Computation of Minimal Projections and Extensions

Simon Foucart[*]— University of Georgia

Abstract

The determination of minimal projections is examined from an optimization theory viewpoint. It is first shown how to transform the problem into a linear program for the coordinate spaces ℓ_1^n and ℓ_∞^n. It is then shown how to transform the problem into a linear program for the matrix spaces $\mathcal{M}_{1\to 1}^{k\times k}$ and $\mathcal{M}_{\infty\to\infty}^{k\times k}$. The procedure is exploited to experimentally determine minimal projections onto various matrix subspaces. Moreover, a fully theoretical determination of minimal projections into zero-trace matrices is proposed when the matrix norm is unitarily invariant. Next, for polynomial spaces, it is shown how to approximate the problem by a linear program or by a semidefinite program using techniques from robust optimization. It allows us to tabulate the relative projection constants of several polynomial subspaces. The article finishes by illustrating that the underlying method also applies to the determination of minimal extensions rather than merely minimal projections. All the programs have been implemented in a publicly available MATLAB package called MinProj, which uses the external packages CVX and Chebfun.

Key words and phrases: projection constants, minimal projections, coordinate spaces, matrix spaces, polynomial spaces, robust optimization, semidefinite programming.

AMS classification: 46B20, 49M29, 65K05, 90C22, 90C47.

1 Introduction

This article reflects primarily on the computability of minimal projections using modern tools from optimization theory. It is accompanied by a collection of MATLAB routines performing the tasks described hereafter. The resulting package can be downloaded under the name MinProj from the author's webpage. It should prove useful for gaining experimental insight even to researchers focusing on theoretical investigations about minimal projections. MinProj relies heavily on CVX, a package for specifying and solving convex programs [1]. The package Chebfun [11] is also required when dealing with polynomial spaces.

Minimal projections are a classical topic in Approximation Theory, see e.g. the survey [10] and the papers [7, 9, 16, 17] for a selection of newer results. Our contribution consists essentially in adding

[*]S. F. partially supported by NSF grant number DMS-1120622

an extra perspective to the body of work that preceded the era of computational convenience. The general problem, which falls naturally in the realm of the modern field of Robust Optimization [2, 3], is described in Section 2, where recurring aspects used for all cases are gathered. In Section 3, we concentrate on the simple case of the coordinate space $\mathbb{R}^n$ equipped with one of the polyhedral norms ℓ_1 or ℓ_∞. The resulting linear program is validated by comparison with known formulas, before being utilized in a numerical exploration of average projection constants. Section 4 considers minimal projections in the matrix space $\mathcal{M}^{k\times k}$, traditionally less emphasized in the theoretical literature. In the case of polyhedral norms, we again propose a linear program which is used to find by experimental means the minimal projections and projection constants for a number of subspaces. In the case of unitarily-invariant norms, we find by theoretical means the minimal projections and projection constants for the subspaces of symmetric matrices and of zero-trace matrices. Section 5 is devoted to minimal projections in spaces of algebraic polynomials equipped with the max-norm on $[-1,1]$. We suggest to approximate the original approximation problem by two quantifiably close problems that are made tractable via ideas from Robust Optimization. The first problem becomes a linear program, while the second problem becomes a semidefinite program (see [18]) and they are both utilized to compute the projection constants of polynomial subspaces. Finally, Section 6 gives a brief illustration of the potential of the main strategy to compute minimal extensions in addition to minimal projections.

The following notational conventions are used throughout the article: a blackboard-bold uppercase letter denotes a vector space, e.g. $\mathbb{U},\mathbb{V},\mathbb{W}$; a curly uppercase letter a linear map, e.g. $\mathcal{A},\mathcal{P}$; a straight uppercase letter the matrix of a linear map, e.g., A,P; a bold lowercase letter a vector, e.g. $\mathbf{u},\mathbf{v}\,\mathbf{w}$; a bold uppercase letter a matrix understood as an element of a linear space, e.g. $\mathbf{M}$. The normed spaces manipulated in the article are real, not complex. They feature an indication of the dimension as a superscript and an indication of the norm as a subscript, e.g., ℓ_1^n, $\mathcal{M}^{k\times k}_{\infty\to\infty}$, or $\mathbb{P}^n_\infty$. The operator norm of a linear map from a normed space $\mathbb{V}$ into itself is denoted with a triple bar, i.e.,

$$|||\mathcal{A}||| = \max_{\mathbf{v}\in\mathbb{V}\setminus\{\mathbf{0}\}} \frac{\|\mathcal{A}\mathbf{v}\|}{\|\mathbf{v}\|},$$

and we may add a subscript referring to the norm used on $\mathbb{V}$, for instance, $|||\mathcal{A}|||_{\infty\to\infty}$ if $\mathbb{V}=\ell_\infty^n$.

2 The generic approach

2.1 Recasting the minimization problem

Throughout the article, let $\mathbb{V}$ be a finite-dimensional normed space and let $\mathbb{U}$ be a subspace of $\mathbb{V}$. A projection $\mathcal{P}$ from $\mathbb{V}$ onto $\mathbb{U}$ is a linear map from $\mathbb{V}$ into $\mathbb{V}$ that satisfies

(i) $\mathcal{P}\mathbf{u} = \mathbf{u}$ for all $\mathbf{u} \in \mathbb{U}$,

(ii) $\mathcal{P}\mathbf{v} \in \mathbb{U}$ for all $\mathbf{v} \in \mathbb{V}$.[1]

We point out that $\mathcal{P}$ can be interpreted as an extension of the canonical injection $\iota : \mathbb{U} \to \mathbb{V}$ to the whole space $\mathbb{V}$. In fact, this article also deals with extension of arbitrary operators: given a linear map $\mathcal{A}$ defined on $\mathbb{U}$ and taking value in a normed space $\mathbb{W}$, an extension $\mathcal{P}$ of $\mathcal{A}$ to $\mathbb{V}$ is a linear map from $\mathbb{V}$ to $\mathbb{W}$ such that

(i') $\mathcal{P}\mathbf{u} = \mathcal{A}\mathbf{u}$ for all $\mathbf{u} \in \mathbb{U}$,

(ii') $\mathcal{P}\mathbf{v} \in \mathrm{ran}(\mathcal{A})$ for all $\mathbf{v} \in \mathbb{V}$.

Our goal is to determine minimal projections/extensions, i.e., projections/extensions that minimize the operator norm

$$|||\mathcal{P}||| = \max_{\mathbf{v} \in \mathbb{V} \setminus \{\mathbf{0}\}} \frac{\|\mathcal{P}\mathbf{v}\|}{\|\mathbf{v}\|}.$$

Since the operator norm is a convex function, minimizing $|||\mathcal{P}|||$ is just a finite-dimensional convex optimization program and can *in principle* be solved efficiently. In the case of projections, the problem takes the form

$$\underset{\mathcal{P}}{\text{minimize}}\ |||\mathcal{P}||| \qquad \text{subject to} \qquad \mathcal{P} \text{ is a projection from } \mathbb{V} \text{ onto } \mathbb{U}.$$

By introducing a slack variable d, we arrive at the following equivalent problem:

$$\text{(P)} \qquad \underset{d,\mathcal{P}}{\text{minimize}}\ d \qquad \text{subject to} \qquad \mathcal{P} \text{ is a projection from } \mathbb{V} \text{ onto } \mathbb{U} \qquad \text{and to} \qquad |||\mathcal{P}||| \le d.$$

This is the formulation used in our practical implementations. In case of extension of $\mathcal{A} : \mathbb{U} \to \mathbb{W}$, an analogous formulation reads

$$\text{(E)} \qquad \underset{d,\mathcal{P}}{\text{minimize}}\ d \qquad \text{subject to} \qquad \mathcal{P} \text{ is an extension of } \mathcal{A} \qquad \text{and to} \qquad |||\mathcal{P}||| \le d.$$

We point out that both (P) and (E) are instances of the generic robust optimization problem [2, 3] with uncertainty sets involving the unit ball $B_{\mathbb{V}}$ of the space $\mathbb{V}$, since the constraint $|||\mathcal{P}||| \le d$ can be rephrased as $\|\mathcal{P}\mathbf{v}\| \le d\|\mathbf{v}\|$ for all $\mathbf{v} \in B_{\mathbb{V}}$.

[1] alternatively, one may consider $\mathcal{P}$ as a linear map from $\mathbb{V}$ into $\mathbb{U}$ and remove this second condition, but we found the current convention to have technical advantages in subsequent computations.

2.2 Expressing the constraints

In programs (P) and (E), the condition that $\mathcal{P}$ is a projection/extension turns into a linear equality constraint independent of the norms involved in the problem. We transform it for computational purposes below. Next, we focus on the condition $\|\mathcal{P}\| \le d$, which is a convex inequality constraint dependent on the norms.

2.2.1 Linear equality constraints

Let us consider bases for the spaces $\mathbb{U}, \mathbb{V}, \mathbb{W}$, say:

$$\begin{aligned} \underline{\mathbf{u}} &= (\mathbf{u}_1, \ldots, \mathbf{u}_m) \text{ is a basis for } \mathbb{U}, \\ \underline{\mathbf{v}} &= (\mathbf{v}_1, \ldots, \mathbf{v}_n) \ \text{ is a basis for } \mathbb{V}, \\ \underline{\mathbf{w}} &= (\mathbf{w}_1, \ldots, \mathbf{w}_\ell) \text{ is a basis for } \mathbb{W}. \end{aligned}$$

We represent the basis $\underline{\mathbf{u}}$ by its matrix U relative to the basis $\underline{\mathbf{v}}$, so that $\mathbf{u}_j = \sum_{i=1}^n U_{i,j} \mathbf{v}_j$, or pictorially

$$U = \begin{array}{c} \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{array} \begin{array}{c} \begin{array}{cccc} \mathbf{u}_1 & \mathbf{u}_2 & & \mathbf{u}_m \end{array} \\ \left[\begin{array}{c|c|c|c} \vdots & \vdots & & \vdots \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_m \\ \vdots & \vdots & & \vdots \end{array} \right] \end{array}.$$

We also represent the linear map $\mathcal{P} : \mathbb{V} \to \mathbb{W}$ as its matrix relative to the bases $\underline{\mathbf{v}}$ and $\underline{\mathbf{w}}$, so that $\mathcal{P}\mathbf{v}_j = \sum_{i=1}^\ell P_{i,j} \mathbf{w}_i$, or pictorially

$$P = \begin{array}{c} \\ \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_\ell \end{array} \begin{array}{c} \begin{array}{cccc} \mathbf{v}_1 & \mathbf{v}_2 & & \mathbf{v}_n \end{array} \\ \left[\begin{array}{c|c|c|c} \vdots & \vdots & & \vdots \\ \mathcal{P}(\mathbf{v}_1) & \mathcal{P}(\mathbf{v}_2) & \cdots & \mathcal{P}(\mathbf{v}_n) \\ \vdots & \vdots & & \vdots \end{array} \right] \end{array}.$$

In a similar way, we represent the linear map $\mathcal{A} : \mathbb{U} \to \mathbb{W}$ as its matrix relative to the bases $\underline{\mathbf{u}}$ and $\underline{\mathbf{w}}$, so that $\mathcal{A}\mathbf{u}_j = \sum_{i=1}^\ell A_{i,j} \mathbf{w}_j$, or pictorially

$$A = \begin{array}{c} \\ \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_\ell \end{array} \begin{array}{c} \begin{array}{cccc} \mathbf{u}_1 & \mathbf{u}_2 & & \mathbf{u}_m \end{array} \\ \left[\begin{array}{c|c|c|c} \vdots & \vdots & & \vdots \\ \mathcal{A}(\mathbf{u}_1) & \mathcal{A}(\mathbf{u}_2) & \cdots & \mathcal{A}(\mathbf{u}_m) \\ \vdots & \vdots & & \vdots \end{array} \right] \end{array}.$$

Such pieces of notation are helpful when expressing the conditions (i)-(ii), or more generally (i')-(ii'), as detailed below.

