in P_\Sigma(z)$, we get

$$(\mathcal{R}e\langle x,z\rangle)^2\left(\|z_{V^\perp}\|_\Sigma^2 + \|z_V\|_{\mathcal{H}}^2\right) \leq \|x\|_{\mathcal{H}}^2\|z_V\|_{\mathcal{H}}^2\left(\|x+z\|_\Sigma^2 - \|x\|_{\mathcal{H}}^2 - 2\mathcal{R}e\langle x,z\rangle\right)$$

$$\frac{(\mathcal{R}e\langle x,z\rangle)^2}{\|x\|_{\mathcal{H}}^2\left(\|x+z\|_\Sigma^2 - \|x\|_{\mathcal{H}}^2 - 2\mathcal{R}e\langle x,z\rangle\right)} \leq \frac{\|z_V\|_{\mathcal{H}}^2}{\left(\|z_{V^\perp}\|_\Sigma^2 + \|z_V\|_{\mathcal{H}}^2\right)} = \frac{1}{\frac{\|z_{V^\perp}\|_\Sigma^2}{\|z_V\|_{\mathcal{H}}^2}+1} \leq \frac{1}{\frac{\|z-P_Vz\|_\Sigma^2}{\|P_\Sigma(z)\|_{\mathcal{H}}^2}+1}.$$

where the last inequality (we could use here the weaker alternative assumption $P_\Sigma(z) \cap \arg\min_{x\in\Sigma}\|x-z\|_\Sigma/\|x\|_{\mathcal{H}} \neq \emptyset$) uses that $z_{V^\perp} = z - P_Vz$ and $\|P_{V_0}z\|_{\mathcal{H}} = \|P_\Sigma(z)\|_{\mathcal{H}} \geq \|P_V(z)\|_{\mathcal{H}} = \|z_V\|_{\mathcal{H}}$. To conclude, we use the additional hypothesis $P_\Sigma(z) \subseteq \arg\min_{x\in\Sigma}\|x-z\|_\Sigma$, which implies $\|z-P_\Sigma(z)\|_\Sigma \leq \|z-P_Vz\|_\Sigma$ since $P_Vz \in \Sigma$

$$\sup_{x\in\Sigma}\frac{-\mathcal{R}e\langle x,z\rangle}{\|x\|_{\mathcal{H}}\sqrt{\|x+z\|_\Sigma^2 - \|x\|_{\mathcal{H}}^2 - 2\mathcal{R}e\langle x,z\rangle}} \leq \frac{1}{\sqrt{\sup\limits_{z\in\mathcal{T}_R(\Sigma)\setminus\{0\}}\frac{\|z-P_\Sigma(z)\|_\Sigma^2}{\|P_\Sigma(z)\|_{\mathcal{H}}^2}+1}}.$$

□

To replicate the proof used in the necessary case, we show a monotony property of $\|\cdot\|_\Sigma$.

Lemma 21 *Consider a model set $\Sigma \subset \mathcal{H}$, $\|\cdot\|_\Sigma$ the atomic "norm" induced by Σ, and $D : \mathcal{H} \to \mathcal{H}$ a linear operator. If $D\Sigma \subseteq \Sigma$ and $\|D\|_{op} := \sup_{\|v\|_{\mathcal{H}}\leq 1}\|Dv\|_{\mathcal{H}} \leq 1$ then*

$$\|Dv\|_\Sigma \leq \|v\|_\Sigma, \quad \forall v \in \mathcal{H}. \tag{113}$$

Proof Let λ_i, u_i such that $u_i \in \Sigma$, $\sum_i \lambda_i = 1$, $\sum_i \lambda_i u_i = v$. Denoting $u_i' = Du_i$ we have $u_i' \in \Sigma$ and $Dv = \sum \lambda_i u_i'$. By Jensen's inequality and the fact that $\|u\|_\Sigma = \|u\|_{\mathcal{H}}$ for any $u \in \Sigma$ (Fact A1), it follows that

$$\|Dv\|_\Sigma^2 \leq \sum \lambda_i \|u_i'\|_\Sigma^2 = \sum \lambda_i \|u_i'\|_{\mathcal{H}}^2 = \sum \lambda_i \|Du_i\|_{\mathcal{H}}^2 \leq \sum \lambda_i \|u_i\|_{\mathcal{H}}^2. \tag{114}$$

With Fact A2, $\|v\|_\Sigma^2$ is the infimum of the right hand side over all such decompositions $v = \sum \lambda_i u_i$. □

Corollary 5 *With $\Sigma := \Sigma_k$ the set of k-sparse vectors in $\mathcal{H} = \mathbb{R}^n$, we have:*

1. *the norm $\|\cdot\|_\Sigma$ is invariant by permutation and coordinate sign changes;*
2. *for any vectors $v, v' \in \mathcal{H}$ such that $|v_j| \leq |v_j'|$ for all j we have $\|v\|_\Sigma \leq \|v'\|_\Sigma$;*
3. *consider any vector z, and T_k a subset indexing k components of largest magnitude,* i.e., $\min_{i\in T}|z_i| \geq \max_{j\notin T}|z_j|$, *with $|T| = k$. Then*

$$\max_{|T|\leq k}\|z_T\|_\Sigma = \|z_{T_k}\|_\Sigma \tag{115}$$

$$\min_{|T|\leq k}\|z - z_T\|_\Sigma = \|z - z_{T_k}\|_\Sigma. \tag{116}$$

Proof We show the three properties separately.

- **Property 1:** Let π be a permutation of $(1,\dots,n)$ and $\epsilon_1,\dots,\epsilon_n \in \{\pm 1\}$. Define D by $(Du)_i = \epsilon_i u_{\pi(i)}$. Observe that $D\Sigma_k \subseteq \Sigma_k$ and $\|D\|_{op} = 1$. Conclude using Lemma 21 that $\|Du\|_\Sigma \leq \|u\|_\Sigma$ for any $u \in \mathcal{H}$. The same holds with $D' = D^{-1}$, hence $\|u\|_\Sigma = \|D^{-1}Du\|_\Sigma \leq \|Du\|_\Sigma$ for any u. This shows $\|D\cdot\|_\Sigma = \|\cdot\|_\Sigma$.
- **Property 2:** Given the assumptions on v, v', the linear operator defined by $(Du)_i = v_i u_i / v_i'$ if $v_i' \neq 0$ (and $(Du)_i = 0$ otherwise) satisfies $D\Sigma \subseteq \Sigma$ and $\|D\|_{op} \leq 1$ hence, using Lemma 21 again, $\|v\|_\Sigma = \|Dv'\|_\Sigma \leq \|v'\|_\Sigma$.
- **Property 3:** By the invariance by permutation and coordinate sign changes of $\|\cdot\|_\Sigma$, it is sufficient to prove the result when $z_1 \geq \dots \geq z_n \geq 0$ and $T_k = \{1,\dots,k\}$. Given T of size k, there is a permutation ϕ of $(1,\dots,n)$ such that $T = \{\phi(1),\dots,\phi(k)\}$ where $\phi(1) < \dots < \phi(k)$. It follows that $z_{\phi(i)} \leq z_i$ for $1 \leq i \leq k$. Hence by Property 2, we have $\|z_T\|_\Sigma = \|(z_{\phi(1)},\dots,z_{\phi(k)},0,\dots,0)\|_\Sigma \leq \|(z_1,\dots,z_k,0,\dots,0)\|_\Sigma = \|z_{T_k}\|_\Sigma$. A similar argument using T^c yields $\|z - z_T\|_\Sigma \geq \|z - z_{T_k}\|_\Sigma$.

□

Corollary 6 *With $\Sigma := \Sigma_r$ the set of matrices of rank lower than r in $\mathcal{H}$ the set of symmetric matrices in $\mathbb{R}^{n\times n}$, we have:*

1. *for any matrices* $V^T\mathtt{diag}(w)V, V^T\mathtt{diag}(w')V$ *with* $V \in O(n)$ *such that* $|w_j| \leq |w'_j|$ *for all* j *we have* $\|V^T\mathtt{diag}(w)V\|_\Sigma \leq \|V^T\mathtt{diag}(w')V\|_\Sigma$*;*
2. *For any symmetric matrix* z*, and* T_r *a subset indexing* r *components of largest magnitude of* $\mathtt{eig}(z)$, i.e.,

$$\min_{i\in T} |\mathtt{eig}(z)_i| \geq \max_{j\notin T} |\mathtt{eig}(z)_j|,$$

with $|T| = r$*. Then*

$$\max_{|T|\leq r} \|z_T\|_\Sigma = \|z_{T_r}\|_\Sigma \tag{117}$$

$$\min_{|T|\leq r} \|z - z_T\|_\Sigma = \|z - z_{T_r}\|_\Sigma. \tag{118}$$

Proof We show the two properties separately.

- **Property 1:** Given the assumptions on w, w', the linear operator defined by $Dz = V^TWVz$ where W is the diagonal matrix such that $W_{ii} = w_i/w'_i$ if $w'_i \neq 0$ (and $W_{ii} = 0$ otherwise) satisfies $D\Sigma \subseteq \Sigma$ and $\|D\|_{op} \leq 1$. We have $D(V^T\mathtt{diag}(w')V) = V^TWw'V = V^TwV$. With Lemma 21, we get $\|V^T\mathtt{diag}(w)V\|_\Sigma = \|D(V^T\mathtt{diag}(w')V)\|_\Sigma \leq \|V^T\mathtt{diag}(w')V\|_\Sigma$.
- **Property 2:** This property is direct using the eigenvalue decomposition

$$z = U^T\mathtt{diag}(\mathtt{eig}(z))U^T = U^T\mathtt{diag}(\mathtt{eig}(z)_T + \mathtt{eig}(z)_{T^c})U^T$$

and Property 1.

□

We now prove Lemma 11.

Proof (Proof of Lemma 11) Consider first $\Sigma = \Sigma_k$. First, the properties of $\|\cdot\|_\Sigma$ established in Corollary 5 directly show that the minimum of $\|x - z\|_\Sigma$ with respect to $x \in \Sigma$ is reached at any $x \in P_\Sigma(z)$. Then, we can write $\Sigma = \cup_{V\in\mathcal{V}} V$ where $V \in \mathcal{V}$ if, and only if there is an index set $T \subseteq \{1, \ldots, n\}$ such that $|T| \leq k$ and $V = \mathrm{span}(e_i)_{i\in T}$. Given $V \in \mathcal{V}$ and $u \in \Sigma_k$, let us show that $P_{V^\perp}u \in \Sigma_k$. Writing $V = \mathrm{span}(e_i)_{i\in T}$ where $|T| \leq k$, we have $P_V(u) = u_T$ and $P_{V^\perp}(u) = u_{T^c}$. As $\mathtt{supp}(u_{T^c}) \subseteq \mathtt{supp}(u)$ it follows that $\|u_{T^c}\|_0 \leq k$, hence $P_{V^\perp}(u) \in \Sigma_k$.

In the case of low rank matrices $\Sigma = \Sigma_r$. We take $\mathcal{V} = \{\mathrm{span}(U_i)_{i\in I}, |I| \leq r, \|U_i\|_F = 1, \mathtt{rank}(U_i) = 1, \langle U_i, U_j\rangle = 0, i \neq j\}$. With Corollary 6, the minimum of $\|x - z\|_\Sigma$ with respect to $x \in \Sigma$ is reached at any $x \in P_\Sigma(z)$. Let $z \in \Sigma_r$ and $V \in \mathcal{V}$. We have $P_V(z) = V_1^TS_1V_1$ has rank r' lower than r. We can write $z = V_1^TS_1V_1 + V_2^TS_2V_2$ with $V_1V_2^T = 0$. Hence $P_{V^\perp}(z)$ has rank at most $r - r' \leq r$ and $P_{V^\perp}(z) \in \Sigma_r$ otherwise z would be of rank greater than r.

□

We need the following Lemma to control $\|\cdot\|_\Sigma$.