(i'): The condition $\mathcal{P}\mathbf{u} = \mathcal{A}\mathbf{u}$ for all $\mathbf{u} \in \mathbb{U}$ is equivalent to $\mathcal{P}\mathbf{u}_j = \mathcal{A}\mathbf{u}_j$ for all $j \in [\![1 : m]\!]$, i.e., $P\mathrm{col}_j(U) = \mathrm{col}_j(A)$ for all $j \in [\![1 : m]\!]$, or

$$PU = A. \tag{1}$$

Since optimization variables are typically organized as column vectors rather than matrices, we reformulate (1) as a constraint on the vectorization $\mathrm{vec}(P) = [\mathrm{col}_1(P); \ldots ; \mathrm{col}_n(P)]$ of P, i.e., on the vector formed by the columns of P stacked on top of each other. With $\otimes$ denoting the Kroenecker product, we invoke the property $\mathrm{vec}(MXN) = (N^\top \otimes M)\mathrm{vec}(X)$ to express $\mathrm{vec}(PU)$ as $(U^\top \otimes I_\ell)\mathrm{vec}(P)$, so that (1) becomes

$$(U^\top \otimes I_\ell)\mathrm{vec}(P) = \mathrm{vec}(A).$$

We stress once again that this is a linear condition on $\mathrm{vec}(P)$. In the case of projections, we have $\mathbb{W} = \mathbb{V}$, $A = \iota : \mathbb{U} \to \mathbb{V}$, and we can take $\underline{\mathbf{w}} = \underline{\mathbf{v}}$, so that the matrix A reduces to the matrix U, hence the constraint simplifies to

$$(U^\top \otimes I_n)\mathrm{vec}(P) = \mathrm{vec}(U). \tag{2}$$

(ii'): With the above constraint (i') being satisfied, the condition $\mathcal{P}\mathbf{v} \in \mathrm{ran}(\mathcal{A})$ is equivalent to $\mathcal{P}\widetilde{\mathbf{u}}_j \in \mathrm{ran}(\mathcal{A})$ for all $j \in [\![m+1 : n]\!]$, where $\underline{\widetilde{\mathbf{u}}} = (\widetilde{\mathbf{u}}_{m+1}, \ldots, \widetilde{\mathbf{u}}_n)$ is a basis of an orthogonal complement of $\mathbb{U}$ in $\mathbb{V}$. We represent $\underline{\widetilde{\mathbf{u}}}$ via its matrix $\widetilde{U}$ relative to $\underline{\mathbf{v}}$, that is

$$\widetilde{U} = \begin{array}{c} \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{array} \begin{array}{c} \begin{array}{cccc} \widetilde{\mathbf{u}}_{m+1} & \widetilde{\mathbf{u}}_{m+2} & \quad & \widetilde{\mathbf{u}}_n \end{array} \\ \left[\begin{array}{c|c|c|c} \vdots & \vdots & & \vdots \\ \widetilde{\mathbf{u}}_{m+1} & \widetilde{\mathbf{u}}_{m+2} & \cdots & \widetilde{\mathbf{u}}_n \\ \vdots & \vdots & & \vdots \end{array}\right] \end{array}.$$

Likewise, let $\widetilde{W}$ represent the matrix of a basis $\underline{\widetilde{\mathbf{w}}} = (\widetilde{\mathbf{w}}_{r+1}, \ldots, \widetilde{\mathbf{w}}_\ell)$ of an orthogonal complement of $\mathrm{ran}(\mathcal{A})$ in $\mathbb{W}$ — here the inner product $\langle \cdot, \cdot \rangle$ is the one making $\underline{\mathbf{w}}$ an orthonormal basis of $\mathbb{W}$ — that is

$$\widetilde{W} = \begin{array}{c} \\ \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_\ell \end{array} \begin{array}{c} \begin{array}{cccc} \widetilde{\mathbf{w}}_{r+1} & \widetilde{\mathbf{w}}_{r+2} & \quad & \widetilde{\mathbf{w}}_\ell \end{array} \\ \left[\begin{array}{c|c|c|c} \vdots & \vdots & & \vdots \\ \widetilde{\mathbf{w}}_{r+1} & \widetilde{\mathbf{w}}_{r+2} & \cdots & \widetilde{\mathbf{w}}_\ell \\ \vdots & \vdots & & \vdots \end{array}\right] \end{array}.$$

The condition $\mathcal{P}\widetilde{\mathbf{u}}_j \in \mathrm{ran}(\mathcal{A})$ for all $j \in [\![m+1 : n]\!]$ becomes $\langle \mathcal{P}\widetilde{\mathbf{u}}_j, \widetilde{\mathbf{w}}_i \rangle = 0$ for all $i \in [\![r+1 : \ell]\!]$ and $j \in [\![m+1 : n]\!]$. This reads $\mathrm{col}_i(\widetilde{W})^\top P\mathrm{col}_j(\widetilde{U}) = 0$ for all $i \in [\![r+1 : \ell]\!]$ and $j \in [\![m+1 : n]\!]$, i.e., $\widetilde{W}^\top P \widetilde{U} = 0$. After vectorization, we obtain

$$(\widetilde{U}^\top \otimes \widetilde{W}^\top)\mathrm{vec}(P) = 0.$$

We stress once again that this is a linear condition on $\mathrm{vec}(P)$. In case of projections, we have $\mathbb{V} = \mathbb{W}$, $\mathrm{ran}(\mathcal{A}) = \mathbb{U}$, and we can take $\underline{\widetilde{\mathbf{u}}} = \underline{\widetilde{\mathbf{w}}}$, so that the matrix $\widetilde{W}$ reduces to the matrix $\widetilde{U}$, hence the constraint simplifies to

$$(\widetilde{U}^\top \otimes \widetilde{U}^\top)\mathrm{vec}(P) = 0. \tag{3}$$

Remark. A convenient choice for the matrix $\widetilde{U}$ exploits the QR-factorization to U. Indeed, if

$$U = Q \left[\begin{array}{c} R \\ \hline 0 \end{array}\right],$$

then we can take

$$\widetilde{U} = Q \left[\begin{array}{c} 0 \\ \hline I \end{array}\right],$$

since the columns of $\widetilde{U}$ are linearly independent and since they are orthogonal to the columns of U, which can be seen from

$$\widetilde{U}^\top U = \left[\begin{array}{c|c} 0 & I \end{array}\right] Q^\top Q \left[\begin{array}{c} R \\ \hline 0 \end{array}\right] = \left[\begin{array}{c|c} 0 & I \end{array}\right] \left[\begin{array}{c} R \\ \hline 0 \end{array}\right] = 0.$$

2.2.2 Convex inequality constraints

We now look at convenient ways to express the condition $\|\mathcal{P}\| \le d$ in terms of the optimization variables d and P. In the best scenario, the operator norm $\|\mathcal{P}\|$ is explicitly computable. For projections, this is the case when

$$\mathbb{V} = \mathbb{W} = \ell_\infty^n : \qquad \|\mathcal{P}\| = \max_{i \in [\![1:n]\!]} \sum_{j=1}^n |P_{i,j}|,$$

$$\mathbb{V} = \mathbb{W} = \ell_1^n : \qquad \|\mathcal{P}\| = \max_{j \in [\![1:n]\!]} \sum_{i=1}^n |P_{i,j}|.$$

This will be exploited in Section 3. Otherwise, with $B_{\mathbb{V}}$ denoting the unit ball of the normed space $\mathbb{V}$ and with $\mathrm{Ex}(B_{\mathbb{V}})$ denoting its set of extreme points, the fact that

$$\|\mathcal{P}\| = \max_{\mathbf{v} \in B_{\mathbb{V}}} \|\mathcal{P}\mathbf{v}\| = \max_{\mathbf{v} \in \mathrm{Ex}(B_{\mathbb{V}})} \|\mathcal{P}\mathbf{v}\|$$

leads to the reformulation of the constraint $\|\mathcal{P}\| \le d$ as

$$\|\mathcal{P}\mathbf{v}^*\| \le d \qquad \text{for all } \mathbf{v}^* \in \mathrm{Ex}(B_{\mathbb{V}}).$$

This proves useful when the set $\mathrm{Ex}(B_{\mathbb{V}})$ is finite, e.g. for $\mathbb{V} = \mathcal{M}^{k\times k}_{\infty\to\infty}$ or $\mathbb{V} = \mathcal{M}^{k\times k}_{1\to 1}$, since

$$\mathrm{Ex}(B_{\mathcal{M}^{k\times k}_{\infty\to\infty}}) = \{\varepsilon_1 \mathbf{E}_{1,j_1} + \varepsilon_2 \mathbf{E}_{2,j_2} + \cdots + \varepsilon_k \mathbf{E}_{k,j_k},\ \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k \in \{\pm 1\},\ j_1, j_2, \ldots, j_k \in [\![1:k]\!]\},$$

$$\mathrm{Ex}(B_{\mathcal{M}^{k\times k}_{1\to 1}}) = \{\varepsilon_1 \mathbf{E}_{j_1,1} + \varepsilon_2 \mathbf{E}_{j_2,2} + \cdots + \varepsilon_k \mathbf{E}_{j_k,k},\ \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k \in \{\pm 1\},\ j_1, j_2, \ldots, j_k \in [\![1:k]\!]\}.$$

In addition, the choice of the basis $\underline{\mathbf{w}}$ for $\mathbb{W}$ is critical. It may allow for an explicit determination of $\|\mathcal{P}\mathbf{v}^*\|$ in terms of P — this will be exploited in Section 4. Failing this, it may still allow for a semidefinite relaxation in terms of P of the condition $\|\mathcal{P}\mathbf{v}^*\| \leq d$ — this will be exploited in Section 6.

3 Minimal projections in coordinate spaces

In this section, we consider projections onto subspaces of the spaces ℓ_1^n and ℓ_∞^n, i.e., of the coordinate space $\mathbb{R}^n$ equipped with the norms

$$\|\mathbf{x}\|_1 := \sum_{j=1}^n |x_j|,$$
$$\|\mathbf{x}\|_\infty := \max_{j \in [\![1:n]\!]} |x_j|.$$

3.1 Formulation of the minimization problem

According to the previous considerations, finding a minimal projection from ℓ_∞^n into $\mathbb{U} \subseteq \ell_\infty^n$ amounts to solving the optimization problem

$$\underset{d,P}{\text{minimize}}\ d \qquad \text{subject to} \qquad (U^\top \otimes I_n)\text{vec}(P) = \text{vec}(U), \quad (\widetilde{U}^\top \otimes \widetilde{U}^\top)\text{vec}(P) = 0,$$
$$\text{and to} \qquad \sum_{j=1}^n |P_{i,j}| \leq d \quad \text{for all } i \in [\![1:n]\!]$$

Likewise, for a minimal projection from ℓ_1^n into $\mathbb{U} \subseteq \ell_1^n$, we solve

$$\underset{d,P}{\text{minimize}}\ d \qquad \text{subject to} \qquad (U^\top \otimes I_n)\text{vec}(P) = \text{vec}(U), \quad (\widetilde{U}^\top \otimes \widetilde{U}^\top)\text{vec}(P) = 0,$$
$$\text{and to} \qquad \sum_{i=1}^n |P_{i,j}| \leq d \quad \text{for all } j \in [\![1:n]\!].$$

These optimization programs can be fed directly in the above form to CVX, a package for specifying and solving convex programs [1]. In our implementation, they are called via the commands

```
>>MinProjCoor(U,inf)
>>MinProjCoor(U,1)
```

In fact, the above convex programs have been transformed into linear programs whose sizes are doubled. This can be done either by replacing each $P_{i,j}$ by $P_{i,j}^+ \geq 0$ and $P_{i,j}^- \geq 0$ such that $P_{i,j} = P_{i,j}^+ - P_{i,j}^-$ and $|P_{i,j}| = P_{i,j}^+ + P_{i,j}^-$, or by introducing slack variables $c_{i,j}$ with $c_{i,j} \geq |P_{i,j}|$, i.e., $-c_{i,j} \leq P_{i,j} \leq c_{i,j}$. MinProj incorporates the latter option.

3.2 Exact formulas: verification

The article [5] established exact formulas for the projection constant of the hyperplane $\mathbb{U}_{\mathbf{a}}$ defined for $\mathbf{a} \in \mathbb{R}^n$ (assumed without loss of generality to satisfy $\|\mathbf{a}\|_1 = 1$) by

$$\mathbb{U}_{\mathbf{a}} = \left\{ \mathbf{u} \in \mathbb{R}^n : \langle \mathbf{a}, \mathbf{u} \rangle := \sum_{i=1}^{n} a_i u_i = 0 \right\}.$$

Precisely, with $\alpha_i = |a_i|/(1-2|a_i|)$ (interpreted as ∞ if $|a_i| \geq 1/2$), the projection constant of $\mathbb{U}_{\mathbf{a}}$ relative to ℓ_∞^n is (see [5, Theorem 2])

$$\lambda(\mathbb{U}_{\mathbf{a}}, \ell_\infty^n) = 1 + \frac{1}{\sum_{i=1}^{n} \alpha_i} = \begin{cases} 1 & \text{if } \|\mathbf{a}\|_\infty \geq 1/2, \\ 1 + \dfrac{1}{\sum_{i=1}^{n} \dfrac{|a_i|}{1-2|a_i|}} & \text{if } \|\mathbf{a}\|_\infty < 1/2. \end{cases}$$

A more complicated expression for the projection constant of $\mathbb{U}_{\mathbf{a}}$ relative to ℓ_1^n was also given in [5, Theorem 7]. Furthermore, in the case of two-codimensional subspaces, Lewicki obtained the expression of the projection constant relative to ℓ_∞^n for $n = 4$ in some special instances, see [14, Theorem 3.1] where the parameters α_i appear again. The consistency between these theoretical formulas and our computational outcomes is verified in the reproducible file for this article. It is furthermore expected that using the software MinProj will help to infer explicit expressions of projection constants in other specific situations.

3.3 Expected projection constants

We now use the software MinProj to explore the expected value of the projection constant of a random m-dimensional subspace $\mathbb{U}^m$ of ℓ_∞^n. The subspaces are generated as the span of m vectors in $\mathbb{R}^n$ with independent standard normal entries. Their computed projection constants are averaged to give an estimate of

$$\overline{\lambda}(m,n) := \mathbb{E}[\lambda(\mathbb{U}^m, \ell_\infty^n)].$$

The results are summarized on Figure 1. From the leftmost plot, one can hypothesize the behavior $\overline{\lambda}(m,n) \approx m^{\gamma(n)}$ with $0 < \gamma(n) \leq 1/2$, which is consistent with the known fact that the projection constant of an m-dimensional space is always bounded above by $\sqrt{m}$. On the rightmost plot, we examined the coefficient $\gamma(n)$ by representing $\ln[1/2 - \ln(\overline{\lambda}(m,n))/\ln(m)]$ as a function of $\ln(n)$ to possibly detect a behavior $\gamma(n) \approx 1/2 - \varepsilon(n)$ for a function $\varepsilon(n)$ decreasing as a power of n, but the outcome is not strikingly conclusive. In summary, the behavior $\overline{\lambda}(m,n) \approx m^{\gamma(n)}$ is postulated for the expected projection constant of a random m-dimensional subspace of ℓ_∞^n, without additional insight on the coefficient $\gamma(n)$ besides $\gamma(n) \leq 1/2$.

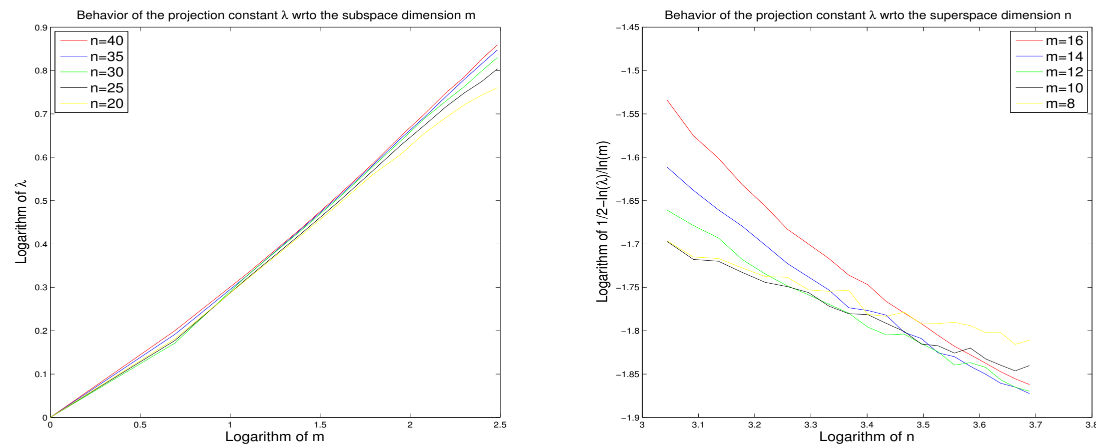

Figure 1: Behavior of the average projection constant as a function of m and n when m varies (left) and when n varies (right)

4 Minimal projections in matrix spaces

In this section, we consider projections onto subspaces of the spaces $\mathcal{M}^{k\times k}_{\infty\to\infty}$, $\mathcal{M}^{k\times k}_{1\to 1}$, and $\mathcal{M}^{k\times k}_{S_p}$, i.e., of the matrix space $\mathbb{R}^{k\times k}$ equipped with the norms

$$\|\mathbf{M}\|_{\infty\to\infty} := \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{M}\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} = \max_{i\in[\![1:k]\!]} \sum_{j=1}^{k} |M_{i,j}|,$$

$$\|\mathbf{M}\|_{1\to 1} := \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{M}\mathbf{x}\|_1}{\|\mathbf{x}\|_1} = \max_{j\in[\![1:k]\!]} \sum_{i=1}^{k} |M_{i,j}|,$$

$$\|\mathbf{M}\|_{S_p} := \left[\sum_{\ell=1}^{k} \sigma_\ell(\mathbf{M})^p\right]^{1/p}, \qquad \text{where } \sigma_1(\mathbf{M}) \geq \ldots \geq \sigma_k(\mathbf{M}) \text{ are the singular values of } \mathbf{M}.$$

The results are experimental in Subsection 4.1 dealing with the spaces $\mathcal{M}^{k\times k}_{\infty\to\infty}$ and $\mathcal{M}^{k\times k}_{1\to 1}$ that have polyhedral norms. Concerning the spaces $\mathcal{M}^{k\times k}_{S_p}$, $p \in [1,\infty]$, that have unitarily invariant norms, Subsection 4.2 presents some theoretical results for the subspaces of symmetric matrices and of zero-trace matrices.

4.1 Polyhedral norms

We reformulate the minimization program following the considerations of Section 2. Here, the basis $\underline{\mathbf{v}}$ of the k^2-dimensional space $\mathbb{R}^{k\times k}$ is chosen to be $(\mathbf{E}_{1,1}, \mathbf{E}_{1,2}, \ldots, \mathbf{E}_{1,k}, \mathbf{E}_{2,1}, \mathbf{E}_{2,2}, \ldots, \mathbf{E}_{2,k}, \ldots)$ in this order, that is

$$v_{(i-1)k+j} = \mathbf{E}_{i,j}, \qquad i,j \in [\![1:k]\!].$$

The subspace $\mathbb{U}$ is still represented by a basis $\underline{\mathbf{u}} = (\mathbf{u}_1, \ldots, \mathbf{u}_m)$ with matrix U relative to $\underline{\mathbf{v}}$ — a step that is not always straightforward. The linear map $\mathcal{P} : \mathbb{V} \to \mathbb{V}$ is also represented as before by the matrix P (or by its vectorized form). Then the linear constraint that $\mathcal{P}$ is a projection onto $\mathbb{U}$ is still expressed as (2)-(3).

Let us now focus on the convex constraint $\|\mathcal{P}\| \le d$ in the case $\mathbb{V} = \mathcal{M}^{k\times k}_{\infty\to\infty}$. This constraint reads

$$\|\mathcal{P}(\mathbf{V})\|_{\infty\to\infty} \le d \qquad \text{for all } \mathbf{V} \in \mathrm{Ex}(B_{\mathcal{M}^{k\times k}_{\infty\to\infty}}). \tag{4}$$

It is not hard to see that this set of extreme points is given by

$$\mathrm{Ex}(B_{\mathcal{M}^{k\times k}_{\infty\to\infty}}) = \{\varepsilon_1 \mathbf{E}_{1,j_1} + \varepsilon_2 \mathbf{E}_{2,j_2} + \cdots + \varepsilon_k \mathbf{E}_{k,j_k},\ \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k \in \{\pm 1\},\ j_1, j_2, \ldots, j_k \in [\![1:k]\!]\}.$$

Since it has cardinality $(2k)^k$ and since each constraint in (4) represents in fact the k constraints

$$\sum_{j=1}^k |[\mathcal{P}(\mathbf{V})]_{i,j}| \le d, \qquad \text{all } i \in [\![1:k]\!],$$

this results in a number of constraints equal to $2^k k^{k+1}$, which is much too large even for moderate values of k. We shall reduce the problem to arrive at a manageable number of constraints. First, for $i,j \in [\![1:k]\!]$ and $\mathbf{V} = \varepsilon_1 \mathbf{E}_{1,j_1} + \varepsilon_2 \mathbf{E}_{2,j_2} + \cdots + \varepsilon_k \mathbf{E}_{k,j_k} \in \mathrm{Ex}(B_{\mathcal{M}^{k\times k}_{\infty\to\infty}})$, notice that

$$\begin{aligned}
[\mathcal{P}(\mathbf{V})]_{i,j} &= \varepsilon_1 [\mathcal{P}(\mathbf{E}_{1,j_1})]_{i,j} + \varepsilon_2 [\mathcal{P}(\mathbf{E}_{2,j_2})]_{i,j} + \cdots + \varepsilon_k [\mathcal{P}(\mathbf{E}_{k,j_k})]_{i,j} \\
&= \varepsilon_1 P_{(i-1)k+j,j_1} + \varepsilon_2 P_{(i-1)k+j,k+j_2} + \cdots + \varepsilon_k P_{(i-1)k+j,(k-1)k+j_k}.
\end{aligned}$$

Thus, (4) becomes the constraint that, for all $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k \in \{\pm 1\}$, all $j_1, j_2, \ldots, j_k \in [\![1:k]\!]$, and all $i \in [\![1:k]\!]$,

$$\sum_{j=1}^k \left| \varepsilon_1 P_{(i-1)k+j,j_1} + \varepsilon_2 P_{(i-1)k+j,k+j_2} + \cdots + \varepsilon_k P_{(i-1)k+j,(k-1)k+j_k} \right| \le d.$$

This implies that, for all $j_1, j_2, \ldots, j_k \in [\![1:k]\!]$ and all $i \in [\![1:k]\!]$,

$$\sum_{j=1}^k \left(|P_{(i-1)k+j,j_1}| + |P_{(i-1)k+j,k+j_2}| + \cdots + |P_{(i-1)k+j,(k-1)k+j_k}| \right) \le d.$$

This is seemingly still k^{k+1} constraints, but if we write them as, for all $i \in [\![1:k]\!]$,

$$\max_{j_1 \in [\![1:k]\!]} \sum_{j=1}^k |P_{(i-1)k+j,j_1}| + \max_{j_2 \in [\![1:k]\!]} \sum_{j=1}^k |P_{(i-1)k+j,k+j_2}| + \cdots + \max_{j_k \in [\![1:k]\!]} \sum_{j=1}^k |P_{(i-1)k+j,(k-1)k+j_k}| \le d$$

and introduce slack variables $D_{i,1}, D_{i,2}, \ldots, D_{i,k}$ and impose, for all $i \in [\![1:k]\!]$,

$$D_{i,1} + D_{i,2} + \cdots + D_{i,k} \le d,$$

as well as, for all $i \in [\![1:k]\!]$ and all $j_1, j_2, \ldots, j_k \in [\![1:k]\!]$,

$$\sum_{j=1}^{k} |P_{(i-1)k+j,j_1}| \le D_{i,1}, \quad \sum_{j=1}^{k} |P_{(i-1)k+j,k+j_2}| \le D_{i,2}, \quad \ldots, \quad \sum_{j=1}^{k} |P_{(i-1)k+j,(k-1)k+j_k}| \le D_{i,k},$$

the number of constraints has been reduced to $k + k^3$ by adding k^2 variables. The complexity now appears polynomial in k rather than exponential in k. To sum up, the minimization program reads

$$\begin{aligned} \underset{d\in\mathbb{R},P\in\mathbb{R}^{k^2\times k^2},D\in\mathbb{R}^{k\times k}}{\text{minimize}} \; d \quad & \text{subject to} \quad (U^\top \otimes I_{k^2})\mathrm{vec}(P) = \mathrm{vec}(U), \quad (\widetilde{U}^\top \otimes \widetilde{U}^\top)\mathrm{vec}(P) = 0, \\ & \text{and to} \quad D_{i,1} + D_{i,2} + \cdots + D_{i,k} \le d \quad \text{for all } i \in [\![1:k]\!], \\ & \text{and to} \quad \sum_{j=1}^{k} |P_{(i-1)k+j,(h-1)k+\ell}| \le D_{i,h} \quad \text{for all } i,h,\ell \in [\![1:k]\!]. \end{aligned}$$

Likewise, for $\mathcal{M}^{k\times k}_{1\to 1}$, we would arrive at the following optimization program:

$$\begin{aligned} \underset{d\in\mathbb{R},P\in\mathbb{R}^{k^2\times k^2},D\in\mathbb{R}^{k\times k}}{\text{minimize}} \; d \quad & \text{subject to} \quad (U^\top \otimes I_{k^2})\mathrm{vec}(P) = \mathrm{vec}(U), \quad (\widetilde{U}^\top \otimes \widetilde{U}^\top)\mathrm{vec}(P) = 0, \\ & \text{and to} \quad D_{1,j} + D_{2,j} + \cdots + D_{k,j} \le d \quad \text{for all } j \in [\![1:k]\!], \\ & \text{and to} \quad \sum_{i=1}^{k} |P_{(i-1)k+j,(h-1)k+\ell}| \le D_{\ell,j} \quad \text{for all } j,h,\ell \in [\![1:k]\!]. \end{aligned}$$

These optimization programs can be fed directly in the above form to CVX. But in fact, as for coordinate spaces, we take a further step to transform them into linear programs by introducing slack variables $c_{s,t}$ with $c_{s,t} \ge |P_{s,t}|$ for all $s,t \in [\![1:k^2]\!]$. In the package MinProj, these programs are solved by calling the commands:

```
>>MinProjMatr(U,inf)
>>MinProjMatr(U,1)
```