Lemma 22 *Let* $\Sigma = \Sigma_k \subset \mathbb{R}^n$*. Then for any* v

$$\|v\|_\Sigma^2 \geq \frac{\|v\|_1^2}{k}. \tag{119}$$

Let $\Sigma = \Sigma_r$*. Then for any* v

$$\|v\|_\Sigma^2 \geq \frac{\|v\|_*^2}{r}. \tag{120}$$

Proof Case $\Sigma = \Sigma_k$: Let $\lambda_i \geq 0, u_i \in \Sigma$ such that $\|v\|_\Sigma^2 = \sum \lambda_i\|u_i\|_2^2$ and $v = \sum \lambda_i u_i$ from Fact A2. We have, by convexity

$$\|v\|_1 = \left\|\sum_i \lambda_i u_i\right\|_1 \leq \sum_i \lambda_i \|u_i\|_1. \tag{121}$$

Using the fact that $\|x\|_1 \leq \sqrt{k}\|x\|_2$ if $|\mathtt{supp}(x)| \leq k$ and the concavity of the square root,

$$\|v\|_1 \leq \sqrt{k}\sum_i \lambda_i\|u_i\|_2 \leq \sqrt{k}\sqrt{\sum_i \lambda_i\|u_i\|_2^2} = \sqrt{k}\|v\|_\Sigma. \tag{122}$$

Case $\Sigma = \Sigma_r$: Let $\lambda_i \geq 0, u_i \in \mathcal{S}$ such that $\|v\|_\Sigma^2 = \sum \lambda_i \|u_i\|_F^2$ and $v = \sum \lambda_i u_i$ from Fact A2. We have, by convexity

$$\|v\|_1 = \left\| \sum_i \lambda_i u_i \right\|_* \leq \sum_i \lambda_i \|u_i\|_*. \tag{123}$$

Using the fact that $\|x\|_* \leq \sqrt{r}\|x\|_F$ if $\mathtt{rank}(x) \leq r$ and the concavity of the square root,

$$\|v\|_* \leq \sqrt{k} \sum_i \lambda_i \|u_i\|_F \leq \sqrt{k} \sqrt{\sum_i \lambda_i \|u_i\|_F^2} = \sqrt{k}\|v\|_\Sigma. \tag{124}$$

□

A.5.1 Sparsity

We prove several intermediates lemmas to obtain $D_\Sigma(\|\cdot\|_1)$.

Lemma 23 *Consider $\Sigma = \Sigma_k$ the set of k-sparse vectors in $\mathcal{H} = \mathbb{R}^n$, and $0 \leq L \geq n-k$. We have*

$$D_\Sigma^{k+L}(\|\cdot\|_1) := \sup_{z \in \mathcal{T}_{\|\cdot\|_1}(\Sigma)\setminus\{0\} : |\mathtt{supp}(z)| = k+L} \frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_2^2} = \min\left(1, \frac{L}{k}\right). \tag{125}$$

Proof It was already proven in [36, Theorem 4.1] that $\delta_\Sigma^{\mathtt{suff}}(\|\cdot\|_1) \geq \frac{1}{\sqrt{2}}$ hence by Lemma 10

$$\sup_{z \in \mathcal{T}_{\|\cdot\|_1}(\Sigma)\setminus\{0\}} \frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_2^2} = D_\Sigma(\|\cdot\|_1) \leq 1. \tag{126}$$

Hence, $D_\Sigma^{k+L}(\|\cdot\|_1) \leq 1$

Consider H_0 of cardinality k, H_1 of cardinality L such that $H_0 \cap H_1 = \emptyset$ (this is possible as $k + L \leq n$), and define $z = \alpha 1_{H_0} + 1_{H_1}$ where $\alpha = \max(1, L/k)$. As $\alpha \geq 1$, a set of k largest components of z is $T = H_0$. Moreover, $\|z_{H_0}\|_1 = \alpha k = \max(k, L) \geq L = \|z_{H_1}\|_1 = \|z_{H_0^c}\|_1$.

We distinguish two cases:

- **Case 1:** $L \geq k$, from Lemma 22, $\|z_{T^c}\|_\Sigma^2 \geq \frac{1}{k}\|z_{H_1}\|_1^2 = L^2/k$. Moreover $\|z_T\|_2^2 = k\alpha^2 = L^2/k$, thus $\|z_{T^c}\|_\Sigma^2/\|z_T\|_2^2 \geq 1$. Combining with (126) yields $D_\Sigma^{k+L}(\|\cdot\|_1) = 1 = \min(1, L/k)$.
- **Case 2:** $L < k$**,** we have $z_{T^c} = z_{H_1} \in \Sigma_k$ hence $\|z_{T^c}\|_\Sigma^2 = \|z_{T^c}\|_2^2 = \|z_{H_1}\|_2^2 = L$ and $\|z_{T^c}\|_\Sigma^2/\|z_T\|_2^2 = L/k$. This shows that $D_\Sigma^{k+L}(\|\cdot\|_1) \geq L/k = \min(1, L/k)$. To conclude, we show that $D_\Sigma^{k+L}(\|\cdot\|_1) \leq L/k$. Consider *any* $z' \in \mathcal{T}_{\|\cdot\|_1}(\Sigma)$ such that $|\mathtt{supp}(z')| = k+L$, with Lemma 13, there is a support H of size lower than k such that, $\|z'_H\|_1 \geq \|z'_{H^c}\|_1$, let T a set of k largest components of z'. We have $\|z'_T\|_1 - \|z'_{T^c}\|_1 \geq \|z'_H\|_1 - \|z'_{H^c}\|_1$. As $\|z'\|_0 \leq k + L$ and $L < k$, $z'_{T^c} \in \Sigma_L \subset \Sigma_k$ hence $\|z'_{T^c}\|_\Sigma = \|z'_{T^c}\|_2$. Moreover, $|z'_i| \geq \|z'_{T^c}\|_\infty$ for any $i \in T$, hence $\|z'_T\|_2^2 \geq k\|z'_{T^c}\|_\infty^2$. As a result

$$\frac{\|z'_{T^c}\|_\Sigma^2}{\|z'_T\|_2^2} = \frac{\|z'_{T^c}\|_2^2}{\|z'_T\|_2^2} \leq \frac{L\|z'_{T^c}\|_\infty^2}{k\|z'_{T^c}\|_\infty^2} = L/k.$$

□

Lemma 24 *Let $\Sigma = \Sigma_k$ be the set of k-sparse vectors in $\mathbb{R}^n$ with $k < n/2$ and $1 \leq L \leq n-k$. Assume that R is positively homogeneous, subadditive and nonzero.*

Consider

$$(H_0, v_0) \in \arg \max_{\substack{H \subseteq \{1,\dots,n\}: \ |H|=k \\ v \in Q_H}} R(v) \tag{127}$$

$$(H_1, v_1) \in \arg \min_{\substack{H \subseteq \{1,\dots,n\}\setminus H_0, |H|=L \\ v \in Q_H}} R(v). \tag{128}$$

We have

$$D_\Sigma^{k+L}(R) := \sup_{z \in \mathcal{T}_R(\Sigma)\setminus\{0\} : |\mathtt{supp}(z)| = k+L} \frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_2^2} \geq \min\left(1, \frac{L}{k}\right). \tag{129}$$

Proof From Lemma 15, $R^\star(v_1) = \frac{L}{k} R^\star(v_0)$. Since $k + L \leq n$ there is indeed some H of cardinality L such that $H \cap H_0 = \emptyset$, hence H_1 is well defined. By construction, $H_1 \cap H_0 = \emptyset$. From Lemma 15, we also have $R(v_0) > 0$ and $R(v_1)/R(v_0) \leq L/k$.

Since $R(v_0) > 0$, R is positively homogeneous and Σ is homogeneous, by Lemma 14, $z = -\alpha v_0 + v_1 \in \mathcal{T}_R(\Sigma)$ with $\alpha := \max(R(v_1)/R(v_0), 1)$. Observe that $|\mathtt{supp}(z)| = |H_0| + |H_1| = k + L$. Since $\alpha \geq 1$ and all nonzero entries of v_0, v_1 have magnitude one, a set of k largest components of z is $T = H_0$. We have

$$\frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_2^2} = \frac{\|v_1\|_\Sigma^2}{k\alpha^2}. \tag{130}$$

With Lemma 22, $\|v_1\|_\Sigma^2 \geq \frac{\|v_1\|_1^2}{k} \geq \frac{L^2}{k}$ if $L \geq k$ and $\|v_1\|_\Sigma^2 = \|v_1\|_2^2$ otherwise (Fact A1). If $L \geq k$

$$\frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_2^2} \geq \frac{L^2}{k^2\alpha^2} \geq \frac{L^2}{k^2 \max(L/k, 1)^2} = 1. \tag{131}$$

If $L < k$,

$$\frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_2^2} = \frac{L}{k\alpha^2} \geq \frac{L}{k} \tag{132}$$

which leads to the conclusion.

□

A.5.2 Low rank

Lemma 25 *Consider $\Sigma = \Sigma_r$ the set of symmetric matrices of rank lower than r. For any $L \geq 0$ such that $r + L \leq n$ we have,*

$$D_\Sigma^{r+L}(\|\cdot\|_*) := \sup_{z \in \mathcal{T}_{\|\cdot\|_*}(\Sigma) \setminus \{0\} : rank(z) = r+L} \frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_F^2} = \min\left(1, \frac{L}{r}\right) \tag{133}$$

where z_T is z restricted to its r biggest eigenvalues, and $z_{T^c} = z - z_T$

Proof It was already proven in [36, Theorem 4.1] that $\delta_\Sigma^{\mathtt{suff}}(\|\cdot\|_*) \geq \frac{1}{\sqrt{2}}$ hence by Lemma 10

$$\sup_{z \in \mathcal{T}_{\|\cdot\|_*}(\Sigma) \setminus \{0\}} \frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_F^2} = D_\Sigma(\|\cdot\|_*) \leq 1. \tag{134}$$

Consider $H_0 = \{1, ..r\}$, $H_1 = \{r+1, .., r+L\}$, let $U \in O(n)$ and define $z = U^T \mathtt{diag}(\alpha 1_{H_0} + 1_{H_1})U$ where $\alpha = \max(1, L/r)$. As $\alpha \geq 1$, a set of r largest components of $\mathtt{eig}(z)$ is $T = H_0$. Moreover, $\|z_T\|_* = \alpha r = \max(r, L) \geq L = \|z - z_T\|_* = \|z_{T^c}\|_*$.

If $L \geq r$, from Lemma 22, $\|z_{T^c}\|_\Sigma^2 \geq \frac{1}{r}(\|z_{T^c}\|_*)^2 = L^2/r$. Moreover $\|z_T\|_F^2 = r\alpha^2 = L^2/r$, thus $\|z_{T^c}\|_\Sigma^2 / \|z_T\|_F^2 \geq 1$. Combining with (134) yields $D_L(\|\cdot\|_*) = 1 = \min(1, L/r)$.

If $L < r$, we have $z_{T^c} \in \Sigma_r$ hence $\|z_{T^c}\|_\Sigma^2 = L$ and $\|z_{T^c}\|_\Sigma^2 / \|z_T\|_2^2 = L/r$. This shows that $D_L(\|\cdot\|_*) \geq L/r = \min(1, L/r)$. To conclude, we show that $D_L(\|\cdot\|_*) \leq L/r$. Consider *any* $z' \in \mathcal{T}_{\|\cdot\|_*}(\Sigma)$ such that $|\mathtt{supp}(z')| = r + L$, with Lemma 18, there is a support r' and $H = 1, .., r'$ such that $\|z'_H\|_* \geq \|z'_{H^c}\|_*$, let T a set of r largest components of z'. We have $\|z'_T\|_* - \|z'_{T^c}\|_* \geq \|z'_H\|_* - \|z'_{H^c}\|_*$. As $\|\mathtt{eig}(z')\|_0 \leq r + L$ and $L < r$, $z'_{T^c} \in \Sigma_L \subset \Sigma_r$ hence $\|z'_{T^c}\|_\Sigma = \|z'_{T^c}\|_F$. Moreover, $|\mathtt{eig}(z')_i| \geq \|\mathtt{eig}(z'_{T^c})\|_\infty$ for any $i \in T$, hence $\|z'_T\|_F^2 \geq r\|\mathtt{eig}(z'_{T^c})\|_\infty^2$. As a result

$$\frac{\|z'_{T^c}\|_\Sigma^2}{\|z'_T\|_F^2} = \frac{\|z'_{T^c}\|_F^2}{\|z'_T\|_F^2} \leq \frac{L\|\mathtt{eig}(z'_{T^c})\|_\infty^2}{r\|\mathtt{eig}(z'_{T^c})\|_\infty^2} = L/r.$$

□

Lemma 26 *Let $\Sigma = \Sigma_r$ be the set of $n \times n$ symmetric matrices with rank at most r with $r < n/2$, and $1 \leq L \leq n - r$. Assume R is positively homogeneous, subadditive and nonzero. Consider the supports $H_0 = \{1, 2, .., r\}$ and $H_1 = \{r+1, \ldots, r+L\}$.*

$$(U_0, v_0) \in \arg \max_{U \in O(n), v \in Q_{H_0}} \|U^T \mathtt{diag}(v) U\|_{\mathcal{A}} \tag{135}$$

$$(U_1, v_1) \in \arg \min_{U \in O(n), v \in Q_{H_1} :\, U_{0,1:r} U_{r+1:r+L}^T = 0} \|U^T \mathtt{diag}(v) U\|_{\mathcal{A}}. \tag{136}$$

We have

$$D_\Sigma^{r+L}(R) := \sup_{z \in \mathcal{T}_R(\Sigma) \setminus \{0\} : |\mathtt{supp}(z)| = r+L} \frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_F^2} \geq \min\left(1, \frac{L}{r}\right). \tag{137}$$