These commands were used in a number of experiments designed to infer some particular projection constants. The results of these experiments, which can be duplicated by running the reproducible file for this article, are summarized in Table 1 for the space $\mathcal{M}^{k\times k}_{\infty\to\infty}$ (for the space $\mathcal{M}^{k\times k}_{1\to 1}$ instead of $\mathcal{M}^{k\times k}_{\infty\to\infty}$, the results are similar after exchanging the role of rows and columns). The notation used in the table is the following: $\mathbf{D}_\ell$ represents the matrix with 1's on the ℓth diagonal above the main diagonal and 0's everywhere else (so that $\langle \mathbf{M}, \mathbf{D}_\ell \rangle_F = \mathrm{tr}(\mathbf{D}_\ell^\top \mathbf{M})$ is the sum of the entries of $\mathbf{M}$ on the ℓth diagonal above the main diagonal), the matrix $\mathbf{R}_i$ represents the matrix with 1's in the ith row and 0's everywhere else, the matrix $\mathbf{C}_j$ represents the matrix with 1's in the jth column and 0's everywhere else, $\mathbf{X}$ represents the matrix with 1's on the main diagonal and main antidiagonal and 0's everywhere else, and $\widetilde{\mathbf{X}}$ represents the sum of the matrix with 1's on the main diagonal and 0's everywhere else and of the matrix with 1's on the main antidiagonal and 0's everywhere else (so that $\widetilde{\mathbf{X}} = \mathbf{X}$ if k is even and $\widetilde{\mathbf{X}} = \mathbf{X} + \mathbf{E}_{(k+1)/2,(k+1)/2}$ if k is odd). The orthogonal complement is understood with respect to the Frobenius inner product. Note that once the projection constant is determined, verifying that the projection given in the third column is indeed minimal only requires to bound its norm from above by the value from the second column. It is particularly instructive to perform this exercise by hand for the subspace $\{\mathbf{C}_1 + \cdots + \mathbf{C}_\ell\}^\perp$ to explain the difference between the cases $\ell = 1$ and $\ell \geq 2$.

Table 1: Empirically determined projection constants in $\mathcal{M}^{k\times k}_{\infty\to\infty}$, $k \geq 2$

Subspace	Projection Constant	Minimal Projection
Symmetric matrices	$\dfrac{1+k}{2}$	$\mathbf{M} \mapsto \dfrac{1}{2}\left(\mathbf{M} + \mathbf{M}^\top\right)$
Zero-trace matrices	$2 - \dfrac{1}{k}$	$\mathbf{M} \mapsto \mathbf{M} - \dfrac{1}{k}\mathrm{tr}(\mathbf{M})\mathbf{I}$
$\{\mathbf{D}_\ell\}^\perp$	$2 - \dfrac{1}{k-\ell}$	$\mathbf{M} \mapsto \mathbf{M} - \dfrac{1}{k-\ell}\mathrm{tr}(\mathbf{D}_\ell^\top \mathbf{M})\mathbf{D}_\ell$
$\{\mathbf{R}_1 + \cdots + \mathbf{R}_\ell\}^\perp$	$2 - \dfrac{2}{\ell k}$	$\mathbf{M} \mapsto \mathbf{M} - \dfrac{1}{\ell k}\left[\sum_{i=1}^{\ell}\sum_{j=1}^{k} M_{i,j}\right](\mathbf{R}_1 + \cdots + \mathbf{R}_\ell)$
$\{\mathbf{C}_1 + \cdots + \mathbf{C}_\ell\}^\perp$	$2 - \dfrac{2}{\max\{2,\ell\}k}$	$\mathbf{M} \mapsto \mathbf{M} - \dfrac{1}{\ell k}\left[\sum_{i=1}^{k}\sum_{j=1}^{\ell} M_{i,j}\right](\mathbf{C}_1 + \cdots + \mathbf{C}_\ell)$
$\{\mathbf{X}\}^\perp$	$2 - \dfrac{1}{k}$	$\mathbf{M} \mapsto \mathbf{M} - \dfrac{1}{2k}\mathrm{tr}(\mathbf{X}\mathbf{M})\widetilde{\mathbf{X}}$

Although it is conceivable that extensive experiments could allow for the empirical determination of a general formula for the projection constant of a hyperplane[2] in $\mathcal{M}^{k\times k}_{\infty\to\infty}$ or $\mathcal{M}^{k\times k}_{1\to 1}$, the case of the orthogonal complement of the matrix with 1's on the first row and first column and 0's everywhere else is already uncertain.

4.2 Unitarily invariant norms

In this subsection, we consider the space $\mathbb{R}^{k\times k}$ equipped with a unitarily invariant norm $\|\cdot\|$, i.e., one that satisfies

$$\|\mathbf{U}\mathbf{M}\mathbf{V}\| = \|\mathbf{M}\| \qquad \text{for all } \mathbf{M}, \mathbf{U}, \mathbf{V} \in \mathbb{R}^{k\times k} \text{ with } \mathbf{U}, \mathbf{V} \text{ orthogonal.}$$

For such a norm, the quantity $\|\mathbf{M}\|$ depends only on the singular values of $\mathbf{M}$ (as a consequence of [4, Theorem IV.2.1]). Besides the straightforward observation that $\mathbf{M} \mapsto (\mathbf{M} + \mathbf{M}^\top)/2$ is always

[2] the projection constant of a hyperplane is always bounded above by 2, as confirmed by Table 1.

a minimal projection onto the subspace of symmetric matrices in this setting (since it has norm one), we shall establish the following result.

Theorem 1. If the matrix space $\mathbb{R}^{k\times k}$ is equipped with a unitarily invariant norm, then the map

$$\mathbf{M} \mapsto \mathbf{M} - \frac{\mathrm{tr}(\mathbf{M})}{k}\mathbf{I}_k$$

is a minimal projection onto the subspace of zero-trace matrices.

We first isolate the following fact about hyperplanes of $\mathbb{R}^{k\times k}$, which are all described for some $\mathbf{B} \in \mathbb{R}^{k\times k}$ as

$$\mathbb{U}_{\mathbf{B}} := \left\{ \mathbf{M} \in \mathbb{R}^{k\times k} : \mathrm{tr}(\mathbf{BM}) = 0 \right\}.$$

Proposition 2. If the matrix $\mathbf{B} \in \mathbb{R}^{k\times k}$ is symmetric, then there is a minimal projection from $\mathbb{R}^{k\times k}$ onto $\mathbb{U}_{\mathbf{B}}$ of the form

$$\mathbf{M} \mapsto \mathbf{M} - \mathrm{tr}(\mathbf{BM})\mathbf{A} \qquad \text{for some symmetric } \mathbf{A} \in \mathbb{R}^{k\times k} \text{ with } \mathrm{tr}(\mathbf{BA}) = 1.$$

Proof. We first remark that $\mathbb{U}_{\mathbf{B}}$ is transposition-invariant, since

$$\mathrm{tr}(\mathbf{BM}^\top) = \mathrm{tr}(\mathbf{B}^\top\mathbf{M}^\top) = \mathrm{tr}((\mathbf{MB})^\top) = \mathrm{tr}(\mathbf{MB}) = \mathrm{tr}(\mathbf{BM}).$$

If $\mathcal{P}$ is a projection onto $\mathbb{U}_{\mathbf{B}}$, we define $\mathcal{P}' : \mathbb{R}^{k\times k} \to \mathbb{R}^{k\times k}$ by $\mathcal{P}'(\mathbf{M}) = [\mathcal{P}(\mathbf{M}^\top)]^\top$, $\mathbf{M} \in \mathbb{R}^{k\times k}$. Using the transposition-invariance of $\mathbb{U}_{\mathbf{B}}$, we notice that $\mathcal{P}'$ has values in $\mathbb{U}_{\mathbf{B}}$ and that it acts as the identity on $\mathbb{U}_{\mathbf{B}}$, since

$$\mathbf{M} \in \mathbb{U}_{\mathbf{B}} \Rightarrow \mathbf{M}^\top \in \mathbb{U}_{\mathbf{B}} \Rightarrow \mathcal{P}(\mathbf{M}^\top) = \mathbf{M}^\top \Rightarrow \mathcal{P}'(\mathbf{M}) = [\mathbf{M}^\top]^\top = \mathbf{M}.$$

In short, $\mathcal{P}'$ is a projection onto $\mathbb{U}_{\mathbf{B}}$. It is also plain to see that $\mathcal{P}'$ has the same norm as $\mathcal{P}$, since $|||\mathcal{P}'||| \le |||\mathcal{P}|||$ is apparent from

$$\|\mathcal{P}'(\mathbf{M})\| = \|[\mathcal{P}(\mathbf{M}^\top)]^\top\| = \|\mathcal{P}(\mathbf{M}^\top)\| \le |||\mathcal{P}|||\|\mathbf{M}^\top\| = |||\mathcal{P}|||\|\mathbf{M}\|, \quad \mathbf{M} \in \mathbb{R}^{k\times k},$$

and since the reverse inequality $|||\mathcal{P}||| \le |||\mathcal{P}'|||$ holds because $(\mathcal{P}')' = \mathcal{P}$. Thus, if $\mathcal{P}$ is a minimal projection, then so is $\mathcal{P}'$, and $\mathcal{P}'' := (\mathcal{P} + \mathcal{P}')/2$ is a minimal projection, too. Observing that any projection from $\mathbb{R}^{k\times k}$ takes the form

$$\mathcal{P}(\mathbf{M}) = \mathbf{M} - \mathrm{tr}(\mathbf{BM})\mathbf{A} \qquad \text{for some } \mathbf{A} \in \mathbb{R}^{k\times k} \text{ with } \mathrm{tr}(\mathbf{BA}) = 1,$$

if $\mathcal{P}$ written in this form is a minimal projection, then the minimal projection $\mathcal{P}''$ takes the form

$$\mathcal{P}''(\mathbf{M}) = \frac{1}{2}\left(\mathbf{M} - \mathrm{tr}(\mathbf{BM})\mathbf{A}\right) + \frac{1}{2}\left(\mathbf{M} - \mathrm{tr}(\mathbf{BM}^\top)\mathbf{A}^\top\right) = \mathbf{M} - \mathrm{tr}(\mathbf{BM})\frac{\mathbf{A}+\mathbf{A}^\top}{2}.$$

Our claim is finally justified by the immediate facts that $(\mathbf{A}+\mathbf{A}^\top)/2$ is a symmetric matrix and that $\mathrm{tr}(\mathbf{B}(\mathbf{A}+\mathbf{A}^\top)/2) = 1$. $\square$

With Proposition 2 in place, we continue with the proof of the main result of this subsection.

Proof of Theorem 1. According to Proposition 2 applied to $\mathbf{B} = \mathbf{I}$, there is a projection $\mathcal{P}$ from $\mathbb{R}^{k\times k}$ onto the subspace $\mathbb{U}_\mathbf{I}$ of zero-trace matrices that takes the form

$$\mathbf{M} \mapsto \mathbf{M} - \mathrm{tr}(\mathbf{M})\mathbf{A} \qquad \text{for some symmetric } \mathbf{A} \in \mathbb{R}^{k\times k} \text{ with } \mathrm{tr}(\mathbf{A}) = 1.$$

Now, given an orthogonal matrix $\mathbf{V} \in \mathbb{R}^{k\times k}$, we define

$$\mathcal{P}_\mathbf{V}(\mathbf{M}) := \mathbf{V}^\top \mathcal{P}(\mathbf{V}\mathbf{M}\mathbf{V}^\top)\mathbf{V} = \mathbf{M} - \mathrm{tr}(\mathbf{M})\mathbf{V}^\top\mathbf{A}\mathbf{V}.$$

We easily notice that $\mathcal{P}_\mathbf{V}$ takes values in $\mathbb{U}_\mathbf{I}$ and that it acts as the identity on $\mathbb{U}_\mathbf{I}$, since

$$\mathbf{M} \in \mathbb{U}_\mathbf{I} \Rightarrow \mathbf{V}\mathbf{M}\mathbf{V}^\top \in \mathbb{U}_\mathbf{I} \Rightarrow \mathcal{P}(\mathbf{V}\mathbf{M}\mathbf{V}^\top) = \mathbf{V}\mathbf{M}\mathbf{V}^\top \Rightarrow \mathcal{P}_\mathbf{V}(\mathbf{M}) = \mathbf{V}^\top(\mathbf{V}\mathbf{M}\mathbf{V}^\top)\mathbf{V} = \mathbf{M}.$$

In short, $\mathcal{P}_\mathbf{V}$ is a projection onto $\mathbb{U}_\mathbf{I}$. It is in fact a minimal projection, since unitarily invariance yields $\|\!|\mathcal{P}_\mathbf{V}|\!\| \le \|\!|\mathcal{P}|\!\|$ in view of

$$\|\mathcal{P}_\mathbf{V}(\mathbf{M})\| = \|\mathbf{V}^\top\mathcal{P}(\mathbf{V}\mathbf{M}\mathbf{V}^\top)\mathbf{V}\| = \|\mathcal{P}(\mathbf{V}\mathbf{M}\mathbf{V}^\top)\| \le \|\!|\mathcal{P}|\!\|\|\mathbf{V}\mathbf{M}\mathbf{V}^\top\| \le \|\!|\mathcal{P}|\!\|\|\mathbf{M}\|, \quad \mathbf{M} \in \mathbb{R}^{k\times k}.$$

Then, with μ denoting the Haar measure on the group $\mathcal{O}(k)$ of orthogonal matrices in $\mathbb{R}^{k\times k}$, one readily sees that

$$\widehat{\mathcal{P}}(\mathbf{M}) := \int_{\mathcal{O}(k)} \mathcal{P}_\mathbf{V}(\mathbf{M}) d\mu(\mathbf{V}) = \mathbf{M} - \mathrm{tr}(\mathbf{M})\widehat{\mathbf{A}}, \qquad \widehat{\mathbf{A}} := \int_{\mathcal{O}(k)} \mathbf{V}^\top\mathbf{A}\mathbf{V} d\mu(\mathbf{V}),$$

also defines a minimal projection onto $\mathbb{U}_\mathbf{I}$. The symmetric matrix $\widehat{\mathbf{A}}$ satisfies $\widehat{\mathbf{A}} = \mathbf{U}^\top\widehat{\mathbf{A}}\mathbf{U}$ for all orthogonal matrices $\mathbf{U} \in \mathbb{R}^{k\times k}$. This implies that $\widehat{\mathbf{A}}$ is a diagonal matrix (by taking $\mathbf{U}$ with columns composed of eigenvectors of $\mathbf{A}$), next that the diagonal entries of $\widehat{\mathbf{A}}$ are all equal (by taking $\mathbf{U}$ as a permutation matrix), and finally that $\widehat{\mathbf{A}} = \mathbf{I}_k/k$ (by virtue of $\mathrm{tr}(\widehat{\mathbf{A}}) = \mathrm{tr}(\mathbf{A}) = 1$). We have therefore proved that the minimal projection $\widehat{\mathcal{P}}$ takes the required form. $\square$

It is worth pointing out that the projection exhibited in Theorem 1 is just the orthogonal projection onto the subspace of zero-trace matrices with respect to the Frobenius norm. Its minimality was of course clear when $\mathbb{R}^{k\times k}$ is equipped with the Frobenius norm, i.e., the Schatten 2-norm, but it was a priori not clear when $\mathbb{R}^{k\times k}$ is equipped with the spectral norm (aka operator norm), i.e., the Schatten ∞-norm, or with the nuclear norm (aka trace norm), i.e., the Schatten 1-norm. We may now deduce the value of the projection constants for these norms.

Theorem 3. The projection constants of the subspace $\mathbb{U}_\mathbf{I}$ of zero-trace matrices in $\mathbb{R}^{k\times k}$ equipped with the spectral norm and with the nuclear norm are given by

$$\lambda\left(\mathbb{U}_\mathbf{I}, \mathcal{M}^{k\times k}_{S_\infty}\right) = 2\frac{k-1}{k} \qquad \text{and} \qquad \lambda\left(\mathbb{U}_\mathbf{I}, \mathcal{M}^{k\times k}_{S_1}\right) = 2\frac{k-1}{k}.$$

Moreover, for any $p \in [1,\infty]$, the projection constant of this subspace in $\mathcal{M}^{k\times k}_{S_p}$ satisfies

$$\lambda\left(\mathbb{U}_{\mathbf{I}}, \mathcal{M}^{k\times k}_{S_p}\right) \le \left[2\frac{k-1}{k}\right]^{|p-2|/p}$$

Proof. For the first part, Theorem 1 guarantees that it is enough to prove that, for all $\mathbf{M} \in \mathbb{R}^{k\times k}$,

$$\|\mathcal{P}(\mathbf{M})\|_{S_\infty} \le 2\frac{k-1}{k}\|\mathbf{M}\|_{S_\infty} \qquad \text{and} \qquad \|\mathcal{P}(\mathbf{M})\|_{S_1} \le 2\frac{k-1}{k}\|\mathbf{M}\|_{S_1},$$

where $\mathcal{P}(\mathbf{M}) = \mathbf{M} - (\mathrm{tr}(\mathbf{M})/k)\mathbf{I}_k$, and that these inequalities turn to equalities for some $\mathbf{M} \in \mathbb{R}^{k\times k}$. Let then $\mathbf{M} \in \mathbb{R}^{k\times k}$, or more generally $\mathbf{M} \in \mathbb{C}^{k\times k}$. By Schur's unitary triangularization theorem, there is an upper triangular matrix $\mathbf{T} \in \mathbb{C}^{k\times k}$ and a unitary matrix $\mathbf{U} \in \mathbb{C}^{k\times k}$ such that

$$\mathbf{M} = \mathbf{U}\mathbf{T}\mathbf{U}^*, \qquad \text{hence} \quad \mathcal{P}(\mathbf{M}) = \mathbf{U}[\mathbf{T} - (\mathrm{tr}(\mathbf{T})/k)\mathbf{I}_k]\mathbf{U}^*.$$

With $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ denoting the diagonal entries of $\mathbf{T}$, the diagonal entries of $[\mathbf{T} - (\mathrm{tr}(\mathbf{T})/k)\mathbf{I}_k]$ are $\lambda_1 - \left(\sum_{\ell=1}^k \lambda_\ell\right)/k, \ldots, \lambda_k - \left(\sum_{\ell=1}^k \lambda_\ell\right)/k$. Thus, we need to prove that

$$\max_{i\in[\![1:k]\!]} \left|\lambda_i - \frac{1}{k}\sum_{\ell=1}^k \lambda_\ell\right| \le 2\frac{k-1}{k} \max_{j\in[\![1:k]\!]} |\lambda_j|, \tag{5}$$

$$\sum_{i=1}^k \left|\lambda_i - \frac{1}{k}\sum_{\ell=1}^k \lambda_\ell\right| \le 2\frac{k-1}{k} \sum_{j=1}^k |\lambda_j|, \tag{6}$$

and that these inequalities are sharp for some $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$. To obtain (5), we write that, for all $i \in [\![1:k]\!]$,

$$\begin{aligned}\left|\lambda_i - \frac{1}{k}\sum_{\ell=1}^k \lambda_\ell\right| &= \left|\left(1-\frac{1}{k}\right)\lambda_i + \frac{1}{k}\sum_{\ell=1,\ell\ne i}^n \lambda_\ell\right| \le \left(1-\frac{1}{k}\right)|\lambda_i| + \frac{1}{k}\sum_{\ell=1,\ell\ne i}^n |\lambda_\ell| \\ &\le \left(1-\frac{1}{k}\right)\max_{j\in[\![1:k]\!]}|\lambda_j| + \frac{k-1}{k}\max_{j\in[\![1:k]\!]}|\lambda_j| = 2\frac{k-1}{k}\max_{j\in[\![1:k]\!]}|\lambda_j|,\end{aligned}$$

as expected. Note that equality occurs by taking $i=1$, $\lambda_1 = 1$, and $\lambda_2 = \cdots = \lambda_k = -1$. To obtain (6), we write

$$\begin{aligned}\sum_{i=1}^k \left|\lambda_i - \frac{1}{k}\sum_{\ell=1}^k \lambda_\ell\right| &= \sum_{i=1}^k \left|\left(1-\frac{1}{k}\right)\lambda_i + \frac{1}{k}\sum_{\ell=1,\ell\ne i} \lambda_\ell\right| \le \left(1-\frac{1}{k}\right)\sum_{i=1}^k|\lambda_i| + \frac{1}{k}\sum_{i,\ell=1,i\ne\ell}^k |\lambda_\ell| \\ &\le \left(1-\frac{1}{k}\right)\sum_{i=1}^k|\lambda_i| + \frac{1}{k}\sum_{\ell=1}^n (k-1)|\lambda_\ell| = 2\frac{k-1}{k}\sum_{j=1}^k|\lambda_j|,\end{aligned}$$

as expected. Note that equality occurs by taking $\lambda_1 = 1$ and $\lambda_2 = \cdots = \lambda_k = 0$.

For the second part, we simply apply Riesz–Thorin theorem to the operator

$$F : [\lambda_1; \ldots ; \lambda_k] \in \mathbb{C}^k \mapsto \left[\lambda_1 - \frac{1}{k}\sum_{\ell=1}^{k} \lambda_\ell; \ldots ; \lambda_k - \frac{1}{k}\sum_{\ell=1}^{k} \lambda_\ell\right] \in \mathbb{C}^k,$$

which has just been shown to satisfy $\|F\|_{\infty\to\infty} = \|F\|_{1\to 1} = 2(k-1)/k$ and which also satisfies $\|F\|_{2\to 2} = 1$ (because the orthogonal projection $\mathcal{P}$ has norm one relative to the Schatten 2-norm). It gives the required bound on $\|F\|_{p\to p}$, and in turn on $\|\mathcal{P}\|_{S_p\to S_p}$ by the above arguments. □

Theorem 3 has a direct implication on the value of the projection constant of the subspace of zero-trace matrices relative to Ky-Fan norms, which are defined for $s \in [\![1 : k]\!]$ by

$$\|\mathbf{M}\|_{(s)} = \sum_{i=1}^{s} \sigma_i(\mathbf{M}), \qquad \mathbf{M} \in \mathbb{R}^{k\times k}.$$

We shall rely instead on an alternative characterization of these norms (see [4, Proposition IV.2.3]), namely

$$\|\mathbf{M}\|_{(s)} = \min \left\{ \|\mathbf{M}'\|_{S_1} + s\|\mathbf{M}''\|_{S_\infty}, \mathbf{M} = \mathbf{M}' + \mathbf{M}'' \right\}.$$

Corollary 4. The projection constant of the subspace $\mathbb{U}_{\mathbf{I}}$ of zero-trace matrices in $\mathbb{R}^{k\times k}$ equipped with the Ky-Fan s-norm obeys

$$\lambda\left(\mathbb{U}_{\mathbf{I}}, \mathcal{M}_{(s)}^{k\times k}\right) \le 2\frac{k-1}{k}.$$

Proof. With $\mathcal{P}$ denoting the projection of Theorem 1, we are going to prove that, for all $\mathbf{M} \in \mathbb{R}^{k\times k}$,

$$\|\mathcal{P}(\mathbf{M})\|_{(s)} \le 2\frac{k-1}{k}\|\mathbf{M}\|_{(s)},$$

given that we already know $\|\mathcal{P}\|_{S_1\to S_1} = \|\mathcal{P}\|_{S_\infty\to S_\infty} = 2(k-1)/k$. For $\mathbf{M} \in \mathbb{R}^{k\times k}$, let us consider $\mathbf{M}', \mathbf{M}'' \in \mathbb{R}^{k\times k}$ with $\mathbf{M} = \mathbf{M}' + \mathbf{M}''$ and $\|\mathbf{M}\|_{(s)} = \|\mathbf{M}'\|_{S_1} + s\|\mathbf{M}''\|_{S_\infty}$. Then the equality $\mathcal{P}(\mathbf{M}) = \mathcal{P}(\mathbf{M}') + \mathcal{P}(\mathbf{M}'')$ yields

$$\begin{aligned}\|\mathcal{P}(\mathbf{M})\|_{(s)} &\le \|\mathcal{P}(\mathbf{M}')\|_{S_1} + s\|\mathcal{P}(\mathbf{M}'')\|_{S_\infty} \le \|\mathcal{P}\|_{S_1\to S_1}\|\mathbf{M}'\|_{S_1} + \|\mathcal{P}\|_{S_\infty\to S_\infty}\|\mathbf{M}''\|_{S_\infty}\\ &= 2\frac{k-1}{k}\left(\|\mathbf{M}'\|_{S_1} + s\|\mathbf{M}''\|_{S_\infty}\right) = 2\frac{k-1}{k}\|\mathbf{M}\|_{(s)},\end{aligned}$$

which is the desired result. □

5 Minimal projections in polynomial spaces

In this section, we consider projections onto subspaces of $\mathbb{P}^n_\infty$, i.e., of the n-dimensional space of polynomials of degree at most $n-1$ equipped with the max-norm

$$\|f\|_\infty := \max_{x\in[-1,1]} |f(x)|.$$

5.1 Approximation of the optimization program

Given a subspace $\mathbb{U}$ of $\mathbb{P}^n_\infty$, the original minimization program

$$\text{(7)} \qquad \underset{\mathcal{P}}{\text{minimize}} \; \|\mathcal{P}\|_{\infty\to\infty} \qquad \text{subject to} \qquad \mathcal{P} \text{ is a projection from } \mathbb{P}^n_\infty \text{ onto } \mathbb{U}$$

can be expressed in the equivalent form

$$\text{(8)} \qquad \underset{d,\mathcal{P}}{\text{minimize}} \; d \qquad \text{subject to} \qquad \mathcal{P} \text{ is a projection from } \mathbb{P}^n_\infty \text{ onto } \mathbb{U}$$
$$\text{and to} \qquad \|\mathcal{P}f\|_\infty \le d \quad \text{for all } f \in \mathbb{P}^n_\infty \text{ with } \|f\|_\infty \le 1.$$