Proof From Lemma 19, $R^\star(U_1^T \mathtt{diag}(v_1) U_1 = \frac{L}{r} R^\star(U_0^T \mathtt{diag}(v_0) U_0)$, $R(U_0^T \mathtt{diag}(v_0) U_0) > 0$ and

$$R(U_1^T \mathtt{diag}(v_1) U_1) / R(U_0^T \mathtt{diag}(v_0) U_0) \leq L/r.$$

Since $R(v_0) > 0$, R is positively homogeneous and Σ is homogeneous, by Lemma 14, $z = -\alpha U_0^T \mathtt{diag}(v_0) U_0 + U_1^T \mathtt{diag}(v_1) U_1 \in \mathcal{T}_R(\Sigma)$ with $\alpha := \max(R(U_1^T \mathtt{diag}(v_1) U_1)/R(U_0^T \mathtt{diag}(v_0) U_0), 1)$. Observe that $|\mathtt{supp}(\mathtt{eig}(z))| = |H_0| + |H_1| = r + L$. Since $\alpha \geq 1$ and all nonzero entries of v_0, v_1 have magnitude one, a set of r largest components of z is $T = H_0$. We have

$$\frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_F^2} = \frac{\|U_1^T \mathtt{diag}(v_1) U_1\|_\Sigma^2}{r\alpha^2}. \tag{138}$$

With Lemma 22, we have

$$\begin{cases} \|U_1^T \mathtt{diag}(v_1) U_1\|_\Sigma^2 \geq \frac{1}{r} \|U_1^T \mathtt{diag}(v_1) U_1\|_*^2 = \frac{L^2}{r} & \text{if } L \geq r \\ \|U_1^T \mathtt{diag}(v_1) U_1\|_\Sigma^2 = \|U_1^T \mathtt{diag}(v_1) U_1\|_F^2 & \text{otherwise (Fact A1).} \end{cases} \tag{139}$$

If $L \geq r$

$$\frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_F^2} \geq \frac{L^2}{r^2\alpha^2} \geq \frac{L^2}{r^2 \max(L/r, 1)^2} = 1. \tag{140}$$

If $L < r$,

$$\frac{\|z_{T^c}\|_\Sigma^2}{\|z_T\|_F^2} = \frac{L}{r\alpha^2} \geq \frac{L}{r} \tag{141}$$

which leads to the conclusion.

□

A.6 Proofs for Section 4

We extend notations for classical sparsity to sparsity in levels ($\Sigma = \Sigma_{k_1,k_2}$). For $z = (z_1, z_2) \in \mathcal{H}$, we we define the following projections $P_1(z) := z_1$ and $P_2(z) := z_2$ and denote $T = (S_1, S_2) = T(z)$ where for $i \in \{1, 2\}$, $S_i \subseteq \{1, \dots, n_i\}$ is a support containing k_i largest coordinates (in absolute value) of z_i, i.e. $|S_i| = k_i$ and $\min_{j \in S_1} |z_{i,j}| \geq \max_{j \in S_i^c} |z_{i,j}|$. For every $U = (U_1, U_2)$ where $U_i \subseteq \{1, \dots, n_i\}$ and $|U_i| = k_i$, we also have $\|(z_i)_{S_i}\|_1 \geq \|(z_i)_{U_i}\|_1$ hence $\|z_T\|_w \geq \|z_U\|_w$ and similary $\|z_{T^c}\|_w \leq \|z_{U^c}\|_w$.

We define similarly $T_2 = T_2(z) = (S_1', S_2')$ with S_i' containing $2k_i$ largest coordinates of z_i. We begin by simplifying the condition $z \in \mathcal{T}_{\|\cdot\|_w}(\Sigma) \setminus \{0\}$.

Lemma 27 *Let $w = (w_1, w_2) \in \mathbb{R}_+^2$. Let $\|\cdot\|_w = w_1 \|P_1(\cdot)\|_1 + w_2 \|P_2(\cdot)\|_1$ Let $z \in \mathcal{T}_{\|\cdot\|_w}(\Sigma_{k_1,k_2}) \setminus \{0\}$ then*

$$\|z_{T^c}\|_w \leq \|z_T\|_w. \tag{142}$$

Reciprocally,

$$\|z_{T^c}\|_w \leq \|z_T\|_w \tag{143}$$

implies $z \in \mathcal{T}_{\|\cdot\|_w}(\Sigma_{k_1,k_2})$.

Proof By definition, if $z \in \mathcal{T}_{\|\cdot\|_w}(\Sigma_{k_1,k_2}) \setminus \{0\}$ then there exists $x \in \Sigma_{k_1,k_2}$ and $\gamma \in \mathbb{R} \setminus \{0\}$ such that $z = \gamma y$ and $\|x + y\|_w \leq \|x\|_w$. With $U := \mathtt{supp}(x)$ we have $\|y_{U^c}\|_w + \|(x+y)_U\|_w = \|x + y\|_w \leq \|x\|_w = \|x_U\|_w$. By the triangle inequality this implies

$$\|y_{U^c}\|_w \leq \|x_U\|_w - \|(x+y)_U\|_w \leq \|y_U\|_w. \tag{144}$$

As $\gamma \neq 0$, we obtain $\|z_{U^c}\|_w \leq \|z_U\|_w$. We have

$$\|z_T\|_w \geq \|z_U\|_w \geq \|z_{U^c}\|_w \geq \|z_{T^c}\|_w. \tag{145}$$

□

To calculate $B_\Sigma(\|\cdot\|_w)$ (see definition in Corollary 3), we need a few lemmas.

Lemma 28 *Consider $w_1, w_2, k_1, k_2 > 0$ and $\beta_1, \beta_2, \lambda \geq 0$ and*

$$V := \min_{\alpha_1, \alpha_2 \geq 0} k_1\alpha_1^2 + k_2\alpha_2^2 \quad \text{s.t.} \quad \alpha_1 \geq \beta_1,\ \alpha_2 \geq \beta_2,\ k_1 w_1 \alpha_1 + k_2 w_2 \alpha_2 \geq \lambda \tag{146}$$

- *If* $\lambda < k_1 w_1 \beta_1 + k_2 w_2 \beta_2$ *then* $V = k_1 \beta_1^2 + k_2 \beta_2^2$.
- *If* $\lambda \geq k_1 w_1 \beta_1 + k_2 w_2 \beta_2$ *then the minimum is achieved at* α_1^*, α_2^* *such that* $k_1 w_1 \alpha_1^* + k_2 w_2 \alpha_2^* = \lambda$. *Moreover*
  - *if* $\lambda \geq (k_1 w_1^2 + k_2 w_2^2) \max(\beta_1/w_1, \beta_2/w_2)$ *then*

$$V = \min_{\alpha_1, \alpha_2 \geq 0, w_1 \alpha_1 + k_2 w_2 \alpha_2 = \lambda} k_1 \alpha_1^2 + k_2 \alpha_2^2 = \lambda^2/(k_1 w_1^2 + k_2 w_2^2);$$

  - *otherwise*

$$V = \min\left(k_1 \beta_1^2 + \frac{(\lambda - k_1 w_1 \beta_1)^2}{k_2 w_2^2}, k_2 \beta_2^2 + \frac{(\lambda - k_2 w_2 \beta_2)^2}{k_1 w_1^2} \right) > \lambda^2/(k_1 w_1^2 + k_2 w_2^2).$$

Proof Consider the change of variables $x = \sqrt{k_1}\alpha_1$, $y = \sqrt{k_2}\alpha_2$ and denote $x_0 := \sqrt{k_1}\beta_1$, $y_0 := \sqrt{k_2}\beta_2$, $a := \sqrt{k_1}w_1$, $b := \sqrt{k_2}w_2$. This leads to the equivalent problem

$$\min_{x,y \geq 0} x^2 + y^2 \;\text{ s.t. }\; x \geq x_0, y \geq y_0, ax + by \geq \lambda$$

which involves a convex objective to be minimized over a polyhedral constraint set. If $ax_0 + by_0 > \lambda$, i.e., if $k_1 w_1 \beta_1 + k_2 w_2 \beta_2 > \lambda$, then this problem is equivalent to

$$\min_{x,y \geq 0} x^2 + y^2 \;\text{ s.t. }\; x \geq x_0, y \geq y_0$$

which is minimized at (x_0, y_0), with value $x_0^2 + y_0^2 = k_1 \beta_1^2 + k_2 \beta_2^2$. Otherwise, the candidate optima must satisfy the constraint $ax + by = \lambda$, hence $y = (\lambda - ax)/b$ and the problem is equivalent to

$$\min_{x_0 \leq x \leq (\lambda - by_0)/a} x^2 + (ax - \lambda)^2/b^2. \tag{147}$$

The unconstrained minimum of (147) is at x^* satisfying $2x^* + 2a(ax^* - \lambda)/b^2 = 0$, *i.e.,* , $x^* = \frac{a\lambda}{a^2+b^2}$, leading to $y^* = (\lambda - ax^*)/b = \frac{b\lambda}{a^2+b^2}$ and to an optimal unconstrained problem value

$$(x^*)^2 + (y^*)^2 = \lambda^2/(a^2 + b^2) = \lambda^2/(k_1 w_1^2 + k_2 w_2^2).$$

This is also the value of the constrained minimum of (147), provided that $x_0 \leq x^* \leq (\lambda - by_0)/a$, i.e., that $\lambda \geq (a^2 + b^2)\max(x_0/a, y_0/b) = (k_1 w_1^2 + k_2 w_2^2)\max(\beta_1/w_1, \beta_2/w_2)$. Otherwise, the constrained minimum is either at $x = x_0$ and $y = (\lambda - ax_0)/b$, so that $x^2 + y^2 = x_0^2 + (\lambda - ax_0)^2/b^2$; or at $y = y_0$ and $x = (\lambda - by_0)/a$, so that $x^2 + y^2 = y_0^2 + (\lambda - by_0)^2/a^2$. The value at the optimum is then $\min(x_0^2 + (\lambda - ax_0)^2/b^2, y_0^2 + (\lambda - by_0)^2/a^2)$, which is necessarily larger than that of the unconstrained minimum. Once translated in terms of the original variables, this yields the result. □

Lemma 29 *Let* $\rho \geq 0$, $k_1, k_2, L_1, L_2, w_1, w_2, \lambda > 0$

$$\max_{\beta_1 \geq 0, \beta_2 \geq 0} \frac{L_1 \beta_1^2 + L_2 \beta_2^2}{\rho + k_1 \beta_1^2 + k_2 \beta_2^2} \;\text{ s.t. }\; w_1(k_1 + L_1)\beta_1 + w_2(k_2 + L_2)\beta_2 = \lambda \tag{148}$$

is equal to

$$\max_{i \in \{1,2\}} \frac{L_i \lambda^2}{\rho w_i^2 (k_i + L_i)^2 + k_i \lambda^2}. \tag{149}$$

Denoting i^* *the index maximizing this expression, the maximum is reached for* $\beta_{i^*} = \frac{\lambda}{w_{i^*}(k_{i^*} + L_{i^*})}$ *(and* $\beta_j = 0$ *for* $j \neq i$*).*

Proof Let $c \geq 0$. Observe that

$$\frac{L_1 \beta_1^2 + L_2 \beta_2^2}{\rho + k_1 \beta_1^2 + k_2 \beta_2^2} \geq c \tag{150}$$

is equivalent to

$$(L_1 - ck_1)\beta_1^2 + (L_2 - ck_2)\beta_2^2 \geq c\rho. \tag{151}$$

With the change of variable $b_i = w_i(k_i + L_i)\beta_i$ we have $b_1 + b_2 = \lambda$ and (151) reads

$$\frac{(L_1 - ck_1)}{w_1^2(k_1 + L_1)^2} b_1^2 + \frac{(L_2 - ck_2)}{w_2^2(k_2 + L_2)^2}(b_1 - \lambda)^2 \geq c\rho. \tag{152}$$

The left side is *maximized* (with respect to $0 \leq b_1 \leq \lambda$) for either $b_1 = 0$ or $b_1 = \lambda$. The initial inequality (150) is thus feasible if, and only if, the maximum of the left hand side of (152) over these two values verifies the inequality

$$\max_{i\in\{1,2\}} \frac{(L_i - ck_i)}{w_i^2(k_i + L_i)^2}\lambda^2 \geq c\rho \tag{153}$$

i.e., if there is $i \in \{1, 2\}$ such that $(L_i - ck_i)\lambda^2 \geq c\rho w_i^2(k_i + L_i)^2$. This is equivalent to $L_i\lambda^2 \geq c(\rho w_i^2(k_i + L_i)^2 + k_i\lambda^2)$ and

$$c \leq \frac{L_i\lambda^2}{\rho w_i^2(k_i + L_i)^2 + k_i\lambda^2}. \tag{154}$$

□

Lemma 30 *Consider $w_1, w_2, \beta_1, \beta_2, c \geq 0$ and*

$$V := \sup_{0\leq\theta_i\leq\beta_i, w_1\theta_1+w_2\theta_2\leq c} \theta_1^2 + \theta_2^2. \tag{155}$$

Denoting $(\ell, r) \in \{(1, 2), (2, 1)\}$ such that $w_\ell\beta_\ell \leq w_r\beta_r$, we have

1. *if $c < w_\ell\beta_\ell$ then $V = \max_{i\in\{1,2\}}(c/w_i)^2$;*
2. *if $w_\ell\beta_\ell \leq c < w_r\beta_r$ then $V = \max((c/w_r)^2, \beta_\ell^2 + [(c - w_\ell\beta_\ell)/w_r]^2$;*
3. *if $w_r\beta_r \leq c < w_1\beta_1 + w_2\beta_2$ then $V = \max_{(i,j)\in\{(1,2),(2,1)\}} \beta_i^2 + [(c - w_i\beta_i)/w_j]^2$;*
4. *if $c \geq w_1\beta_1 + w_2\beta_2$ then $V = \beta_1^2 + \beta_2^2$;*

Proof The optimum V is the maximization of a quadratic form within the intersection of a rectangle and a half-space delimited by an affine function. Using standard compactness arguments there exists at least a maximizer (θ_1^*, θ_2^*) of the considered expression. If $\theta_i^* < \beta_i$ for some $i \in \{1, 2\}$ then the constraint $c = w_1\theta_1^* + w_2\theta_2^*$ is satisfied (otherwise, we would have $0 \leq \theta_i^* < \beta_i$ and $w_1\theta_1 + w_2\theta_2 < c$, and we could exhibit other $\theta_i > \theta_i*$ still satisfying the constraints and such that $\theta_1^2 + \theta_2^2$ is increased), hence $w_1\beta_1 + w_2\beta_2 > w_1\theta_1^* + w_2\theta_2^* = c$.
Vice-versa if $w_1\beta_1 + w_2\beta_2 > c$ then since (θ_1^*, θ_2^*) satisfies all constraints we have $w_1\theta_1^* + w_2\theta_2^* \leq c < w_1\beta_1 + w_2\beta_2$, hence there is at least one index $i \in \{1, 2\}$ such that $\theta_i^* < \beta_i$. We can thus consider the following cases (depending on the shape of the domain):

- if $w_1\beta_1 + w_2\beta_2 \leq c$ then for each $i \in \{1, 2\}$, $\theta_i^* = \beta_i$ hence $V = \beta_1^2 + \beta_2^2$ as claimed;
- otherwise, i.e., if $w_1\beta_1 + w_2\beta_2 > c$, we have $w_1\theta_1^* + w_2\theta_2^* = c$ and we distinguish three cases:
  (a) $\theta_1^* < \beta_1$, $\theta_2^* < \beta_2$: then, since $\theta_2^* = (c - w_1\theta_1^*)/w_2$ where θ_1^* is a maximizer of $\theta_1^2 + [(c - w_1\theta_1)/w_2]^2$ under the constraint $0 \leq \theta_1$ and $c - w_1\theta_1 \geq 0$, there is $(i, j) \in \{(1, 2), (2, 1)\}$ such that $\theta_j^* = 0$ and $\theta_i^* = c/w_i$. This is feasible provided that $c/w_i < \beta_i$.
  (b) $\theta_1^* = \beta_1$, $\theta_2^* < \beta_2$, hence $\theta_2^* = (c - w_1\beta_1)/w_2$. This satisfies $0 \leq \theta_2^* < \beta_2$ if, and only if, $c \geq w_1\beta_1$.
  (c) $\theta_1^* < \beta_1$, $\theta_2^* = \beta_2$, hence $\theta_1^* = (c - w_2\beta_2)/w_1$. This is feasible provided that $c \geq w_2\beta_2$.
  We now discuss the possible cases depending on the value of c:
  - $c < w_\ell\beta_\ell$: (a) with any $(i, j) \in \{(1, 2), (2, 1)\}$ is feasible; (b)-(c) are unfeasible, hence $V = \max_{i\in\{1,2\}}(c/w_i)^2$.
  - $c \geq w_r\beta_r$: (a) is unfeasible; (b)-(c) are both feasible, hence the claimed value of V for this case.
  - $w_\ell\beta_\ell \leq c < w_r\beta_r$: (a) is feasible with (i, j) such that $c < w_i\beta_i$, *i.e.*, , with $(i, j) = (r, \ell)$, leading to a value $(\theta_j^*)^2 + (\theta_i^*)^2 = (c/w_i)^2 = (c/w_r)^2$; (b) is feasible provided that $c \geq w_1\beta_1$, i.e., that $(r, \ell) = (2, 1)$, leading to a value $(\theta_1^*)^2 + (\theta_2^*)^2 = \beta_1^2 + [(c - w_1\beta_1)/w_2]^2 = \beta_\ell^2 + [(c - w_\ell\beta_\ell)/w_r]^2$; similarly, (c) is feasible provided that $(r, \ell) = (2, 1)$, leading to a value $(\theta_2^*)^2 + (\theta_1^*)^2 = \beta_2^2 + [(c - w_2\beta_2)/w_1]^2 = \beta_\ell^2 + [(c - w_\ell\beta_\ell)/w_r]^2$. Overall, this leads to $V = \max((c/w_r)^2, \beta_\ell^2 + [(c - w_\ell\beta_\ell)/w_r]^2$.

As in the case of the ℓ^1 norm for sparsity and the nuclear norm for low-rank matrices, we compute $B_\Sigma(\|\cdot\|_w)$ (see definition in Corollary 3) via intermediate quantities $B^{L_1,L_2}(w)$ that we now introduce and control. We find an expression consistent with the ℓ^1 case.

Lemma 31 *Consider weights $w = (w_1, w_2)$ with $w_i > 0$ and integers $k_i \geq 0$. Denote for any integers $L_1, L_2 \geq 0$*

$$B^{L_1,L_2}(w) := \sup_{\substack{\alpha_i\geq\beta_i\geq 0, \beta_1+\beta_2>0 \\ \sum_{i=1}^2 (k_i w_i\alpha_i - w_i(k_i+L_i)\beta_i)=0}} \frac{\sum_{i=1}^2 L_i\beta_i^2}{\sum_{i=1}^2 k_i(\alpha_i^2 + \beta_i^2)}. \tag{156}$$

For $m \in \{1, 2\}$, consider

$$g_m(L_1, L_2, \alpha_1, \alpha_2, \beta_1, \beta_2) := \frac{L_1\beta_1^2 + L_2\beta_2^2 + [(\sum_{i=1}^2 (k_i w_i\alpha_i - (k_i + L_i)w_i\beta_i)/w_m]^2}{\sum_{i=1}^2 k_i(\alpha_i^2 + \beta_i^2)}.$$

We have

$$\sup_{\substack{\alpha_i,\beta_i:0\le\beta_i\le\alpha_i;\beta_1+\beta_2>0\\ \sum_{i=1}^2(k_i+L_i)w_i\beta_i\le\sum_{i=1}^2 k_iw_i\alpha_i}} g_m(L_1,L_2,\alpha_1,\alpha_2,\beta_1,\beta_2) \le B^{L_1,L_2}(w). \tag{157}$$

(158)

Proof First we show that there exist $\alpha_i^* \in \mathbb{R}_+, \beta_i^* \in \mathbb{R}_+$ such that

$$g_m(L_1,L_2,\alpha_1^*,\alpha_2^*,\beta_1^*,\beta_2^*) = \sup_{\substack{\alpha_i,\beta_i:0\le\beta_i\le\alpha_i;\beta_1+\beta_2>0\\ \sum_{i=1}^2(k_i+L_i)w_i\beta_i\le\sum_{i=1}^2 k_iw_i\alpha_i}} g_m(L_1,L_2,\alpha_1,\alpha_2,\beta_1,\beta_2) \tag{159}$$

with $0 \le \beta_i^* \le \alpha_i^*; \beta_1^* + \beta_2^* > 0$, and $\sum_{i=1}^2(k_i+L_i)w_i\beta_i^* \le \sum_{i=1}^2 k_iw_i\alpha_i^*$. Indeed, given any α_i, β_i satisfying these constraints, setting $\beta_j' = \beta_j/(\beta_1+\beta_2)$, $\alpha_j' = \alpha_j/(\beta_i+\beta_j)$, we have $g_m(L_1,L_2,\alpha_1',\alpha_2',\beta_1',\beta_2') = g_m(L_1,L_2,\alpha_1,\alpha_2,\beta_1,\beta_2)$ hence the supremum is unchanged if we impose $\beta_1'+\beta_2' = 1$ instead of $\beta_1+\beta_2 > 0$. Given any such pair β_1', β_2', Lemma 28 yields the optimum over α_i satisfying the constraints, and as the resulting expression is continuous with respect to β_j', the existence of a maximizer follows using a compactness argument.

We will soon prove that $\sum_i(k_i+L_i)w_i\beta_i^* = \sum_i k_iw_i\alpha_i^*$. If this equality is verified, since $0 \le \beta_i^* \le \alpha_i^*$, we obtain the desired result

$$\begin{aligned} g_m(L_1,L_2,\alpha_1^*,\alpha_2^*,\beta_1^*,\beta_2^*) &= \frac{\sum_{i=1}^2(L_i\beta_i^*)^2}{\sum_{i=1}^2 k_i((\alpha_i^*)^2+(\beta_i^*)^2)} \\ &\le \sup_{\substack{\alpha_i,\beta_i:0\le\beta_i\le\alpha_i;\beta_1+\beta_2>0\\ \sum_{i=1}^2(k_i+L_i)w_i\beta_i=\sum_{i=1}^2 k_iw_i\alpha_i}} \frac{\sum_{i=1}^2 L_i\beta_i^2}{\sum_{i=1}^2 k_i((\alpha_i)^2+(\beta_i)^2)} = B^{L_1,L_2}(w). \end{aligned} \tag{160}$$

For the sake of contradiction, assume that $\sum_i(k_i+L_i)w_i\beta_i^* < \sum_i k_iw_i\alpha_i^*$, then with the shorthand $C := g_m(L_1,L_2,\alpha_1^*,\alpha_2^*,\beta_1^*,\beta_2^*)$, we have

$$[(\sum_i k_iw_i\alpha_i^* - \sum_i(k_i+L_i)w_i\beta_i^*)/w_m]^2 + \sum_i(L_i-Ck_i)(\beta_i^*)^2 = C\sum_i k_i(\alpha_i^*)^2. \tag{161}$$

Since $g_m(L_1,L_2,\alpha_1,\alpha_2,\beta_1,\beta_2) \le C$ within the constraints of (157), (β_1^*,β_2^*) maximize

$$h(\beta_1,\beta_2) := [(\sum_i k_iw_i\alpha_i^* - \sum_i(k_i+L_i)w_i\beta_i)/w_v]^2 + \sum_i(L_i-Ck_i)(\beta_i)^2$$

among all β_1,β_2 such that $0 \le \beta_i \le \alpha_i^*$, $\beta_1+\beta_2 > 0$ and $\sum_{i=1}^2(k_i+L_i)w_i\beta_i \le \sum_{i=1}^2 k_iw_i\alpha_i^*$.

Consider $j \in \{1,2\}$.

If $C > L_j/k_j$, then h is decreasing with respect to β_j on the considered range, hence $\beta_j^* = 0$. Otherwise $C \le L_j/k_j$, and since h is a second degree polynomial in β_j with positive leading coefficient, its maximum is at one of the extremities of the optimization interval, i.e., since we assumed $\sum_i(k_i+L_i)w_i\beta_i^* < \sum_i k_iw_i\alpha_i^*$, at least one of the constraints $\beta_j^* = 0$, $\beta_j^* = \alpha_j^*$ is reached.

Since the optimum satisfies all constraints of (157), we have $\beta_1^*+\beta_2^* > 0$, hence in light of the above observations there is at least one index $j \in \{1,2\}$ such that $C \le L_j/k_j$, and for which we have $\beta_j^* = \alpha_j^* > 0$.