As we do not see how to solve the program (8) exactly, we resort to computing the solutions of approximate programs whose closeness to the true solution can be quantified. For this purpose, the max-norm is substituted by the discrete norm

$$\|f\|_{[N]} := \max_{i \in [\![1:N]\!]} |f(t_i)|, \qquad t_i := \cos(\theta_i), \qquad \theta_i := \frac{(N-i)\pi}{N-1}.$$

On the one hand, it is clear that

$$\|f\|_{[N]} \le \|f\|_\infty \qquad \text{for all } f \in \mathbb{P}^n_\infty.$$

On the other hand, given $f \in \mathbb{P}^n_\infty$ and $x \in [-1,1]$ written as $x = \cos(\theta)$ for some $\theta \in [0,\pi]$, we choose $i \in [\![1:N]\!]$ such that $|\theta - \theta_i| \le \pi/(2(N-1))$, and we derive that

$$\begin{aligned} |f(x)| &= |(f\circ\cos)(\theta)| \le |(f\circ\cos)(\theta_i)| + |(f\circ\cos)(\theta) - (f\circ\cos)(\theta_i)| \\ &\le |f(t_i)| + \|(f\circ\cos)'\|_\infty |\theta-\theta_i| \le \|f\|_{[N]} + (n-1)\|f\circ\cos\|_\infty \frac{\pi}{2(N-1)}, \end{aligned}$$

where we have used Bernstein inequality for trigonometric polynomials of degree at most $n-1$ in the last step. Taking the maximum over x yields $\|f\|_\infty \le \|f\|_{[N]} + (\pi(n-1))/(2(N-1))\|f\|_\infty$. Choosing N such that $N-1$ is a multiple of $n-1$, thus introducing a discretization index ν such that

$$N = (n-1)\nu + 1,$$

a rearrangement gives

$$\|f\|_\infty \le \left(1 - \frac{\pi/2}{\nu}\right)^{-1} \|f\|_{[N]} \qquad \text{for all } f \in \mathbb{P}^n_\infty.$$

Thus, we obtain the estimates

$$\max_{f\in\mathbb{P}^n_\infty} \frac{\|\mathcal{P}f\|_\infty}{\|f\|_\infty} \begin{cases} \le & \left(1-\dfrac{\pi/2}{\nu}\right)^{-1} \times \displaystyle\max_{f\in\mathbb{P}^n_\infty} \frac{\|\mathcal{P}f\|_{[N]}}{\|f\|_{[N]}}, \\ \ge & \left(1-\dfrac{\pi/2}{\nu}\right) \times \displaystyle\max_{f\in\mathbb{P}^n_\infty} \frac{\|\mathcal{P}f\|_{[N]}}{\|f\|_{[N]}}, \end{cases}$$

$$\max_{f\in\mathbb{P}^n_\infty} \frac{\|\mathcal{P}f\|_\infty}{\|f\|_\infty} \begin{cases} \le & \left(1-\dfrac{\pi/2}{\nu}\right)^{-1} \times \displaystyle\max_{f\in\mathbb{P}^n_\infty} \frac{\|\mathcal{P}f\|_{[N]}}{\|f\|_\infty}, \\ \ge & 1 \times \displaystyle\max_{f\in\mathbb{P}^n_\infty} \frac{\|\mathcal{P}f\|_{[N]}}{\|f\|_\infty}. \end{cases}$$

It follows that the projection constant $\lambda_n(\mathbb{U}) := \lambda(\mathbb{U}, \mathbb{P}^n_\infty)$ of $\mathbb{U}$ in $\mathbb{P}^n_\infty$, i.e., the solution of (7), satisfies

$$\left(1 - \frac{\pi/2}{\nu}\right) \le \frac{\lambda_n(\mathbb{U})}{\lambda_n(\mathbb{U}; N, N)} \le \left(1 - \frac{\pi/2}{\nu}\right)^{-1} \qquad \text{and} \qquad 1 \le \frac{\lambda_n(\mathbb{U})}{\lambda_n(\mathbb{U}; N)} \le \left(1 - \frac{\pi/2}{\nu}\right)^{-1},$$

where $\lambda_n(\mathbb{U}; N, N)$ is the solution of the program

$$\text{(9)} \qquad \underset{d,\mathcal{P}}{\text{minimize}}\ d \qquad \text{subject to} \qquad \mathcal{P} \text{ is a projection from } \mathbb{P}^n_\infty \text{ onto } \mathbb{U}$$
$$\text{and to} \qquad \|\mathcal{P}f\|_{[N]} \le d \quad \text{for all } f \in \mathbb{P}^n_\infty \text{ with } \|f\|_{[N]} \le 1$$

and $\lambda_n(\mathbb{U}; N)$ is the solution of the program

$$\text{(10)} \qquad \underset{d,\mathcal{P}}{\text{minimize}}\ d \qquad \text{subject to} \qquad \mathcal{P} \text{ is a projection from } \mathbb{P}^n_\infty \text{ onto } \mathbb{U}$$
$$\text{and to} \qquad \|\mathcal{P}f\|_{[N]} \le d \quad \text{for all } f \in \mathbb{P}^n_\infty \text{ with } \|f\|_\infty \le 1.$$

We point out that both these optimization programs take the form

$$\underset{d,\mathcal{P}}{\text{minimize}}\ d \qquad \text{subject to} \qquad \mathcal{P} \text{ is a projection from } \mathbb{P}^n_\infty \text{ onto } \mathbb{U}$$
$$\text{and to} \qquad |(\mathcal{P}f)(t_i)| \le d \quad \text{for all } f \in \Omega \text{ and all } i \in [\![1:N]\!]$$

with different sets Ω for (9) and for (10). In the next two subsections, we show how familiar robust optimization techniques (see e.g. [2, Theorem 1.3.4]) are used to transform (9) and (10) into a linear program and into a semidefinte program, respectively. The details are spelled out for the reader's convenience.

5.2 The first approximate program as a linear programming

We first look at a convenient parametrization of the set Ω corresponding to the program (9), which is initially given by

$$\Omega := \left\{ f \in \mathbb{P}^n_\infty : \|f\|_{[N]} = \max_{i \in [\![1:N]\!]} |f(t_i)| \le 1 \right\}.$$

Let $(\widetilde{L}_1, \ldots, \widetilde{L}_N)$ denote the basis of $\mathbb{P}^N_\infty$ of fundamental Lagrange interpolating polynomials at the Chebyshev points $t_1, \ldots, t_N$ of order N, i.e.,

$$\widetilde{L}_i \in \mathbb{P}^N_\infty \quad \text{satisfies} \quad \widetilde{L}_i(t_j) = \delta_{i,j}.$$

Similarly, let $(L_1, \ldots, L_n)$ denote the basis of $\mathbb{P}^n_\infty$ of fundamental Lagrange interpolating polynomials at the Chebyshev points of order n, which happen to be $t_1, t_{\nu+1}, \ldots, t_{(n-2)\nu+1}, t_{(n-1)\nu+1} = t_N$. Any $f \in \mathbb{P}^N_\infty$ with $|f(t_i)| \le 1$ for all $i \in [\![1:N]\!]$ can be written as $f = \sum_{i=1}^N c_i \widetilde{L}_i$ with $-1 \le c_i \le 1$ for all

$i \in [\![1:N]\!]$, and the fact that $f \in \mathbb{P}^n_\infty$ means that f can also be written as $f = \sum_{j=1}^n c_{(j-1)\nu+1} L_j$. Since

$$L_j = \sum_{i=1}^N B_{i,j} \widetilde{L}_i, \qquad \mathbf{B} \in \mathbb{R}^{N \times n} \text{ having entries } B_{i,j} = L_j(t_i),$$

identifying the two expressions for f gives $c_i = \sum_{j=1}^n B_{i,j} c_{(j-1)\nu+1}$ for all $i \in [\![1:N]\!]$, or in short

$$\mathbf{c} = \widetilde{\mathbf{B}}\mathbf{c}, \qquad \widetilde{\mathbf{B}} \in \mathbb{R}^{N\times N} \text{created by inserting } \nu - 1 \text{ columns of 0's between any two columns of } \mathbf{B}.$$

In summary, the set Ω can be parametrized as

$$\Omega = \left\{ \sum_{j=1}^n c_{(j-1)\nu+1} L_j, \quad \mathbf{c} \in \mathbb{R}^N \text{ satisfying } \mathbf{c} + \mathbf{1} \geq \mathbf{0}, -\mathbf{c} + \mathbf{1} \geq \mathbf{0}, (\mathbf{I} - \widetilde{\mathbf{B}})\mathbf{c} = \mathbf{0} \right\}.$$

It is then convenient to choose the basis $\underline{\mathbf{v}}$ (hence also $\underline{\mathbf{w}}$) as the Lagrange basis $(L_1, \ldots, L_n)$, so that $\mathcal{P}(L_j) = \sum_{h=1}^n P_{h,j} L_h$. Thus, for $f = \sum_{j=1}^n c_{(j-1)\nu+1} L_j \in \Omega$ and for $i \in [\![1:N]\!]$, we have

$$\mathcal{P}(f)(t_i) = \left[\sum_{j=1}^n c_{(j-1)\nu+1} \left(\sum_{h=1}^n P_{h,j} L_h \right) \right](t_i) = \sum_{j=1}^n c_{(j-1)\nu+1} \sum_{h=1}^n B_{i,h} P_{h,j} = \mathbf{a}_i(P)^\top \mathbf{c},$$

where $\mathbf{a}_i(P) \in \mathbb{R}^N$ is the vector with nonzero entries

$$\mathbf{a}_i(P)_{(j-1)\nu+1} = \sum_{h=1}^n B_{i,h} P_{h,j} = (\mathbf{BP})_{i,j}, \qquad j \in [\![1:n]\!],$$

i.e., $\mathbf{a}_i(P)$ is the (transpose of) the ith row of the matrix $\widetilde{\mathbf{BP}} \in \mathbb{R}^{N\times N}$ defined by inserting $\nu - 1$ columns of 0's between any two columns of $\mathbf{BP} \in \mathbb{R}^{N\times n}$. At this point, the program (9) can be expressed as

$$\begin{aligned} \underset{d,P}{\text{minimize}}\ d \qquad & \text{subject to} \qquad (U^\top \otimes I_n)\text{vec}(P) = \text{vec}(U), \quad (\widetilde{U}^\top \otimes \widetilde{U}^\top)\text{vec}(P) = 0, \\ & \text{and to} \qquad |\mathbf{a}_i(P)^\top \mathbf{c}| \leq d \\ & \text{whenever} \qquad \mathbf{c} + \mathbf{1} \geq \mathbf{0}, -\mathbf{c} + \mathbf{1} \geq \mathbf{0}, (\mathbf{I} - \widetilde{\mathbf{B}})\mathbf{c} = \mathbf{0}, \quad \text{all } i \in [\![1:N]\!]. \end{aligned}$$

Following robust optimization techniques, the above condition for a given $i \in [\![1:N]\!]$ (where the absolute value can be discarded by invariance upon the change $\mathbf{c} \leftrightarrow -\mathbf{c}$) is written as

$$\left[\max_{\mathbf{c} \in \mathbb{R}^N} \mathbf{a}_i(P)^\top \mathbf{c} \quad \text{subject to } \mathbf{c} + \mathbf{1} \geq \mathbf{0}, -\mathbf{c} + \mathbf{1} \geq \mathbf{0}, (\mathbf{I} - \widetilde{\mathbf{B}})\mathbf{c} = \mathbf{0} \right] \leq d.$$

By classical duality theory (see e.g. [6, p.224-225]), this is equivalent to

$$\left[\min_{\mathbf{x},\mathbf{y},\mathbf{z} \in \mathbb{R}^N} \mathbf{1}^\top \mathbf{x} + \mathbf{1}^\top \mathbf{y} \quad \text{subject to } \mathbf{x} \geq \mathbf{0}, \mathbf{y} \geq \mathbf{0}, \mathbf{x} - \mathbf{y} + (\mathbf{I} - \widetilde{\mathbf{B}})^\top \mathbf{z} = -\mathbf{a}_i(P) \right] \leq d,$$

i.e., to the existence of $\mathbf{x}^i, \mathbf{y}^i, \mathbf{z}^i \in \mathbb{R}^N$ with $\mathbf{x}^i \geq \mathbf{0}$, $\mathbf{y}^i \geq \mathbf{0}$, and $\mathbf{x}^i - \mathbf{y}^i + (\mathbf{I} - \widetilde{\mathbf{B}})^\top \mathbf{z}^i = -\mathbf{a}_i(P)$ such that $\mathbf{1}^\top \mathbf{x}^i + \mathbf{1}^\top \mathbf{y}^i \leq d$. In conclusion, the program (9) is recast as the linear program

$$
\begin{aligned}
(11) \qquad \underset{d,P,\{\mathbf{x}^i,\mathbf{y}^i,\mathbf{z}^i\}_{i=1}^N}{\text{minimize}} \; d \qquad & \text{subject to} && (U^\top \otimes I_n)\mathrm{vec}(P) = \mathrm{vec}(U), \quad (\widetilde{U}^\top \otimes \widetilde{U}^\top)\mathrm{vec}(P) = 0,\\
& \text{to} && \mathbf{x}^i \geq \mathbf{0}, \mathbf{y}^i \geq \mathbf{0},\\
& \text{to} && \mathbf{x}^i - \mathbf{y}^i + (\mathbf{I} - \widetilde{\mathbf{B}})^\top \mathbf{z}^i = -\mathbf{a}_i(P),\\
& \text{and to} && \mathbf{1}^\top \mathbf{x}^i + \mathbf{1}^\top \mathbf{y}^i \leq d, \quad \text{all } i \in [\![1:N]\!].
\end{aligned}
$$