Since $\sum_{i=1}^2 k_iw_i\beta_i^* \le \sum_{i=1}^2(k_i+L_i)w_i\beta_i^* < \sum_{i=1}^2 k_iw_i\alpha_i^*$, both constraints $\beta_1^* = \alpha_1^*, \beta_2^* = \alpha_2^*$ cannot be reached at the same time hence there is $(i,j) \in \{(1,2),(2,1)\}$ such that $\beta_i^* = 0$, $\beta_j^* = \alpha_j^*$ and

$$C = g_m(L_1,L_2,\alpha_1^*,\alpha_2^*,\beta_1^*,\beta_2^*) = \frac{L_j(\beta_j^*)^2 + [(k_iw_i\alpha_i^* + k_jw_j\alpha_j^* - (k_j+L_j)w_j\beta_j^*)/w_m]^2}{k_i(\alpha_i^*)^2 + k_j(\alpha_j^*)^2 + k_j(\beta_j^*)^2} \tag{162}$$

$$= \frac{L_j(\alpha_j^*)^2 + [(k_iw_i\alpha_i^* - L_jw_j\alpha_j^*)/w_m]^2}{k_i(\alpha_i^*)^2 + 2k_j(\alpha_j^*)^2}. \tag{163}$$

This can be rewritten $(L_j-2Ck_j)(\alpha_j^*)^2 + [(k_iw_i\alpha_i^* - L_jw_j\alpha_j^*)/w_m]^2 = Ck_i(\alpha_i^*)^2$. Observe that any $\alpha_1,\alpha_2,\beta_1,\beta_2$ such that $\beta_i = 0$, $\beta_j = \alpha_j > 0$, $\alpha_i = \alpha_i^*$, and $L_jw_j\alpha_j \le k_iw_i\alpha_i^*$ satisfy the constraints of (157), hence $g_m(L_1,L_2,\alpha_1,\alpha_2,\beta_1,\beta_2) \le C$, or equivalently

$$(L_j-2Ck_j)(\alpha_j)^2 + [(k_iw_i\alpha_i^* - L_jw_j\alpha_j)/w_m]^2 \le Ck_i(\alpha_i^*)^2 \tag{164}$$

Thus, α_j^* maximizes the left hand side of (164) under the constraint $0 \leq L_j w_j \alpha_j \leq k_i w_i \alpha_i^*$. If $L - 2Ck_j \leq 0$, then the left hand side of (164) is decreasing with respect to α_j in the considered range, hence $\alpha_j^* = 0$, which is not possible since $0 < \beta_1 + \beta_2 = \beta_j^* = \alpha_j^*$. Therefore we must have $L_j - 2Ck_j > 0$, hence the left hand side of (164) is a second degree polynomial in α_j with positive leading coefficient. Its maximum is achieved at one extremity of the interval constraint : the case $\alpha_j^* = 0$ was already ruled out as impossible, hence $L_j w_j \alpha_j^* = k_i w_i \alpha_i^*$. This implies $(k_i + L_i) w_i \beta_i^* + (k_j + L_j) w_j \beta_j^* = (k_j + L_j) w_j \alpha_j^* = k_j w_j \alpha_j^* + k_i w_i \alpha_i^*$, which yields the desired contradiction to the assumption that $\sum_i (k_i + L_i) w_i \beta_i^* < \sum_i k_i w_i \alpha_i^*$.

□

Lemma 32 *Consider weights $w = (w_1, w_2)$ and integers k_i, n_i such that $1 \leq 2k_i < n_i$ and $\Sigma = \Sigma_{k_1,k_2} \subset \mathbb{R}^{n_1} \times \mathbb{R}_2^n$, $i \in \{1,2\}$. We have*

$$B_\Sigma(\|\cdot\|_w) = \max_{0 \leq L_i \leq n - 2k_i} B^{L_1,L_2}(w) \tag{165}$$

where $B^{L_1,L_2}(w)$ is defined in (156).

Proof We use the same proof method as in Lemma 17. With the notations $T = T(z), T_2 = T_2(z)$ from the beginning of Appendix A.6, denote $T' = T_2 \setminus T$ so that $\|z_{T_2^c}\|_w + \|z_{T'}\|_w = \|z_{T^c}\|_w$. By Lemma 27, we have

$$B_\Sigma(\|\cdot\|_w) = \sup_{z : z \neq 0, \|z_{T_2^c}\|_w + \|z_{T'}\|_w \leq \|z_T\|_w} \frac{\|z_{T_2^c}\|_2^2}{\|z_{T_2}\|_2^2}. \tag{166}$$

We now show that this expression can be simplified by maximizing over vectors z with a particular shape. Consider z a vector satisfying the constraint in (166). Replacing each entry z_i of z with its magnitude $|z_i|$ leaves the constraint (as well as the maximized quantity) unchanged, hence without loss of generality we can assume that z has nonnegative entries $z_i \geq 0$. Similarly we can assume without loss of generality that for each $i \in \{1,2\}$, the index set $S_i = [1, k_i]$ indexes k_i largest entries of $P_i(z)$ and $S_i' = [1, 2k_i]$ indexes $2k_i$ largest entries.

Given some $j \in \{1,2\}$, consider two (equal or distinct) indices in S_j and the vector $\tilde{z}$ obtained by keeping unchanged all entries of z, except those indexed by these indices which are replaced by their average. This has the following effect:

1. Each S_i (resp. S_i'), $i \in \{1,2\}$, is a set of k_i (resp. $2k_i$) largest coordinates of $P_i(\tilde{z})$, hence $T(\tilde{z}) = T = (S_1, S_2)$, $T_2(\tilde{z}) = T_2 = (S_1', S_2')$, $T'(\tilde{z}) = T' = T_2 \backslash T$, $\tilde{z}_{T_2^c} = z_{T_2^c}$, $\tilde{z}_{T'} = z_{T'}$, and the support of $P_i(\tilde{z})$, $i \in \{1,2\}$ is the same as that of $P_i(z)$.
2. Denoting $a, b \geq 0$ the values of the two considered entries, since $(a+b)/2 + (a+b)/2 = a + b$, we have $\|[P_j(\tilde{z})]_{S_j}\|_1 = \|[P_j(z)]_{S_j}\|_1$, and we obtain that $\|\tilde{z}_T\|_w = \|z_T\|_w$, hence $\tilde{z}$ still satisfies the optimization constraint;
3. As $\|\tilde{z}_{T_2^c}\|_2 = \|z_{T_2^c}\|_2$ and $\|\tilde{z}_{T_2}\|_2^2 - \|z_{T_2}\|_2^2 = 2[(a+b)/2]^2 - a^2 - b^2 = -(a-b)^2/2 \leq 0$, hence $\|\tilde{z}_{T_2^c}\|_2^2 / \|\tilde{z}_{T_2}\|_2^2 \geq \|\tilde{z}_{T_2^c}\|_2^2 / \|\tilde{z}_{T_2}\|_2^2$ where the inequality is strict as soon as $a \neq b$.

All the above imply that, without loss of generality, we can restrict the optimization to vectors z such that, for $i \in \{1,2\}$, all entries of z_{S_i} are equal. We denote $\alpha_i > 0$ their common value. A similar reasoning with $S_j' \backslash S_j$ instead of S_j shows that we can also assume without loss of generality that all entries of $z_{S_i' \backslash S_i}$, $i \in \{1,2\}$, are equal. We denote $\beta_i \geq 0$ their common value.

The value of the smallest component of $[P_i(z)]_{S_i}$ is α_i, while the smallest component of $[P_i(z)]_{S_i'}$ is $\min(\alpha_i, \beta_i)$. Denoting $x_i = P_i(z)_{(S_i')^c}$, we have $x_i \in \mathbb{R}_+^{n_i - 2k_i}$ and the largest component of $[P_i(z)]_{(S_i')^c}$ is $\|x_i\|_\infty$. Hence, S_i and S_i' are respectively a set of k_i and $2k_i$ largest components of $P_i(z)$ if, and only if, $\|x_i\|_\infty \leq \beta_i \leq \alpha_i$.

Finally, we observe that $\|z_T\|_w - \|z_{T'}\|_w - \|z_{T_2^c}\|_w = w_1 k_1 \alpha_1 + w_2 k_2 \alpha_2 - w_1 k_1 \beta_1 - w_2 k_2 \beta_2 - w_1 \|x_1\|_1 - w_2 \|x_2\|_1$, $\|z_{T_2^c}\|_2^2 = \|x_1\|_2^2 + \|x_2\|_2^2$ and $\|z_{T_2}\|_2^2 = k_1 \alpha_1^2 + k_2 \alpha_2^2 + k_1 \beta_1^2 + k_2 \beta_2^2$. This establishes

$$B_\Sigma(\|\cdot\|_w) = \sup_{\substack{\beta_i : \beta_i \geq 0 \\ \beta_1 + \beta_2 > 0}} \sup_{\alpha_i : \alpha_i \geq \beta_i} \sup_{\substack{\|x_i\|_\infty \leq \beta_i \\ \sum_{i=1}^2 w_i \|x_i\|_1 \leq \sum_{i=1}^2 k_i w_i (\alpha_i - \beta_i)}} \frac{\sum_{i=1}^2 \|x_i\|_2^2}{\sum_{i=1}^2 k_i (\alpha_i^2 + \beta_i^2)}, \tag{167}$$

where the restriction $\beta_1 + \beta_2 > 0$ simply follows from the fact that when $\beta_1 + \beta_2 = 0$ we have $x_1 = x_2 = 0$ which leads to a sub-optimal objective value. To show that the supremum in (167) is achieved, observe that both the constraints on $y := (\alpha_1, \alpha_2, \beta_1, \beta_2, x_1, x_2)$ and the quantity $f(y)$ that is maximized are invariant by multiplication by a positive constant factor. Hence, the supremum is unchanged if we add a scaling constraint. e.g. by fixing $\|y\|_\infty$. This leads to the supremum of a continuous function over a compact set (the unit ℓ_∞ ball), hence there exists $\alpha_i^*, \beta_i^*, x_i^*$ reaching the supremum in (167).

Thanks to Lemma 16, given the constraints (depending on α_i and β_i), the maximisation w.r.t x_i is reached with vectors with the shape

$$(\underbrace{\beta_i,\dots,\beta_i}_{L_i},\theta_i,\quad \underbrace{0,\dots,0}_{n_i-2k_i-(L_i+1)\geq 0}\quad)$$

with $0\leq\theta_i\leq\beta_i$, $0\leq L_i\leq n_i-2k_i-1$, including potentially $L_i=0$ (case of vector x_i with a single nonzero coordinate θ_i). We deduce

$$B_\Sigma(\|\cdot\|_w)=\sup_{\substack{\beta_i:\beta_i\geq 0\\ \beta_1+\beta_2>0}}\sup_{\alpha_i:\alpha_i\geq\beta_i}\sup_{\substack{L_i,\theta_i:0\leq L_i\leq n-2k_i-1,0\leq\theta_i\leq\beta_i\\ \sum_{i=1}^2 w_i\theta_i\leq\sum_{i=1}^2(k_iw_i\alpha_i-w_i(k_i+L_i)\beta_i)}}\frac{\sum_{i=1}^2 L_i\beta_i^2+\theta_i^2}{\sum_{i=1}^2 k_i(\alpha_i^2+\beta_i^2)}. \tag{168}$$

Hence, denoting

$$f(L_1,L_2,\alpha_1,\alpha_2,\beta_1,\beta_2):=\sup_{\substack{\theta_i:0\leq\theta_i\leq\beta_i\\ \sum_{i=1}^2 w_i\theta_i\leq\sum_{i=1}^2(k_iw_i\alpha_i-w_i(k_i+L_i)\beta_i)}}\frac{\sum_{i=1}^2 L_i\beta_i^2+\theta_i^2}{\sum_{i=1}^2 k_i(\alpha_i^2+\beta_i^2)}, \tag{169}$$

for parameters α_i,β_i,L_i such that $c:=\sum_{i=1}^2(k_iw_i\alpha_i-w_i(k_i+L_i)\beta_i)\geq 0$, we have

$$B_\Sigma(\|\cdot\|_w)=\max_{0\leq L_i\leq n_i-2k_i-1}\underbrace{\sup_{\substack{\beta_1,\beta_2\geq 0\\ \beta_1+\beta_2>0}}\sup_{\substack{\alpha_i:\alpha_i\geq\beta_i\\ \sum_{i=1}^2 k_iw_i\alpha_i\geq\sum_{i=1}^2(k_i+L_i)w_i\beta_i}} f(L_1,L_2,\alpha_1,\alpha_2,\beta_1,\beta_2)}_{f(L_1,L_2)}. \tag{170}$$

To continue, we bound $f(L_1,L_2)$ via characterizations of $f(L_1,L_2,\alpha_1,\alpha_2,\beta_1,\beta_2)$ in different parameter ranges. The supremum in (169) is covered by Lemma 30 hence we need to primarily distinguish cases depending on relative order of $c=\sum_{i=1}^2(k_iw_i\alpha_i-w_i(k_i+L_i)\beta_i)\geq 0$, $w_1\beta_1+w_2\beta_2$, $w_1\beta_1$, and $w_2\beta_2$. This suggests to write $f(L_1,L_2)=\max_{u\in\{0,1\}}f_u(L_1,L_2)$ where

$$f_0(L_1,L_2):=\sup_{\substack{\beta_i,\alpha_i:0\leq\beta_i\leq\alpha_i,\beta_1+\beta_2>0\\ \sum_{i=1}^2 k_iw_i\alpha_i\geq\sum_{i=1}^2(k_i+L_i+1)w_i\beta_i}} f(L_1,L_2,\alpha_1,\alpha_2,\beta_1,\beta_2) \tag{171}$$

$$f_1(L_1,L_2):=\sup_{\substack{\beta_i,\alpha_i:0\leq\beta_i\leq\alpha_i,\beta_1+\beta_2>0\\ \sum_{i=1}^2(k_i+L_i)w_i\beta_i\leq\sum_{i=1}^2 k_iw_i\alpha_i<\sum_{i=1}^2(k_i+L_i+1)w_i\beta_i}} f(L_1,L_2,\alpha_1,\alpha_2,\beta_1,\beta_2). \tag{172}$$

To express $f_0(L_1,L_2)$ and bound $f_1(L_1,L_2)$, we use the functions g_m, $m\in\{1,2\}$, from Lemma 31.

Expressing and bounding f_0: if $\sum_{i=1}^2 k_iw_i\alpha_i\geq\sum_{i=1}^2(k_i+L_i+1)w_i\beta_i$ then $c\geq w_1\beta_1+w_2\beta_2$ hence Lemma 30, case 4 yields

$$f(L_1,L_2,\alpha_1,\alpha_2,\beta_1,\beta_2)=\frac{\sum_{i=1}^2(L_i+1)\beta_i^2}{\sum_{i=1}^2 k_i(\alpha_i^2+\beta_i^2)} \tag{173}$$

$$f_0(L_1,L_2)=\sup_{\substack{0\leq\beta_i\leq\alpha_i,\beta_1+\beta_2>0\\ \sum_{i=1}^2 k_iw_i\alpha_i\geq\sum_{i=1}^2(k_i+L_i+1)w_i\beta_i}}\frac{\sum_{i=1}^2(L_i+1)\beta_i^2}{\sum_{i=1}^2 k_i(\alpha_i^2+\beta_i^2)} \tag{174}$$

$$\overset{\text{Lemma 28}}{=}\sup_{\substack{0\leq\beta_i\leq\alpha_i,\beta_1+\beta_2>0\\ \sum_{i=1}^2 k_iw_i\alpha_i=\sum_{i=1}^2(k_i+L_i+1)w_i\beta_i}}\frac{\sum_{i=1}^2(L_i+1)\beta_i^2}{\sum_{i=1}^2 k_i(\alpha_i^2+\beta_i^2)}=B^{L_1+1,L_2+1}(w). \tag{175}$$

As a result

$$f_0(L_1,L_2)\leq\max_{0\leq L_i\leq n_i-2k_i-1}B^{L_1+1,L_2+1}(w)\leq\max_{0\leq L_i'\leq n_i-2k_i}B^{L_1',L_2'}(w) \tag{176}$$

Bounding f_1: we denote $(\ell,r)\in\{(1,2),(2,1)\}$ a pair such that $w_\ell\beta_\ell=\min_i w_i\beta_i\leq\max_i w_i\beta_i=w_r\beta_r$.

When $\sum_{i=1}^2(k_i+L_i)w_i\beta_i\leq\sum_{i=1}^2 k_iw_i\alpha_i<\sum_{i=1}^2(k_i+L_i+1)w_i\beta_i$ we can distinguish three cases.

1. if $(k_\ell + L_\ell)w_\ell\beta_\ell + (k_r + L_r + 1)w_r\beta_r \leq \sum_{i=1}^2 k_i w_i \alpha_i < \sum_{i=1}^2 (k_i + L_i + 1)w_i\beta_i$ then $\max(w_1\beta_1, w_2\beta_2) = w_r\beta_r \leq c < w_1\beta_1 + w_2\beta_2$ hence Lemma 30, case 3 yields

$$f(L_1, L_2, \alpha_1, \alpha_2, \beta_1, \beta_2) = \max_{(u,v)\in\{(1,2),(2,1)\}} \underbrace{\frac{(L_u+1)\beta_u^2 + L_v\beta_v^2 + [(c - w_u\beta_u)/w_v]^2}{\sum_{i=1}^2 k_i(\alpha_i^2 + \beta_i^2)}}_{=g_v(L_1', L_2', \alpha_1, \alpha_2, \beta_1, \beta_2),\ L_u' = L_u+1, L_v' = L_v}. \tag{177}$$

2. if $(k_\ell + L_\ell + 1)w_\ell\beta_\ell + (k_r + L_r)w_r\beta_r \leq \sum_{i=1}^2 k_i w_i \alpha_i < (k_\ell + L_\ell)w_\ell\beta_\ell + (k_r + L_r + 1)w_r\beta_r$ then $\min(w_1\beta_1, w_2\beta_2) = w_\ell\beta_\ell \leq c < w_r\beta_r = \max(w_1\beta_1, w_2\beta_2)$ hence Lemma 30, case 2 yields

$$f(L_1, L_2, \alpha_1, \alpha_2, \beta_1, \beta_2) = \max\left(\underbrace{\frac{L_1\beta_1^2 + L_2\beta_2^2 + (c/w_r)^2}{\sum_{i=1}^2 k_i(\alpha_i^2 + \beta_i^2)}}_{g_r(L_1, L_2, \alpha_1, \alpha_2, \beta_1, \beta_2)}, \underbrace{\frac{(L_\ell+1)\beta_\ell^2 + L_r\beta_r^2 + [(c - w_\ell\beta_\ell)/w_r]^2}{\sum_{i=1}^2 k_i(\alpha_i^2 + \beta_i^2)}}_{g_r(L_1', L_2', \alpha_1, \alpha_2, \beta_1, \beta_2),\ L_\ell' = L_\ell+1, L_r' = L_r} \right). \tag{178}$$

3. otherwise $\sum_{i=1}^2 (k_i + L_i)w_i\beta_i \leq \sum_{i=1}^2 k_i w_i \alpha_i < (k_\ell + L_\ell + 1)w_\ell\beta_\ell + (k_r + L_r)w_r\beta_r$, hence $c < \min(w_1\beta_1, w_2\beta_2)$ and by Lemma 30, case 1

$$f(L_1, L_2, \alpha_1, \alpha_2, \beta_1, \beta_2) = \max\left(\underbrace{\frac{L_1\beta_1^2 + L_2\beta_2^2 + (c/w_1)^2}{\sum_{i=1}^2 k_i(\alpha_i^2 + \beta_i^2)}}_{g_1(L_1, L_2, \alpha_1, \alpha_2, \beta_1, \beta_2)}, \underbrace{\frac{L_1\beta_1^2 + L_2\beta_2^2 + (c/w_2)^2}{\sum_{i=1}^2 k_i(\alpha_i^2 + \beta_i^2)}}_{g_2(L_1, L_2, \alpha_1, \alpha_2, \beta_1, \beta_2)} \right). \tag{179}$$

Thus, in the range of α_i, β_i involved in the definition of $f_1(L_1, L_2)$ as a supremum, there are integers $0 \leq L_i' \leq n_i - 2k_i$ and $v \in \{1, 2\}$ such that $f(L_1, L_2, \alpha_1, \alpha_2, \beta_1, \beta_2) = g_v(L_1', L_2', \alpha_1, \alpha_2, \beta_1, \beta_2)$. We will shortly prove that given the relations between L_i' and the considered range of α_i, β_i we have

$$\sum_{i=1}^2 (k_i + L_i')w_i\beta_i \leq \sum_{i=1}^2 k_i w_i \alpha_i. \tag{180}$$

hence using Lemma 31 we obtain $g_v(L_1', L_2', \alpha_1, \alpha_2, \beta_1, \beta_2) \leq B^{L_1', L_2'}(w)$.

This implies

$$f_1(L_1, L_2) \leq \max_{0\leq L_i' \leq n_i - 2k_i} B^{L_1', L_2'}(w)$$

and, combined with (170)-(176), yields the upper bound

$$B_\Sigma(\|\cdot\|_w) = \max_{0\leq L_i \leq n_i - 2k_i} \max(f_0(L_1, L_2), f_1(L_1, L_2)) \leq \max_{0\leq L_i' \leq n_i - 2k_i} B^{L_1', L_2'}(w). \tag{181}$$

Proof of (180). We treat separately the three cases respectively associated to (177), (178), (179).

1. When $\sum_{i=1}^2 (k_i + L_i)w_i\beta_i \leq \sum_{i=1}^2 k_i w_i \alpha_i < (k_\ell + L_\ell + 1)w_\ell\beta_\ell + (k_r + L_r)w_r\beta_r$, by (179) there is $v \in \{1, 2\}$ such that $f(L_1, L_2, \alpha_1, \alpha_2, \beta_1, \beta_2) = g_v(L_1', L_2', \alpha_1, \alpha_2, \beta_1, \beta_2)$ with $(L_1', L_2') = (L_1, L_2)$. We observe that $\sum_{i=1}^2 (k_i + L_i')w_i\beta_i = \sum_{i=1}^2 (k_i + L_i)w_i\beta_i \leq \sum_{i=1}^2 k_i w_i \alpha_i$.
2. When $(k_\ell + L_\ell)w_\ell\beta_\ell + (k_r + L_r + 1)w_r\beta_r \leq \sum_{i=1}^2 k_i w_i \alpha_i < \sum_{i=1}^2 (k_i + L_i + 1)w_i\beta_i$, by (177), we have $f(L_1, L_2, \alpha_1, \alpha_2, \beta_1, \beta_2) = g_v(L_1', L_2', \alpha_1, \alpha_2, \beta_1, \beta_2)$ where $(L_1', L_2', v) \in \{(L_1 + 1, L_2, 2), (L_1, L_2 + 1, 1)\}$. If $(L_\ell', L_r') = (L_\ell, L_r + 1)$ then $\sum_{i=1}^2 (k_i + L_i')w_i\beta_i = (k_\ell + L_\ell)w_\ell\beta_\ell + (k_r + L_r + 1)w_r\beta_r$. Otherwise we have $(L_\ell', L_r') = (L_\ell + 1, L_r)$, hence $\sum_{i=1}^2 (k_i + L_i')w_i\beta_i = (k_\ell + L_\ell + 1)w_\ell\beta_\ell + (k_r + L_r)w_r\beta_r \leq (k_\ell + L_\ell)w_\ell\beta_\ell + (k_r + L_r + 1)w_r\beta_r$ since $w_\ell\beta_\ell \leq w_r\beta_r$ by definition of r, ℓ. In both cases we get $\sum_{i=1}^2 (k_i + L_i')w_i\beta_i \leq \sum_{i=1}^2 k_i w_i \alpha_i$.
3. When $(k_\ell + L_\ell + 1)w_\ell\beta_\ell + (k_r + L_r)w_r\beta_r \leq \sum_{i=1}^2 k_i w_i \alpha_i < (k_\ell + L_\ell)w_\ell\beta_\ell + (k_r + L_r + 1)w_r\beta_r$, (178) yields $f(L_1, L_2, \alpha_1, \alpha_2, \beta_1, \beta_2) = g_r(L_1', L_2', \alpha_1, \alpha_2, \beta_1, \beta_2)$ with $(L_\ell', L_r') \in \{(L_\ell, L_r), (L_\ell + 1, L_r)\}$, hence we have $\sum_{i=1}^2 (k_i + L_i')w_i\beta_i \leq (k_\ell + L_\ell + 1)w_\ell\beta_\ell + (k_r + L_r)w_r\beta_r \leq \sum_{i=1}^2 k_i w_i \alpha_i$.

As these three cases cover all possibilities, we deduce bound (180) as claimed.

To conclude, we obtain a lower bound on $B_\Sigma(\|\cdot\|_w)$. Consider any integers $0 \le L_i \le n_i - 2k_i$ and any scalars α_i, β_i such that $0 \le \beta_i \le \alpha_i$, $\beta_1 + \beta_2 > 0$ and $\sum_{i=1}^2 (k_i + L_i) w_i \beta_i = \sum_{i=1}^2 k_i w_i \alpha_i$, and let $z = (z_1, z_2)$ where

$$z_i = (\underbrace{\alpha_i, \dots, \alpha_i}_{k_i}, \underbrace{\beta_i, \dots, \beta_i}_{k_i + L_i}, \underbrace{0, \dots, 0}_{n_i - (2k_i + L_i)}).$$

We have $\|z_T\|_w = k_1 w_1 \alpha_1 + k_2 w_2 \alpha_2 = (k_1 + L_1) w_1 \beta_1 + (k_2 + L_2) w_2 \beta_2 = \|z_{T^c}\|_w$ hence, by Lemma 27 and the definition of $B_\Sigma(\|\cdot\|_w)$,

$$B_\Sigma(\|\cdot\|_w) \ge \frac{\|z_{T_2^c}\|_2^2}{\|z_{T_2}\|_2^2} = \frac{\sum_{i=1}^2 L_i \beta_i^2}{\sum_{i=1}^2 k_i (\alpha_i^2 + \beta_i^2)}. \tag{182}$$

Taking the supremum over α_i, β_i under the considered constraints yields $B_\Sigma(\|\cdot\|_w) \ge B^{L_1, L_2}(w)$. We deduce

$$B_\Sigma(\|\cdot\|_w) \ge \max_{0 \le L_i \le n_i - 2k_i} B^{L_1, L_2}(w).$$

□

We give a characterization/lower bound (depending on w) of the intermediate $B^{L_1, L_2}(w)$.