This is one of the programs implemented in the package MinProj. It is run by calling the command:

```
>>MinProjPoly(u,n,nu,'LP')
```

The subspace $\mathbb{U}$ of $\mathbb{V} = \mathbb{P}^n_\infty$ is not input via the matrix of a basis of $\mathbb{U}$ relative to a basis of $\mathbb{V}$ as done for coordinate and matrix spaces, but more conveniently as a polynomial basis of $\mathbb{U}$ listed in a MATLAB cell `u` whose elements are classic chebfuns (the objects manipulated by the package Chebfun). In this case, the dimension `n` of the polynomial superspace also needs to be specified. The input `nu` represents the discretization index ν introduced earlier and `'LP'` indicates that the user chooses to solve linear program (11). The output is a couple of values bounding the projection constant of $\mathbb{U}$ in $\mathbb{P}^n_\infty$ from below and above, respectively. An optional output can also be returned, namely the pseudo minimal projection that solves the approximate program (11). Although the latter has the appeal of being a linear program, we find the semidefinite program described below more efficient.

5.3 The second approximate program as a semidefinite programming

The set Ω corresponding to the program (10) is initially given by

$$\Omega = \{f \in \mathbb{P}^n_\infty : \|f\|_\infty \leq 1\}.$$

Note that the condition $\|f\|_\infty \leq 1$ is equivalent to the two conditions $1 - f(x) \geq 0$ for all $x \in [-1, 1]$ and $1 + f(x) \geq 0$ for all $x \in [-1, 1]$. In turn, expanding f on the basis of Chebyshev polynomials of the first kind as $f = \sum_{k=1}^n c_k T_{k-1}$, these two nonnegativity conditions are equivalent (see e.g. [12, Theorem 3]) to the existence of positive semidefinite matrices $\mathbf{Q}^-, \mathbf{Q}^+ \in \mathbb{R}^{n \times n}$ such that

$$\sum_{i-j=k-1} Q^{\pm}_{i,j} = \begin{Bmatrix} \pm \dfrac{c_k}{2}, & k \in [\![2:n]\!], \\ 1 \pm c_1, & k = 1. \end{Bmatrix}.$$

This is equivalent to

$$\mathrm{tr}(\mathbf{J}_k(\mathbf{Q}^+ - \mathbf{Q}^-)) = c_k \qquad \text{and} \qquad \mathrm{tr}(\mathbf{J}_k(\mathbf{Q}^+ + \mathbf{Q}^-)) = \delta_{1,k},$$

where $\mathbf{J}_1$ is half the identity matrix (it has $1/2$'s on the main (first) diagonal and 0's everywhere else), and for $k \in [\![2:n]\!]$, $\mathbf{J}_k$ is the symmetric matrix with $1/2$'s on the kth subdiagonal and kth superdiagonal and 0's everywhere else. Thus, the set Ω can be parametrized as

$$\Omega = \Bigg\{ \sum_{k=1}^{n} \mathrm{tr}(\mathbf{J}_k(\mathbf{Q}^+ - \mathbf{Q}^-))T_{k-1}, \quad \mathbf{Q}^-, \mathbf{Q}^+ \in \mathbb{R}^{n\times n} \text{ satisfying } \mathbf{Q}^- \succcurlyeq \mathbf{0}, \mathbf{Q}^+ \succcurlyeq \mathbf{0}, \\ \text{and } \mathrm{tr}(\mathbf{J}_k(\mathbf{Q}^+ + \mathbf{Q}^-)) = \delta_{1,k}, \quad k \in [\![1:n]\!] \Bigg\}.$$

It is then convenient to choose the basis $\underline{\mathbf{v}}$ (hence also $\underline{\mathbf{w}}$) as the Chebyshev basis $(T_0, \ldots, T_{n-1})$, so that $\mathcal{P}(T_{k-1}) = \sum_{h=1}^n P_{h,k} T_{h-1}$. Thus, for $f = \sum_{k=1}^n \mathrm{tr}(\mathbf{J}_k(\mathbf{Q}^+ - \mathbf{Q}^-))T_{k-1}$ and for $i \in [\![1:N]\!]$, we have

$$\mathcal{P}(f)(t_i) = \left[\sum_{k=1}^{n} \mathrm{tr}(\mathbf{J}_k(\mathbf{Q}^+ - \mathbf{Q}^-)) \left(\sum_{h=1}^{n} P_{h,k} T_{h-1} \right) \right](t_i) = \mathrm{tr}(\mathbf{A}_i(P)(\mathbf{Q}^+ - \mathbf{Q}^-)),$$

where $\mathbf{A}_i(P) \in \mathbb{R}^{n\times n}$ is the symmetric Toeplitz matrix depending linearly on P defined by

$$\mathbf{A}_i(P) = \sum_{k=1}^{n} \sum_{h=1}^{n} P_{h,k} T_{h-1}(t_i) \mathbf{J}_k.$$

The program (10) can now be expressed as

$$\begin{aligned} \underset{d,P}{\text{minimize}}\; d \quad & \text{subject to} \quad (U^\top \otimes I_n)\mathrm{vec}(P) = \mathrm{vec}(U), \quad (\widetilde{U}^\top \otimes \widetilde{U}^\top)\mathrm{vec}(P) = 0, \\ & \text{and to} \quad |\mathrm{tr}(\mathbf{A}_i(P)(\mathbf{Q}^+ - \mathbf{Q}^-))| \le d \\ & \text{whenever} \quad \mathbf{Q}^- \succcurlyeq \mathbf{0}, \mathbf{Q}^+ \succcurlyeq \mathbf{0}, \mathrm{tr}(\mathbf{J}_k(\mathbf{Q}^+ + \mathbf{Q}^-)) = \delta_{1,k},\ k \in [\![1:n]\!], \quad \text{all } i \in [\![1:N]\!]. \end{aligned}$$

Following robust optimization techniques, the above condition for a given $i \in [\![1:N]\!]$ (where the absolute value can be discarded by invariance upon the change $\mathbf{Q}^\pm \leftrightarrow \mathbf{Q}^\mp$) is written as

$$\Bigg[\max_{\mathbf{Q}^-, \mathbf{Q}^+ \in \mathbb{R}^{n\times n}} \mathrm{tr}(\mathbf{A}_i(P)(\mathbf{Q}^+ - \mathbf{Q}^-)) \quad \text{subject to } \mathbf{Q}^- \succcurlyeq \mathbf{0}, \mathbf{Q}^+ \succcurlyeq \mathbf{0}, \\ \text{and to } \mathrm{tr}(\mathbf{J}_k(\mathbf{Q}^+ + \mathbf{Q}^-)) = \delta_{1,k},\ k \in [\![1:n]\!] \Bigg] \le d.$$

By classical duality theory (see e.g. [6, p.264-265]), this is equivalent to

$$\left[\min_{\mathbf{z} \in \mathbb{R}^n} z_1 \quad \text{subject to } \sum_{k=1}^n z_k \mathbf{J}_k \succcurlyeq -\mathbf{A}_i(P),\ \sum_{k=1}^n z_k \mathbf{J}_k \succcurlyeq \mathbf{A}_i(P) \right] \le d,$$

i.e., to the existence of $\mathbf{z}^i \in \mathbb{R}^n$ with $\sum_{k=1}^n z_k^i \mathbf{J}_k \succcurlyeq -\mathbf{A}_i(P)$ and $\sum_{k=1}^n z_k^i \mathbf{J}_k \succcurlyeq \mathbf{A}_i(P)$ such that $z_1^i \le d$. In conclusion, the program (10) is recast as the semidefinite program

$$\begin{aligned} (12) \qquad \underset{d,P,\{\mathbf{z}^i\}_{i=1}^N}{\text{minimize}}\; d \qquad & \text{subject to} \qquad (U^\top \otimes I_n)\mathrm{vec}(P) = \mathrm{vec}(U), \quad (\widetilde{U}^\top \otimes \widetilde{U}^\top)\mathrm{vec}(P) = 0, \\ & \text{to} \qquad \sum_{k=1}^n z_k^i \mathbf{J}_k \succcurlyeq -\mathbf{A}_i(P), \quad \sum_{k=1}^n z_k^i \mathbf{J}_k \succcurlyeq \mathbf{A}_i(P), \\ & \text{and to} \qquad z_1^i \le d, \quad \text{all } i \in [\![1:N]\!]. \end{aligned}$$

This is the other program implemented in the package MinProj to compute projection constants in polynomial spaces. It is run by calling the command

```
>>MinProjPoly(u,n,nu,'SDP')
```

As before, the input `u` is a cell listing a polynomial basis for the subspace $\mathbb{U}$ as classic chebfuns, `n` is the dimension of the polynomial superspace, `nu` is the discretization parameter, and `'SDP'` indicates that the user chooses to solve the semidefinite program (12). A lower and upper bound for the projection constant are output by default, and the pseudo minimal projection solving (12) can also be returned optionally.

5.4 Computations of some projection constants

In this subsection, we exploit MinProjPoly to compute some lower and upper bounds for the projection constants of some polynomial subspaces. The outcomes can be duplicated by running the reproducible file of this article. We start by considering the 2-codimensional subspace $\mathbb{V}_{0,0}^{n-2}$ of $\mathbb{P}_\infty^n$ consisting of polynomials vanishing at the endpoints -1 and 1. The solutions of the semidefinite program from $n = 3$ to $n = 11$ are summarized in Table 2.

Table 2: Projection constants for polynomials vanishing at the endpoints (index $\nu = 200$).

$\lambda(\mathbb{V}_{0,0}^{n-2}, \mathbb{P}_\infty^n)$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$	$n=11$
Lower bound	1.0000	1.1106	1.1358	1.1390	1.1374	1.1349	1.1323	1.1299	1.1277
Upper bound	1.0079	1.1194	1.1448	1.1480	1.1464	1.1439	1.1412	1.1388	1.1366

We continue by tabulating the values of lower and upper bounds for the projection constants of $\mathbb{P}_\infty^m$ interpreted as a subspace of $\mathbb{P}_\infty^n$ for $2 \le m < n \le 11$, see Table 3. It is instructive to compare these values with the bounds for the absolute projection constants $\lambda(\mathbb{P}_\infty^m) = \max_{n>m} \lambda(\mathbb{P}_\infty^m, \mathbb{P}_\infty^n)$ that were tabulated up to $m = 11$ in [13] (the exact value of $\lambda(\mathbb{P}_\infty^m)$ is only known for $m = 3$, see [8]). It is also interesting to verify the statement of [15] that the minimal projection P^* from $\mathbb{P}_\infty^{m+1}$ to $\mathbb{P}_\infty^m$ is an interpolating projection. Note that checking that there are indeed m points $x_1, \ldots, x_m \in [-1, 1]$ such that $P^*(f)(x_i) = f(x_i)$ for all $f \in \mathbb{P}_\infty^{m+1}$ reduces to verifying the fact that $P^*(T_m) - T_m$ has m zeros in $[-1, 1]$. For $m = 5$, say, this fact is substantiated by Figure 2, which plots $P(T_m) - T_m$ where P approximates the minimal projection P^* as the projection output by MinProjPoly.

6 Minimal extensions from coordinates spaces

We conclude the article by showing how minimal extensions can also be computed within the framework we have presented. Since this is not the main topic of this work, the procedure has not

Table 3: Projection constants for m-dimensional polynomial spaces in n-dimensional polynomial spaces (index $\nu = 200$). Each box contains an lower bound (top) and an upper bound (bottom).