Lemma 33 *Consider $w = (w_1, w_2)$, $0 \le L_i \le n_i - 2k$, and $B^{L_1, L_2}(w)$ defined as in Lemma 31. We have*

$$\max_{(i,j) \in \{(1,2),(2,1)\}} \frac{L_i/k_i}{\frac{\nu_i}{1-\nu_i}(L_i/k_i)^2 + 2} \le B^{L_1, L_2}(w) \le \max_{(i,j) \in \{(1,2),(2,1)\}} \frac{L_i/k_i}{\nu_i (L_i/k_i + 1)^2 + 1} \tag{183}$$

with $\nu_i = \frac{1}{1 + k_j/(k_i \mu_i^2)}$ and $\mu_i = \frac{w_i}{w_j}$ for $(i,j) \in \{(1,2),(2,1)\}$. The rhs is an equality if $\nu_i \ge \frac{k_i}{k_i + L_i}$, $\forall i \in \{1,2\}$.

Proof For L_1, L_2 such that $L_1 + L_2 > 0$, we rewrite B^{L_1, L_2} defined in (156) as

$$B^{L_1, L_2}(w) = \sup_{\lambda > 0} \sup_{\substack{\beta_i : \beta_1, \beta_2 \ge 0 \\ \sum_{i=1}^2 w_i (k_i + L_i) \beta_i = \lambda}} \sup_{\substack{\alpha_i : \alpha_i \ge \beta_i \\ \sum_{i=1}^2 k_i w_i \alpha_i = \lambda}} \frac{L_1 \beta_1^2 + L_2 \beta_2^2}{k_1 \alpha_1^2 + k_2 \alpha_2^2 + k_1 \beta_1^2 + k_2 \beta_2^2}. \tag{184}$$

For fixed $\lambda > 0$ and β_1, β_2 such that $\sum_{i=1}^2 w_i (k_i + L_i) \beta_i = \lambda$, we have $\lambda > \sum_{i=1}^2 w_i k_i \beta_i$ hence, by Lemma 28,

$$B^{L_1, L_2}(w) \le \sup_{\lambda > 0} \sup_{\substack{\beta_1 \ge 0, \beta_2 \ge 0 \\ w_1 (k_1 + L_1) \beta_1 + w_2 (k_2 + L_2) \beta_2 = \lambda}} \frac{L_1 \beta_1^2 + L_2 \beta_2^2}{\frac{\lambda^2}{k_1 w_1^2 + k_2 w_2^2} + k_1 \beta_1^2 + k_2 \beta_2^2}. \tag{185}$$

with equality if the maximizers $\hat\lambda, \hat\beta_i$ of the right side satisfy the constraints $\hat\lambda \ge (k_1 w_1^2 + k_1 w_2^2) \max(\hat\beta_1 / w_1, \hat\beta_2 / w_2)$.

Consider $(i,j) \in \{(1,2),(2,1)\}$. Since $\nu_i := \frac{1}{1 + k_j w_j^2 / (k_i w_i^2)} = \frac{k_i w_i^2}{k_i w_i^2 + k_j w_j^2}$, we obtain by Lemma 29

$$B^{L_1, L_2}(w) \le \sup_{\lambda > 0} \max_{i \in \{1,2\}} \frac{L_i \lambda^2}{\frac{\lambda^2}{k_1 w_1^2 + k_2 w_2^2} w_i^2 (k_i + L_i)^2 + k_i \lambda^2} = \max_{(i,j) \in \{(1,2),(2,1)\}} \frac{L_i / k_i}{\nu_i (L_i/k_i + 1)^2 + 1}. \tag{186}$$

This establishes the upper bound in (183). Denoting (i^*, j^*) maximizing the right-hand-side expression above, and using the optimal values from Lemma 29, $\hat\beta_{i^*} = \frac{\hat\lambda}{w_{i^*}(k_{i^*} + L_{i^*})}$ (with $\hat\beta_{j^*} = 0$ and an arbitrary $\hat\lambda > 0$), we have $\max(\hat\beta_1 / w_1, \hat\beta_2 / w_2) = \hat\beta_{i^*} / w_{i^*} = \frac{\hat\lambda}{w_{i^*}^2 (k_{i^*} + L_{i^*})}$ hence equality holds in (186) if the following inequality is satisfied

$$(k_1 w_1^2 + k_1 w_2^2) \frac{1}{w_{i^*}^2 (k_{i^*} + L_{i^*})} \le 1,$$

or equivalently if $\frac{k_{i^*}}{\nu_{i^*}(k_{i^*} + L_{i^*})} \le 1$. This is guaranteed as soon as $\nu_\ell \ge \frac{k_\ell}{k_\ell + L_\ell}$ for every $\ell \in \{1,2\}$. This establishes the equality case in the rhs of (183).

We now treat the lower bound in (183). For fixed $\beta_i \ge 0$ and $\lambda > 0$ such that $(k_1 + L_1) w_1 \beta_1 + (k_2 + L_2) w_2 \beta_2 = \lambda$, we still have $\lambda > k_1 w_1 \beta_1 + k_2 w_2 \beta_2$. By Lemma 28, letting

$$V = \min_{\substack{\alpha_i : \alpha_i \ge \beta_i \\ \sum_{i=1}^2 k_i w_i \alpha_i = \lambda}} k_1 \alpha_1^2 + k_2 \alpha_2^2, \tag{187}$$

we either have

– $V = \min\left(k_1\beta_1^2 + k_2(\frac{\lambda - k_1w_1\beta_1}{k_2w_2})^2, k_2\beta_2^2 + k_1(\frac{\lambda-k_2w_2\beta_2}{k_1w_1})^2\right)$;
– or

$$V = \lambda^2/(k_1w_1^2 + k_2w_2^2) = \min_{\substack{\alpha_i:\alpha_i\geq 0\\ \sum_{i=1}^2 k_i\alpha_i=\lambda}} k_1\alpha_1^2 + k_2\alpha_2^2$$
$$\leq \min\left(k_1\beta_1^2 + k_2(\frac{\lambda - k_1w_1\beta_1}{k_2w_2})^2, k_2\beta_2^2 + k_1(\frac{\lambda-k_2w_2\beta_2}{k_1w_1})^2\right) \quad (188)$$

where the last inequality was obtained by evaluating $k_1\alpha_1^2 + k_2\alpha_2^2$ at $\alpha_1 = \beta_1$ (resp. at $\alpha_2 = \beta_2$) with α_2 (resp. α_1) tuned so that $k_1w_1\alpha_1 + k_2w_2\alpha_2 = \lambda$.

We deduce that $V \leq \min\left(k_1\beta_1^2 + k_2(\frac{\lambda-k_1w_1\beta_1}{k_2w_2})^2, k_2\beta_2^2 + k_1(\frac{\lambda-k_2w_2\beta_2}{k_1w_1})^2\right)$ and it follows using (184) that

$$\begin{aligned}
B^{L_1,L_2}(w) &\geq \sup_{\lambda>0} \sup_{\substack{\beta_i:\beta_1,\beta_2\geq 0\\ \sum_{i=1}^2 w_i(k_i+L_i)\beta_i=\lambda}} \frac{L_1\beta_1^2 + L_2\beta_2^2}{\min\left(k_1\beta_1^2 + k_2(\frac{\lambda-k_1w_1\beta_1}{k_2w_2})^2, k_2\beta_2^2 + k_1(\frac{\lambda-k_2w_2\beta_2}{k_1w_1})^2\right) + k_1\beta_1^2 + k_2\beta_2^2}\\
&= \sup_{\substack{\beta_i:\beta_1,\beta_2\geq 0\\ \sum_{i=1}^2 w_i(k_i+L_i)\beta_i=1}} \frac{L_1\beta_1^2 + L_2\beta_2^2}{\min\left(k_1\beta_1^2 + k_2(\frac{1-k_1w_1\beta_1}{k_2w_2})^2, k_2\beta_2^2 + k_1(\frac{1-k_2w_2\beta_2}{k_1w_1})^2\right) + k_1\beta_1^2 + k_2\beta_2^2}\\
&= \sup_{\substack{\beta_i:\beta_1,\beta_2\geq 0\\ \sum_{i=1}^2 w_i(k_i+L_i)\beta_i=1}} \max\left(\frac{L_1\beta_1^2 + L_2\beta_2^2}{2k_1\beta_1^2 + k_2\beta_2^2 + k_2(\frac{1-k_1w_1\beta_1}{k_2w_2})^2}, \frac{L_1\beta_1^2 + L_2\beta_2^2}{k_1\beta_1^2 + 2k_2\beta_2^2 + k_1(\frac{1-k_2w_2\beta_2}{k_1w_1})^2}\right).
\end{aligned} \quad (189)$$

For $(i,j) \in \{(1,2),(2,1)\}$, using the values $\tilde{\beta}_i = \frac{1}{w_i(k_i+L_i)}, \tilde{\beta}_j = 0$, we have

$$B^{L_1,L_2}(w) \geq \frac{L_1\tilde{\beta}_i^2}{2k_i\tilde{\beta}_i^2 + k_j(\frac{1-k_iw_i\tilde{\beta}_i}{k_jw_j})^2}. \quad (190)$$

Since $1 - k_iw_i\tilde{\beta}_i = w_i(k_i+L_i)\tilde{\beta}_i - k_iw_i\tilde{\beta}_i = w_iL_i\tilde{\beta}_i$, we have

$$\frac{L_i\tilde{\beta}_i^2}{2k_i\tilde{\beta}_i^2 + k_j(\frac{1-k_iw_i\tilde{\beta}_i}{k_jw_j})^2} = \frac{L_i\tilde{\beta}_i^2}{2k_i\tilde{\beta}_i^2 + k_j(\frac{w_iL_i}{k_jw_j})^2\beta_i^2} = \frac{L_i}{2k_i + k_j(\frac{w_iL_i}{k_jw_j})^2}.$$

Since $\nu_i = \frac{1}{1+k_jw_j^2/(k_iw_i^2)}$, we have $(1-\nu_i)/\nu_i = 1/\nu_i - 1 = k_jw_j^2/k_iw_i^2$. We deduce

$$B^{L_1,L_2}(w) \geq \frac{L_i}{2k_i + \frac{(w_iL_i)^2}{k_jw_j^2}} = \frac{L_i/k_i}{\frac{k_iw_i^2}{k_jw_j^2}(L_i/k_i)^2 + 2} = \frac{L_i/k_i}{\frac{\nu_i}{1-\nu_i}(L_i/k_i)^2 + 2}. \quad (191)$$

□

The following function study will be used to deal with the optimization of the $B^{L_1,L_2}(w)$.

Lemma 34 *Consider a such that $0 < a \leq 1$. The function*

$$g_1 : u \geq 0 \mapsto g_1(u;a) := \frac{u}{a(u+1)^2 + 1} \quad (192)$$

is maximized at $u_1^* = \sqrt{1+1/a}$, increasing for $u \leq u_1^*$, decreasing for $u \geq u_1^*$ and

$$g_1(u_1^*;a) = \frac{1}{2}(\sqrt{1+1/a} - 1). \quad (193)$$

Proof Since $g_1'(u;a) = \frac{-au^2+a+1}{(a(u+1)^2+1)^2}$, the equality $g_1'(u_1^*;a) = 0$ implies $a(u_1^*)^2 = a+1$ and $u_1^* = \sqrt{1+1/a}$. Given the sign of $g_1'(u;a)$, $g_1(\cdot;a)$ is increasing for $u \leq u_1^*$ and decreasing for $u \geq u_1^*$. As $a = [(u_1^*)^2 - 1]^{-1}$ we get

$$f_1(a) := g_1(u_1^*, a) = \frac{u_1^*((u_1^*)^2-1)}{(u_1^*+1)^2 + (u_1^*)^2 - 1} = \frac{u_1^*((u_1^*)^2-1)}{2(u_1^*)^2 + 2u_1^*} = \frac{1}{2}(u_1^* - 1) = \frac{1}{2}(\sqrt{1+1/a} - 1) \quad (194)$$

which is decreasing with respect to a.

Lemma 35 *Consider* $0 \le a \le 1$, $g_1(u;a) := \frac{u}{a(u+1)^2+1}$ *and* $g_2(u;a) := \frac{u}{\frac{a}{1-a}u^2+2}$ *and define*

$$h_i(u,v;a) := \max(g_i(u;a), g_i(v;1-a)), i \in \{1,2\}. \tag{195}$$

1. *For* $a = 1/2$ *we have* $h_1(u,v;1/2) \le \frac{\sqrt{3}-1}{2}$ *for every* $u, v \ge 0$.
2. *If* $a \notin [\tilde{a}, 1-\tilde{a}]$, *where* $\tilde{a} := 2\sqrt{3} - 3 \approx 0.46$ *then* $h_2(2,2,a) > \frac{\sqrt{3}-1}{2}$.

Consider integers such that $k_i \ge 2$, $1 \le 4k_i \le n_i$ *for* $i \in \{1,2\}$ *and*

$$H_1(a) := \max_{0 \le L_i \le n_i - 2k_i} h_1(L_1/k_1, L_2/k_2; a).$$

3. *There exists* $a^* \in [\tilde{a}, 1-\tilde{a}]$ *such that* $H_1(a^*) = \min_{a \in [\tilde{a}, 1-\tilde{a}]} H_1(a)$.
4. *Consider* $a \in [\tilde{a}, 1-\tilde{a}]$ *then*

$$H_1(a) = \max\left(\max_{L_1 \in \{\lfloor k_1\sqrt{1+1/a}\rfloor; \lceil k_1\sqrt{1+1/a}\rceil\}} g_1(L_1/k_1; a), \max_{L_2 \in \{\lfloor k_2\sqrt{1+1/(1-a)}\rfloor; \lceil k_2\sqrt{1+1/(1-a)}\rceil\}} g_1(L_2/k_2; 1-a) \right).$$

where $\lfloor\cdot\rfloor$ *and* $\lceil\cdot\rceil$ *denote the lower and upper integer part. Moreover, the* L_i^* *maximizing the above expression are such* $L_1^*/k_1 \ge 1/a - 1$ *and* $L_2^*/k_1 \ge 1/(1-a) - 1$.

Proof **Item 1.** By Lemma 34, with $a = 1/2$ and any $u, v \ge 0$ we have $g_1(u;a) \le g_1(u_1^*;a) = \frac{1}{2}(\sqrt{1+1/a} - 1) = (\sqrt{3}-1)/2$ and similarly $g_1(v;1-a) = (\sqrt{3}-1)/2$ hence $h_1(u,v;a) = (\sqrt{3}-1)/2$.

Item 2. We prove the inequality for $a < \tilde{a}$. Since $h_2(2,2;1-a) = h_2(2,2;a)$ by definition of h_2, the same inequality holds if $a > 1-\tilde{a}$. For $a < \tilde{a}$ since $a < 1/2$ we have $a/(1-a) < (1-a)/a$ hence using the definition of g_2 we have $g_2(2,a) > g_2(2,1-a)$. By monotonicity of $a \mapsto (1-a)/(1+a)$ we get

$$h_2(2,2;a) = \max(g_2(2,a), g_2(2,1-a)) = g_2(2,a) = \frac{2}{4\frac{a}{1-a}+2} = \frac{1-a}{1+a} > \frac{1-\tilde{a}}{1+\tilde{a}} = g_2(2,\tilde{a}).$$

Finally we compute

$$g_2(2,\tilde{a}) = \frac{1-(2\sqrt{3}-3)}{1+2\sqrt{3}-3} = \frac{4-2\sqrt{3}}{2\sqrt{3}-2} = \frac{(4-2\sqrt{3})(2\sqrt{3}+2)}{(2\sqrt{3})^2-2^2} = \frac{8\sqrt{3}-12+8-4\sqrt{3}}{8} = \frac{4\sqrt{3}-4}{8} = \frac{\sqrt{3}-1}{2}.$$

Item 3. The function H_1 is defined as the maximum of a finite number of continuous functions of a. By continuity of the maximum, H_1 is continuous and its minimum on the compact set $[\tilde{a}, 1-\tilde{a}]$ is reached.

Item 4. Consider $a \in [\tilde{a}, 1-\tilde{a}]$. By Lemma 34 the function $u \mapsto g_1(u;a)$ is maximized at a $u_1^* = \sqrt{1+1/a}$. Similarly, $v \mapsto g_1(v;1-a)$ is maximized at $u_2^* = \sqrt{1+1/(1-a)}$. Since $1/3 \le \tilde{a} \le 1/2$, with $\nu_1 = a$, $\nu_2 = 1-a$ we have $\nu_i \ge 1/3$ hence $u_i^* = \sqrt{1+1/\nu_i} \le 2$. Moreover since we assume $n_i \ge 4k_i$, we have $n_i - 2k_i \ge 2k_i \ge k_i\sqrt{1+1/\nu_i} = k_i u_i^*$. It follows that for $i \in \{1,2\}$ we have

$$\max_{0 \le L_i \le n_i - 2k_i} g_1(L_i/k_i; \nu_i) = \max_{L_i \in \{\lfloor k_i u_i^* \rfloor; \lceil k_i u_i^* \rceil\}} g_1(L_i/k_i; \nu_i).$$

As a result we have hence $H_1(a) = \max(g_1(L_1^*/k_1; a), g_1(L_2^*/k_2; 1-a))$ with

$$L_1^* \in \arg\max_{L_1 \in \{\lfloor k_1 u_1^* \rfloor; \lceil k_1 u_1^* \rceil\}} g_1(L_1/k_1; a)$$
$$L_2^* \in \arg\max_{L_2 \in \{\lfloor k_2 u_2^* \rfloor; \lceil k_2 u_2^* \rceil\}} g_1(L_2/k_2; 1-a).$$

There remains to show that $L_i^*/k_i \ge 1/\nu_1 - 1$. For this, we first observe that since $k_i \ge 2$ we have

$$L_i^*/k_i \ge \lfloor k_i u_i^* \rfloor / k_i \ge (k_i u_i^* - 1)/k_i = u_i^* - 1/k_i \ge u_i^* - 1/2 = \sqrt{1+1/\nu_i} - 1/2.$$

The derivative of $x \to \sqrt{1+x} - 1/2 - (x-1) = \sqrt{1+x} - x + 1/2$ at any $x \ge 0$ is $1/(2\sqrt{1+x}) - 1 \le -1/2$ hence this function is monotonically decreasing. Since $a \in [\tilde{a}, 1-\tilde{a}]$ and $\nu_1 = a$, $\nu_2 = 1-a$ we have $\nu_i \ge \tilde{a}$ hence $1/\nu_i \le 1/\tilde{a}$ for $i \in \{1,2\}$, hence

$$\sqrt{1+1/\nu_i} - 1/2 - (1/\nu_i - 1) \ge \sqrt{1+1/\tilde{a}} - 1/2 - (1/\tilde{a} - 1) \approx 0.12 > 0.$$

We deduce that $L_i^*/k_i \ge 1/\nu_i - 1$ as claimed. □

We can conclude with the proof of Theorem 7.

Proof (Proof of Theorem 7) The proof starts from the fact (Corollary 3) that

$$\arg\max_{R\in\mathcal{C}'} \delta_\Sigma^{\mathtt{nec}}(R) = \arg\min_{R\in\mathcal{C}'} B_\Sigma(R) \tag{196}$$

with $\mathcal{C}' = \{R(\cdot) = \|\cdot\|_w : w = (w_1, w_2), w_1 > 0, w_2 > 0\}$. Using Lemma 32, for each w we have

$$B_\Sigma(\|\cdot\|_w) = \max_{0\leq L_i\leq n-2k_i, i\in\{1,2\}} B^{L_1,L_2}(w). \tag{197}$$

With the notations of Lemma 33 we have $\mu_1 = w_1/w_2$ and $\mu_2 = w_2/w_1$ hence $\mu_1 = 1/\mu_2$, and one can check that $\nu_1 + \nu_2 = 1$ where $\nu_1 = \nu_1(w) := (1 + k_2 w_2^2/(k_1 w_1^2))^{-1}$. Hence, by Lemma 33 (taking $u = L_1/k_1, v = L_2/k_2, a = \nu_1$ and using (195)) and with the notation of Lemma 35, for all integers $0 \leq L_i \leq n_i - 2k_i$ we have

$$h_2(L_1/k_1, L_2/k_2; \nu_1) \leq B^{L_1,L_2}(w) \leq h_1(L_1/k_1, L_2/k_2; \nu_1) \tag{198}$$

with equality in the right hand s if for each $i \in \{1, 2\}$ we have $\nu_i \geq L_i/(k_i + L_i)$, i.e., $L_i/k_i \geq 1/\nu_i - 1$. Using (197) we get

$$\max_{0\leq L_i\leq n-2k_i, i\in\{1,2\}} h_2(L_1/k_1, L_2/k_2; \nu_1) \leq B_\Sigma(\|\cdot\|_w) \leq \max_{0\leq L_i\leq n-2k_i, i\in\{1,2\}} h_1(L_1/k_1, L_2/k_2; \nu_1), \tag{199}$$

and if the maximizers L_i^* of the right hand side of (199) satisfy $L_i^*/k_i \geq 1/\nu_i - 1$ for each $i \in \{1, 2\}$ then in fact

$$B_\Sigma(\|\cdot\|_w) = H_1(\nu_1) := h_1(L_1^*/k_1, L_2^*/k_2; \nu_1) \tag{200}$$

where H_1 is defined as the maximum of h_1 over the L_i/k_i (Lemma 35). Next we proceed in three steps. We set $\tilde{a} := 2\sqrt{3} - 3 \approx 0.46$.

Step 1. We show that if w', w'' are such that $\nu_1(w') \notin [\tilde{a}, 1 - \tilde{a}]$ and $\nu_1(w'') = 1/2 \in [\tilde{a}, 1 - \tilde{a}]$ then

$$B_\Sigma(\|\cdot\|_{w'}) > \frac{\sqrt{3}-1}{2} \geq B_\Sigma(\|\cdot\|_{w''}).$$

Hence the optimization of $w = (w_1, w_2)$ can be restricted to a range corresponding to $\nu_1 = \nu_1(w) \in [\tilde{a}, 1 - \tilde{a}]$.

Indeed, on the one hand, for $\nu_1 = 1/2$, by Lemma 35-Item 1 we have

$$\max_{0\leq L_i\leq n_i-2k_i} h_1(L_1/k_1, L_2/k_2; \nu_1 = 1/2) \leq \frac{\sqrt{3}-1}{2}$$

hence by the right-hand side in (199) we obtain $B_\Sigma(\|\cdot\|_w) \leq (\sqrt{3} - 1)/2$ as claimed.

On the other hand, if $\nu_1 \notin [\tilde{a}, 1 - \tilde{a}]$ then by Lemma 35-Item 2 we have $h_2(2, 2; \nu_1) > (\sqrt{3} - 1)/2$. Since $n_i \geq 4k_i$, the integers $L_i := 2k_i$, $i \in \{1, 2\}$ satisfy $0 < L_i \leq n_i - 2k_i$ hence, by the left-hand side in (199),

$$B_\Sigma(\|\cdot\|_{w'}) \geq h_2(L_1/k_1, L_2/k_2, \nu_1) = h_2(2, 2; \nu_1) > \frac{\sqrt{3}-1}{2}.$$

Step 2. We show that if w satisfies $\nu_1 = \nu_1(w) \in [\tilde{a}, 1 - \tilde{a}]$ then $B_\Sigma(\|\cdot\|_w) = H_1(\nu_1(w))$.

Since $k_i \geq 2$ and $n_i \geq 4k_i$, by Lemma 35-Item 4, we have the equality $H_1(\nu_1) = h_1(L_1^*/k_1, L_2^*/k_2, \nu_1)$ where $L_1^* \in \{\lfloor k_1\sqrt{1 + 1/\nu_1}\rfloor; \lceil k_1\sqrt{1 + 1/\nu_1}\rceil\}$, $L_2^* \in \{\lfloor k_2\sqrt{1 + 1/(1 - \nu_1)}\rfloor; \lceil k_2\sqrt{1 + 1/(1 - \nu_1)}\rceil\}$ and $L_i^*/k_i \geq 1/\nu_i - 1$. By (199)- (200) we deduce that the equality $B_\Sigma(\|\cdot\|_w) = H_1(\nu_1(w))$ holds.

Step 3. By Lemma 35-Item 3, there is $a^* \in [\tilde{a}, 1 - \tilde{a}]$ such that $H_1(a^*) = \min_{\tilde{a}\leq a\leq 1-\tilde{a}} H_1(a)$. In light of Steps 1 and 2, the infimum over w of $B_\Sigma(\|\cdot\|_w)$ is thus achieved, and a weight vector w^* satisfies

$$B_\Sigma(\|\cdot\|_{w^*}) = \min_w B_\Sigma(\|\cdot\|_w)) = H_1(a^*) \tag{201}$$

if, and only if $H_1(\nu_1(w^*)) = H_1(a^*)$. Since $\nu_1(w) = \left(1 + \frac{k_2}{k_1}(w_2/w_1)^2\right)^{-1}$, combining all of the above yields

$$\frac{w_2^*}{w_1^*} = \sqrt{\frac{k_1}{k_2}(1/\nu_1^* - 1)}$$

where ν_1^* is an optimum of

$$B_\Sigma(\|\cdot\|_{w^*}) = \min_{\nu_1\in[\tilde{a},1-\tilde{a}], \nu_2 = 1-\nu_1} \max_{i\in\{1,2\}} \max_{x_i\in\{\lfloor k_i\sqrt{1+1/\nu_i}\rfloor; \lceil k_i\sqrt{1+1/\nu_i}\rceil\}} g_1(x_i/k_i; \nu_i).$$

□