$\lambda(\mathbb{P}^m_\infty, \mathbb{P}^n_\infty)$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$	$m=8$	$m=9$	$m=10$
$n=3$	1.0000 1.0079								
$n=4$	1.0000 1.0079	1.1992 1.2087							
$n=5$	1.0000 1.0079	1.1997 1.2092	1.3195 1.3300						
$n=6$	1.0000 1.0079	1.2125 1.2221	1.3195 1.3300	1.4063 1.4174					
$n=7$	1.0000 1.0079	1.2134 1.2230	1.3331 1.3437	1.4063 1.4174	1.4726 1.4842				
$n=8$	1.0000 1.0079	1.2153 1.2249	1.3347 1.3453	1.4168 1.4280	1.4726 1.4843	1.5251 1.5372			
$n=9$	1.0000 1.0079	1.2156 1.2252	1.3360 1.3466	1.4170 1.4282	1.4799 1.4916	1.5252 1.5373	1.5679 1.5803		
$n=10$	1.0000 1.0079	1.2163 1.2259	1.3368 1.3474	1.4208 1.4320	1.4799 1.4916	1.5306 1.5427	1.5688 1.5812	1.6034 1.6161	
$n=11$	1.0000 1.0079	1.2167 1.2263	1.3380 1.3486	1.4225 1.4337	1.4837 1.4954	1.5306 1.5427	1.5738 1.5863	1.6061 1.6188	1.6335 1.6464

been implemented in the package MinProj, but the following example can nonetheless be duplicated from the reproducible file. We focus on the case $\mathbb{V} = \ell_\infty^n$, although any space with a polyhedral norm can in principle be handled in the same way, e.g. $\mathbb{V} = \ell_1^n$, $\mathcal{M}^{k\times k}_{\infty\to\infty}$, or $\mathcal{M}^{k\times k}_{1\to 1}$. We also require the norm on $\mathbb{W}$ to be computable, either explicitly or via a semidefinite relaxation. The latter is the situation encountered below where $\mathbb{W} = \mathbb{P}^\ell_\infty$ is the space of polynomials of degree at most $\ell - 1$ equipped with the max-norm on $[-1,1]$. Taking $n > \ell$, we set $\mathbb{V} = \ell^n_\infty$ and we consider the subspace $\mathbb{U}$ of $\mathbb{V}$ given by

$$\mathbb{U} = \Big\{ [P(t_1), ; \ldots ; P(t_n)], P \in \mathbb{P}^\ell_\infty \Big\}$$

relatively to an arbitrary choice of points $t_1 < \cdots < t_n \in [-1,1]$. With $(L_1, \ldots, L_n)$ denoting the Lagrange basis of $\mathbb{P}^n_\infty$ at these points, we define the linear map $\mathcal{A} : \mathbb{U} \to \mathbb{P}^\ell_\infty$ by

$$\mathcal{A} : [c_1; \ldots ; c_n] \mapsto \sum_{j=1}^n c_j L_j. \tag{13}$$

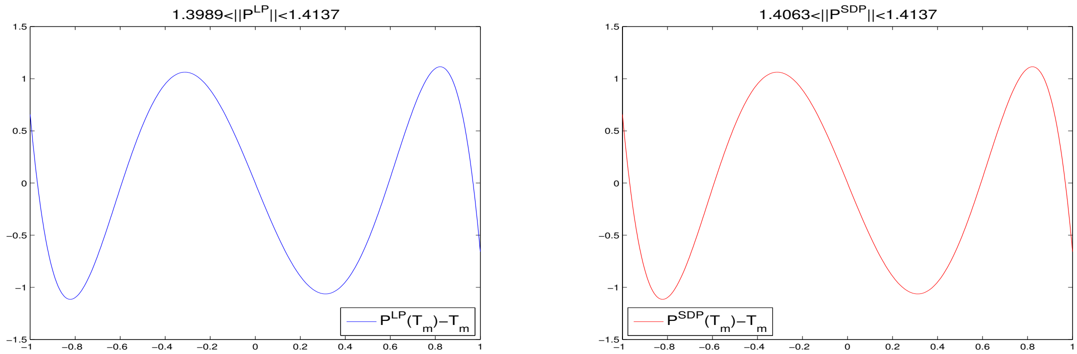

Figure 2: The pseudo minimal projections from $\mathbb{P}^6_\infty$ to $\mathbb{P}^5_\infty$ that solve programs (11) (left) and (12) (right) are interpolating projections (index $\nu = 300$).

The map $\mathcal{A}$ has values in $\mathbb{P}^\ell_\infty$ by virtue of the identity $P = \sum_{j=1}^n P(t_j) L_j$ for all $P \in \mathbb{P}^n_\infty$, and in particular for all $P \in \mathbb{P}^\ell_\infty$. As usual, the problem of minimal extension of $\mathcal{A}$ is transformed into

$$\underset{d,\mathcal{P}}{\text{minimize}}\ d \qquad \text{subject to} \qquad \mathcal{P} \text{ is an extension of } \mathcal{A}$$
$$\text{and} \qquad \|\mathcal{P}\boldsymbol{\varepsilon}\|_\infty \le d \quad \text{for all } \boldsymbol{\varepsilon} \in \{\pm 1\}^n.$$

Choosing $\underline{\mathbf{u}} = \{[T_{i-1}(t_1); \ldots; T_{i-1}(t_n)], i \in [\![1 : \ell]\!]\}$ as the basis for $\mathbb{U}$ and $\underline{\mathbf{w}} = \{T_{j-1}, j \in [\![1 : \ell]\!]\}$ as the basis for $\mathbb{W}$, the map $\mathcal{A}$ is clearly represented by the matrix $A = I_\ell$. Therefore, with $U = [\mathbf{u}_1 | \cdots | \mathbf{u}_\ell] \in \mathbb{R}^{n \times \ell}$, the constraint (i') reads

$$(U^\top \otimes I_\ell)\mathrm{vec}(P) = \mathrm{vec}(I_\ell),$$

while the constraint (ii') is absent in view of $\mathrm{ran}(\mathcal{A}) = \mathbb{P}^\ell_\infty$. The other constraints are, for all $\boldsymbol{\varepsilon} \in \{\pm 1\}^n$,

$$|\mathcal{P}(\boldsymbol{\varepsilon})(x)| = \left| \sum_{j=1}^n \varepsilon_j \left(\sum_{i=1}^\ell P_{i,j} T_{i-1} \right)(x) \right| \le d \qquad \text{for all } x \in [-1, 1].$$

After eliminating the absolute value in view of invariance up to the change $\boldsymbol{\varepsilon} \leftrightarrow -\boldsymbol{\varepsilon}$, they reduce to the 2^n constraints

$$\sum_{i=1}^\ell \left(\sum_{j=1}^n \varepsilon_j P_{i,j} \right) T_{i-1}(x) \ge -d \qquad \text{for all } x \in [-1, 1].$$

Each of these constraints, as a nonnegativity condition for a polynomial, admits a semidefinite relaxation (note, however, that the exponential number of them makes the approach impracticable for moderately large n). Precisely, according to [12, Theorem 3], the constraint corresponding to a fixed $\boldsymbol{\varepsilon} \in \{\pm 1\}$ is equivalent to the existence of a positive semidefinite matrix $\mathbf{Q}^{\boldsymbol{\varepsilon}} \in \mathbb{R}^{\ell \times \ell}$ such that, for any $k \in [\![1 : \ell]\!]$,

$$\sum_{i-j=k-1} Q^{\varepsilon}_{i,j} = \begin{cases} d + \sum_{j=1}^{n} P_{1,j}\varepsilon_j, & k = 1, \\ \dfrac{\sum_{j=1}^{n} P_{k,j}\varepsilon_j}{2}, & k > 1. \end{cases} \tag{14}$$

All is all, a minimal extension of $\mathcal{A}$ is determined by solving the semidefinite program

$$\underset{d,P,\{\mathbf{Q}^{\varepsilon}\}_{\boldsymbol{\varepsilon}}}{\text{minimize}}\; d \qquad \text{subject to} \qquad (U^{\top} \otimes I_\ell)\mathrm{vec}(P) = \mathrm{vec}(I_\ell)$$

$$\text{and to} \qquad \mathbf{Q}^{\boldsymbol{\varepsilon}} \succcurlyeq \mathbf{0} \text{ together with (14)}, \quad \text{all } \boldsymbol{\varepsilon} \in \{\pm 1\}^n.$$

The extension constants of $\mathcal{A}$ are given for different values n in Table 4. The first row was computed for $\ell = 3$ with $t_1 < \cdots < t_n$ chosen as the locations of the extrema of the Legendre polynomial P_{n-1}, while the second row was computed for $\ell = 4$ with $t_1 < \cdots < t_n$ chosen as the locations of the extrema of the Chebyshev polynomial T_{n-1}. We point out that each of these extension constants is an upper bound for the absolute projection constant $\lambda(\mathbb{P}^{\ell}_{\infty}) = \lambda(\mathbb{P}^{\ell}_{\infty}, \mathcal{C}[-1,1])$ of the space $\mathbb{P}^{\ell}_{\infty}$. Indeed, with $\mathcal{P} : \ell^n_\infty \to \mathbb{P}^{\ell}_{\infty}$ denoting a minimal extension of $\mathcal{A}$ and with $\mathcal{D} : \mathcal{C}[-1,1] \to \ell^n_\infty$ defined by $\mathcal{D}(f) = [f(t_1); \cdots; f(t_n)]$, it is not hard to see that $\mathcal{PD}$ is a projection from $\mathcal{C}[-1,1]$ onto $\mathbb{P}^{\ell}_{\infty}$ and that

$$\|\mathcal{PD}f\|_\infty \le \|\!|\mathcal{P}|\!\| \|\mathcal{D}f\|_\infty \le \|\!|\mathcal{P}|\!\| \|f\|_\infty \qquad \text{for all } f \in \mathcal{C}[-1,1],$$

so that $\lambda(\mathbb{P}^{\ell}_{\infty}) \le \|\!|\mathcal{PD}|\!\| \le \|\!|\mathcal{P}|\!\|$. In particular, the upper bound 1.2206 for $\ell = 3$ is quite close to the true value ≈ 1.2201 obtained in [8] for the absolute projection constant of the quadratics and the upper bound 1.3653 for $\ell = 4$ is not too far either from the range $[1.3539, 1.3576]$ obtained in [13] for the absolute projection constant of the cubics.

Table 4: Extension constants of the map (13) with Legendre ($\ell = 3$) and Chebyshev ($\ell = 4$) points.

	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$	$n=11$	$n=12$	$n=13$	$n=14$
$\ell=3$	1.2282	1.2736	1.2238	1.2479	1.2221	1.2374	1.2211	1.2320	1.2206	1.2288
$\ell=4$	1.7370	1.5106	1.4098	1.4183	1.3878	1.3840	1.3789	1.3689	1.3744	1.3653

References

[1] CVX Research, Inc. CVX: MATLAB software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`, 2014.

[2] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization.* Princeton University Press, 2009.

[3] D. Bertsimas, D. B. Brown, and C. Caramanis. *Theory and applications of robust optimization.* SIAM review, 53, 464–501, 2011.

[4] R. Bhatia. *Matrix analysis.* Springer, 1997.

[5] J. Blatter and E. W. Cheney. *Minimal projections on hyperplanes in sequence spaces.* Annali di Matematica Pura ed Applicata, 101, 215–227, 1974.

[6] S. Boyd and L. Vandenberghe. *Convex optimization.* Cambridge University Press, 2009.

[7] B. L. Chalmers and G. Lewicki. *Symmetric spaces with maximal projection constants.* Journal of Functional Analysis, 200, 1–22, 2003.

[8] B. L. Chalmers and F. T. Metcalf. *Determination of a minimal projection from* $C[-1,1]$ *onto the quadratics.* Numerical Functional Analysis and Optimization, 11, 1–10, 1990.

[9] B. L. Chalmers and B. Shekhtman. *Minimal projections and absolute projection constants for regular polyhedral spaces.* Proceedings of the American Mathematical Society, 95, 449–452, 1985.

[10] E. W. Cheney and K. H. Price. *Minimal projections.* In: Approximation Theory (1970), 261–289.

[11] T. A. Driscoll, N. Hale, and L. N. Trefethen, editors. Chebfun *Guide.* Pafnuty Publications, Oxford, 2014.

[12] S. Foucart and V. Powers. Basc*: constrained approximation by semidefinite programming.* Submitted, 2014.

[13] C. Helzel and K. Petras. *Numerical estimation of projection constants.* Numerical Functional Analysis and Optimization, 18, 555–566, 1997.

[14] G. Lewicki. *Minimal Projections onto two dimensional subspaces of* $\ell_\infty^{(4)}$. Journal of Approximation Theory, 88, 92–108, 1997.

[15] G. Lewicki and M. Prophet. *Codimension-one minimal projections onto Haar subspaces.* Journal of Approximation Theory, 127, 198–206, 2004.

[16] G. Lewicki and L. Skrzypek. *Chalmers–Metcalf operator and uniqueness of minimal projections.* Journal of Approximation Theory, 148, 71–91, 2007.

[17] B. Shekhtman and L. Skrzypek. *Uniqueness of minimal projections onto two-dimensional subspaces.* Studia Mathematica, 168, 273–284, 2005.

[18] L. Vandenberghe and S. Boyd. *Semidefinite programming.* SIAM review, 38, 49–95, 1996.